A NEW UPPER BOUND FOR THE GROWTH FACTOR IN GAUSSIAN ELIMINATION WITH COMPLETE PIVOTING

ANKIT BISAIN, ALAN EDELMAN, AND JOHN URSCHEL

Abstract. The growth factor in Gaussian elimination measures how large the entries of an LU factorization can be relative to the entries of the original matrix. It is a key parameter in error estimates, and one of the most fundamental topics in numerical analysis. We produce an upper bound of $n^{0.2079 \ln n + 0.91}$ for the growth factor in Gaussian elimination with complete pivoting – the first improvement upon Wilkinson's original 1961 bound of $2\, n^{0.25 \ln n + 0.5}$.

1. Introduction

The solution of a linear system $Ax = b$ is one of the oldest problems in mathematics. One of the most fundamental and important techniques for solving a linear system is Gaussian elimination, in which a matrix is factored into the product of a lower and upper triangular matrix. Given an $n \times n$ matrix A, Gaussian elimination performs a sequence of rank-one transformations, resulting in the sequence of matrices $A^{(k)} \in \mathbb{C}^{k \times k}$ for k equals n to 1, satisfying

$$A^{(k)} = M^{(2,2)} - M^{(2,1)}[M^{(1,1)}]^{-1}M^{(1,2)}, \qquad \text{where } A = \begin{array}{cc} \begin{array}{cc} \scriptstyle n-k & \scriptstyle k \end{array} & \\ \begin{bmatrix} M^{(1,1)} & M^{(1,2)} \\ M^{(2,1)} & M^{(2,2)} \end{bmatrix} & \begin{array}{c} \scriptstyle n-k \\ \scriptstyle k \end{array} \end{array} .$$

The resulting LU factorization of A is encoded by the first row and column of each of the iterates $A^{(k)}$, $k = 1, ..., n$. Not all matrices have an LU factorization, and a permutation of the rows (or columns) of the matrix may be required. In addition, performing computations in finite precision can elicit issues due to round-off error. The error due to rounding in Gaussian elimination for a matrix A in some fixed precision is controlled by the growth factor of the Gaussian elimination algorithm, defined by

$$g(A) := \frac{\max_k |A^{(k)}|_\infty}{|A|_\infty},$$

where $|\cdot|_\infty$ is the entry-wise matrix infinity norm (see [8, Theorem 3.3.1] for details). For this reason, understanding the growth factor is of both theoretical and practical importance. Complete pivoting, famously referred to as "customary" by von Neumann [19], is a strategy for permuting the rows and columns of A so that, at each step, the pivot (the top-left entry of $A^{(k)}$) is the largest magnitude entry of $A^{(k)}$. Complete pivoting remains the premier theoretical permutation strategy for performing Gaussian elimination. Despite its popularity, the worst-case behavior of the growth factor under complete pivoting is poorly understood.

1.1. **Historical Overview and Relevant Results.** The appearance of the computer in the aftermath of the Second World War created a new branch of mathematics now known as numerical analysis. In their seminal 1947 paper *Numerical Inverting of Matrices of High Order*, von Neumann and Goldstine studied the stability of Gaussian elimination with complete pivoting [19]. This work was motivated by their development of the first stored-program digital

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, 02139 USA.
E-mail addresses: ankitb12@mit.edu, edelman@mit.edu, urschel@mit.edu.
2020 *Mathematics Subject Classification.* Primary 65F05, 15A23.

computer and desire to understand the effect of rounding in computations on it [13]. Goldstine later wrote:

> Indeed, von Neumann and I chose this topic for the first modern paper on numerical analysis ever written precisely because we viewed the topic as being absolutely basic to numerical mathematics [7].

However, it was not until Wilkinson's 1961 paper *Error Analysis of Direct Methods of Matrix Inversion* that a more rigorous analysis of the backward error in Gaussian elimination due to rounding errors occurred. Indeed, Wilkinson was the first to fully recognize the dependence of this error on the growth factor. Let $g_n(\mathbb{R})$ and $g_n(\mathbb{C})$ denote the maximum growth factor under complete pivoting over all non-singular $n \times n$ real and complex matrices, respectively. Wilkinson produced a bound for the growth factor under complete pivoting using only Hadamard's inequality [20, Equation 4.15]:

$$g_n(\mathbb{C}) \leq \sqrt{n}\big(2\ 3^{1/2} \ldots n^{1/(n-1)}\big)^{1/2} \leq 2\sqrt{n}\, n^{\ln(n)/4}, \tag{1.1}$$

where the second inequality is asymptotically tight. This estimate was considered extremely pessimistic, with Wilkinson himself noting that "no matrix has been encountered for which [the growth factor for complete pivoting] was as large as 8 [20]." A conjecture that the growth factor for complete pivoting of a real $n \times n$ matrix was at most n was eventually formed.[1] Many researchers attempted to upper bound the growth factor, with $g_n(\mathbb{R})$ computed exactly for $n = 1, 2, 3, 4$ and shown to be strictly less than five for $n = 5$ (see the works of Tornheim [15, 16, 17, 18], Cryer [4], and Cohen [3] for details). However, no progress was made on improving the bound for arbitrary n. Many years later, in 1991, Gould found a 13×13 matrix with growth factor larger than 13 in finite precision [9] (extended to exact arithmetic by Edelman [5]), providing a counterexample to the conjecture for $n = 13$. Recently, Edelman and Urschel improved the best-known lower bounds for all $n > 8$ and showed that

$$g_n(\mathbb{R}) \geq 1.0045\, n \ \text{ for all } n \geq 11, \qquad \text{and} \qquad \limsup_n \big(g_n(\mathbb{R})/n\big) \geq 3.317,$$

thus disproving the aforementioned conjecture for all $n \geq 11$ by a multiplicative factor [6]. However, for the upper bound, to date no improvement has been made to Wilkinson's bound.

1.2. **Our Contributions.** In this work, we improve Wilkinson's upper bound by an exponential constant, the first improvement in over sixty years. In particular, we prove the following theorem, obtaining a leading exponential constant of $\frac{1}{2[2+(2-\sqrt{2})\ln 2]} \approx 0.20781$.

Theorem 1.1. $g_n(\mathbb{C}) \leq n^{\frac{\ln n}{2[2+(2-\sqrt{2})\ln 2]}+0.91}$.

Our proof consists of four parts:

1. A Generalized Hadamard's inequality: We prove a tighter version of Hadamard's famous inequality for matrices with a large low-rank component. This generalization allows for a more sophisticated analysis of the iterates of Gaussian elimination, providing additional constraints on the pivots of a matrix. (Subsection 3.1)
2. An Improved Optimization Problem: Applying the improved determinant bounds produces an optimization problem that can be considered a refinement of the optimization problem associated with Wilkinson's proof. Unfortunately, this refinement is no longer linear upon a logarithmic transformation. (Subsection 3.2)
3. From Non-Linear to Linear: We relax the logarithmic transformation of our optimization problem to a linear program, and prove that the optimal value of our relaxation has the same asymptotic behavior. (Subsection 3.3)

[1] See [6, Section 1.1] for a detailed discussion of the conjecture and its possible misattribution to both Cryer and Wilkinson.

(4) An Asymptotic Analysis: Finally, we analyze the asymptotic behavior of our linear program by converting it into a continuous program and applying a duality argument, thus producing the improved bound in Theorem 1.1. (Section 4)

Our proof considers the same information regarding the underlying matrix as Wilkinson's original bound, using only the pivots at each step of elimination, and reveals further structure regarding the relationships between them. Improved estimates on the explicit constants in Theorem 1.1 can be obtained through a refinement of the techniques presented herein. However, tight estimates on the maximum growth factor will likely require further information regarding matrix entries.

Finally, we note that the analysis associated with the proof of Theorem 1.1 improves worst-case estimates for small n (see Figure 1) and that the upper bound of Theorem 1.1 (a pessimistic asymptotic bound) is even superior to Wilkinson's bound for practical values of n. For instance, for $n = 24219648$, the largest dimension for which an LU factorization has been performed on a general dense matrix as of November 2023 [14], Theorem 1.1 gives a 99% improvement over Wilkinson's bound. Further details regarding the practical computation of upper bounds for finite n are given in Subsection 3.4.

2. Wilkinson's Bound Viewed as a Linear Program

The proof of Wilkinson's 1961 bound is incredibly short, requiring one page of mathematics and using only Hadamard's inequality applied to the matrix iterates $A^{(k)} \in \mathbb{C}^{k\times k}$ of Gaussian elimination. Letting $p_k = |A^{(k)}|_\infty$, $k = 1, ..., n$, denote the pivots of Gaussian elimination, Hadamard's inequality implies that

$$\prod_{i=1}^{k} p_i = \det(A^{(k)}) \le k^{k/2} |A^{(k)}|_\infty^k = k^{k/2} p_k^k. \tag{2.1}$$

The maximum growth factor is a non-decreasing function in n, and so the maximum value of p_1/p_n under these constraints provides an upper bound for the maximum growth factor:

Wilkinson's Optimization Problem

$$\begin{array}{ll} \max & p_1/p_n \\ \text{s.t.} & \prod_{i=1}^k p_i \le k^{k/2} p_k^k \quad \text{for } k = 1, ..., n. \end{array} \tag{2.2}$$

Performing the transformation $q_k = \ln(p_k)$ for $k = 1, ..., n$ produces the linear program:

Wilkinson's Linear Program

$$\begin{array}{ll} \max & q_1 - q_n \\ \text{s.t.} & \sum_{i=1}^k q_i \le \frac{k}{2} \ln k + k\, q_k \quad \text{for } k = 1, ..., n. \end{array} \tag{2.3}$$

Wilkinson's proof, though never stated in the context of linear programming, can be viewed as a simple LP duality argument. The inequalities can be written in matrix form as

$$Ax = \begin{pmatrix} 1 & & & & \\ 1 & -1 & & & \\ 1 & 1 & -2 & & \\ \vdots & \vdots & \ddots & \ddots & \\ 1 & 1 & \cdots & 1 & -(n-1) \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ \vdots \\ q_n \end{pmatrix} \le \begin{pmatrix} 0 \\ \frac{2}{2}\ln 2 \\ \frac{3}{2}\ln 3 \\ \vdots \\ \frac{n}{2}\ln n \end{pmatrix} = b,$$

where the additional constraint $q_1 \le 0$ plays no role, as the feasible region of Program 2.3 is shift-independent. The matrix A has an easily computable inverse with $A^{-1}_{i1} = 1$ for $i = 1, ..., n$, $A^{-1}_{ii} = -\frac{1}{i-1}$ for $i = 2, ..., n$, and $A^{-1}_{ij} = -\frac{1}{j(j-1)}$ for $i > j$. The quantity

$$[A^{-1}]^T c = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ & -1 & -\frac{1}{2} & \cdots & -\frac{1}{2} \\ & & -\frac{1}{2} & & \vdots \\ & & & \ddots & -\frac{1}{(n-2)(n-1)} \\ & & & & -\frac{1}{n-1} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{2} \\ \vdots \\ \frac{1}{(n-2)(n-1)} \\ \frac{1}{n-1} \end{pmatrix}$$

is entry-wise non-negative, implying Wilkinson's bound

$$q_1 - q_n = \left([A^{-1}]^T c\right)^T A x \le \left([A^{-1}]^T c\right)^T b = \frac{1}{2}\left[\ln n + \sum_{k=2}^{n} \frac{\ln k}{k-1}\right].$$

This bound is the exact solution to Program 2.3, evidenced by the matching feasible point $x = A^{-1} b$. The ease with which the optimal point of the dual program can be obtained is due to the simple structure of the constraints. Our improved linear program, described in Subsection 3.3, has a more complicated set of constraints, requiring a more complex duality argument (given in Section 4).

This same argument also immediately produces bounds for the geometric mean growth factor of the iterates $A^{(k)}$, a key quantity in our proof of Theorem 1.1 that may be of independent interest. Indeed, the quantity $\frac{1}{n}\sum_{k=1}^{n}(q_1 - q_k)$ can be upper bounded by analyzing the linear program:

Geometric Mean Growth LP

$$\begin{aligned} \max \quad & \tfrac{1}{n}\textstyle\sum_{k=1}^{n}(q_1 - q_k) \\ \text{s.t.} \quad & \textstyle\sum_{i=1}^{k} q_i \le \tfrac{k}{2}\ln k + k\, q_k \quad \text{for } k = 1, ..., n. \end{aligned} \tag{2.4}$$

The constraints of this linear program are identical to those of Program 2.3. The only difference is in the objective; here we have $c = \left(\frac{n-1}{n}, -\frac{1}{n}, ..., -\frac{1}{n}\right)^T$. Nevertheless, the quantity

$$[A^{-1}]^T c = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ & -1 & -\frac{1}{2} & \cdots & -\frac{1}{2} \\ & & -\frac{1}{2} & & \vdots \\ & & & \ddots & -\frac{1}{(n-2)(n-1)} \\ & & & & -\frac{1}{n-1} \end{pmatrix} \begin{pmatrix} \frac{n-1}{n} \\ -\frac{1}{n} \\ \vdots \\ -\frac{1}{n} \\ -\frac{1}{n} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{2} \\ \vdots \\ \frac{1}{(n-2)(n-1)} \\ \frac{1}{(n-1)n} \end{pmatrix}$$

is entry-wise non-negative, implying the bound

$$\frac{1}{n}\sum_{k=1}^{n}(q_1 - q_k) = \left([A^{-1}]^T c\right)^T A x \le \left([A^{-1}]^T c\right)^T b = \frac{1}{2}\sum_{k=2}^{n}\frac{\ln k}{k-1} \le \frac{\ln^2 n}{4} + \ln 2, \tag{2.5}$$

which can be easily generalized further to any weighted average $\sum_{k=1}^{n} w_k (q_1 - q_k)$ of the logarithmic growth factors.

3. An Improved Linear Program

In this section, we produce additional constraints that the pivots must satisfy by generalizing Hadamard's inequality for matrices with a large low-rank component. These constraints, applied to the matrix $A^{(k)}$ (viewed as a sub-matrix of $A^{(k+\ell)}$ plus a rank ℓ matrix), lead to a new linear program with optimal value at most $0.2079 \ln^2 n + O(\ln n)$, the first improvement to the exponential constant of 0.25 in Wilkinson's bound (Inequality 1.1).

3.1. **Improved Determinant Bounds.** First, we recall the following basic proposition, itself a corollary of [11, Theorem 1].[2]

Proposition 3.1. *$|\det(A+B)| \le \prod_{i=1}^n \big(\sigma_i(A) + \sigma_{n-i+1}(B)\big)$ for all $A, B \in \mathbb{C}^{n\times n}$, where $\sigma_1(A) \ge ... \ge \sigma_n(A)$ and $\sigma_1(B) \ge ... \ge \sigma_n(B)$ are the singular values of A and B.*

Next, we produce a generalized version of Hadamard's inequality for matrices with a large low-rank component. Here and in what follows, we use the convention that $0^0 = 1$.

Lemma 3.2. *Let $A, B \in \mathbb{C}^{n\times n}$ with $\|A\|_F \le n$, $\|B\|_F \le Cn$, and $\operatorname{rank}(B) \le \ell$. Then*

$$|\det(A+B)| \le \frac{n^n}{(n-\ell)^{\frac{n-\ell}{2}} \ell^{\frac{\ell}{2}}} \big(1+C\big)^{\ell}.$$

Proof. Let $0 < \ell < n$, and $\sigma_1(A) \ge ... \ge \sigma_n(A)$ and $\sigma_1(B) \ge ... \ge \sigma_n(B)$ denote the singular values of A and B. By Proposition 3.1,

$$\begin{aligned}
|\det(A+B)| &\le \bigg(\prod_{i=1}^{n-\ell} \sigma_i(A)\bigg) \prod_{j=1}^{\ell} \big(\sigma_j(B) + \sigma_{n-j+1}(A)\big) \\
&\le \bigg(\frac{1}{n-\ell}\sum_{i=1}^{n-\ell} \sigma_i^2(A)\bigg)^{\frac{n-\ell}{2}} \bigg(\frac{1}{\ell}\sum_{j=1}^{\ell}\sigma_j(B) + \frac{1}{\ell}\sum_{j=1}^{\ell}\sigma_{n-j+1}(A)\bigg)^{\ell} \\
&\le \bigg(\frac{1}{n-\ell}\sum_{i=1}^{n-\ell} \sigma_i^2(A)\bigg)^{\frac{n-\ell}{2}} \bigg(\frac{1}{\ell^{\frac12}}\bigg[\sum_{j=1}^{\ell}\sigma_j^2(B)\bigg]^{\frac12} + \frac{1}{\ell^{\frac12}}\bigg[\sum_{j=1}^{\ell}\sigma_{n-j+1}^2(A)\bigg]^{\frac12}\bigg)^{\ell} \\
&\le \bigg(\frac{n^2}{n-\ell}\bigg)^{\frac{n-\ell}{2}} \bigg(\frac{Cn}{\ell^{\frac12}} + \frac{n}{\ell^{\frac12}}\bigg)^{\ell} \\
&= \frac{n^n}{(n-\ell)^{\frac{n-\ell}{2}}\ell^{\frac{\ell}{2}}}(1+C)^{\ell},
\end{aligned}$$

where we have used the AM-GM inequality in the second inequality and Cauchy-Schwarz in the third. The result for the cases $\ell = 0$ and $\ell = n$ follows from gently modified versions of the same analysis. $\square$

We note that, when $\ell = 0$, Lemma 3.2 is the well-known corollary $|\det(A)| \le n^{n/2}|A|_\infty$ of Hadamard's inequality. A tighter version of Lemma 3.2 can be obtained at the cost of brevity, by explicitly maximizing with respect to the parameter $x := \sum_{j=1}^{\ell} \sigma_{n-j+1}^2(A)$ rather than upper bounding both $\sum_{i=1}^{n-\ell}\sigma_i^2(A)$ and $\sum_{j=1}^{\ell}\sigma_{n-j+1}^2(A)$ with n^2. However, this optimization does not lead to any improvement in the exponential constant of Theorem 1.1, and so its derivation is left to the interested reader.

3.2. **An Improved Optimization Problem.** Lemma 3.2 applied to the matrix iterates $A^{(k)} \in \mathbb{C}^{k\times k}$ of Gaussian elimination under complete pivoting leads to further constraints on the pivots $p_k = |A^{(k)}|_\infty$. Consider some $0 < \ell < k$ with $k + \ell \le n$. Using block notation, let $M^{(1,1)}$, $M^{(1,2)}$, $M^{(2,1)}$, and $M^{(2,2)}$ denote the upper-left $\ell\times\ell$, upper-right $\ell\times k$, lower-left $k\times\ell$,

[2] Proposition 3.1 also follows from applying standard determinant bounds for Hermitian matrices [2, Theorem VI.7.1] to $\left(\begin{smallmatrix} 0 & A \\ A^* & 0\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & B \\ B^* & 0\end{smallmatrix}\right)$, and using the following well-known rearrangement inequality: for any $a_1 \ge ... \ge a_n \ge 0$, $b_1 \ge ... \ge b_n \ge 0$, and $\pi \in \mathrm{S}_n$, $\prod_{i=1}^n (a_i + b_{\pi(i)}) \le \prod_{i=1}^n (a_i + b_{n-i+1})$.

and lower-right $k \times k$ sub-matrices of $A^{(k+\ell)}$. After ℓ further steps of Gaussian elimination applied to $A^{(k+\ell)}$, we obtain

$$A^{(k+\ell)} = \begin{bmatrix} M^{(1,1)} & M^{(1,2)} \\ M^{(2,1)} & M^{(2,2)} \end{bmatrix} = \begin{bmatrix} \widetilde{L} & 0 \\ M^{(2,1)}\widetilde{U}^{-1} & I \end{bmatrix} \begin{bmatrix} \widetilde{U} & \widetilde{L}^{-1}M^{(1,2)} \\ 0 & M^{(2,2)} - M^{(2,1)}[M^{(1,1)}]^{-1}M^{(1,2)} \end{bmatrix},$$

where $\widetilde{L}\widetilde{U}$ is the LU factorization of $M^{(1,1)}$, implying that

$$A^{(k)} = M^{(2,2)} - M^{(2,1)}[M^{(1,1)}]^{-1}M^{(1,2)}.$$

For the sake of space, let $X := M^{(2,2)}$ and $Y := M^{(2,1)}[M^{(1,1)}]^{-1}M^{(1,2)}$, and note that Y has rank at most ℓ. We may rewrite $A^{(k)}$ as

$$A^{(k)} = \left(X - \frac{\operatorname{Re}\langle X, Y\rangle_F}{\|Y\|_F^2} Y \right) - \left(1 - \frac{\operatorname{Re}\langle X, Y\rangle_F}{\|Y\|_F^2} \right) Y, \tag{3.1}$$

where $\langle \cdot, \cdot \rangle_F$ is the Frobenius inner product. We note that

$$\left\| X - \frac{\operatorname{Re}\langle X, Y\rangle_F}{\|Y\|_F^2} Y \right\|_F^2 = \|X\|_F^2 - \frac{\left(\operatorname{Re}\langle X, Y\rangle_F\right)^2}{\|Y\|_F^2} \le \|X\|_F^2 \le p_{k+\ell}^2 n^2$$

and

$$\begin{aligned} \left\| \left(1 - \frac{\operatorname{Re}\langle X, Y\rangle_F}{\|Y\|_F^2} \right) Y \right\|_F^2 &= \|Y\|_F^2 - 2\operatorname{Re}\langle X, Y\rangle_F + \frac{\left(\operatorname{Re}\langle X, Y\rangle_F\right)^2}{\|Y\|_F^2} \\ &\le \|Y\|_F^2 - 2\operatorname{Re}\langle X, Y\rangle_F + \|X\|_F^2 \\ &= \|X - Y\|_F^2 \le p_k^2 n^2, \end{aligned}$$

as the entries of $A^{(k)}$ and $M^{(2,2)}$ have modulus at most p_k and $p_{k+\ell}$, respectively. Applying Lemma 3.2 to $A^{(k)}$ using the splitting in Equation 3.1, we obtain the bound

$$\frac{\prod_{i=1}^k p_i}{p_{k+\ell}^k} = \frac{\det(A^{(k)})}{p_{k+\ell}^k} \le \frac{k^k}{(k-\ell)^{\frac{k-\ell}{2}} \ell^{\frac{\ell}{2}}} \left(1 + \frac{p_k}{p_{k+\ell}} \right)^{\ell}. \tag{3.2}$$

Making use of these additional constraints gives the following refinement of Optimization Problem 2.2:

Improved Optimization Problem

$$\begin{aligned} \max \quad & p_1/p_n \\ \text{s.t.} \quad & \textstyle\prod_{i=1}^k p_i \le k^{k/2} p_k^k && \text{for } k = 1, ..., n \\ & \textstyle\prod_{i=1}^k p_i \le \dfrac{k^k p_{k+\ell}^{k-\ell} (p_k + p_{k+\ell})^{\ell}}{(k-\ell)^{\frac{k-\ell}{2}} \ell^{\frac{\ell}{2}}} && \text{for } \ell = 1, ..., \min\{k-1, n-k\} \\ & && \quad k = 2, ..., n-1. \end{aligned} \tag{3.3}$$

3.3. **From a Non-Linear to Linear Program.** The additional constraints given by Inequality 3.2 for $k = 2, ..., n-1$ and $\ell = 1, ..., \min\{k-1, n-k\}$ produce an optimization problem (Optimization Problem 3.3) that is no longer linear upon the transformation $q_k = \ln(p_k)$, $k = 1, ..., n$. For this reason, we relax Optimization Problem 3.3 in order to maintain linearity. For simplicity, we do so while giving only minor attention to lower-order terms (e.g., terms that do not affect the leading exponential constant). More complicated linear programs with improved behavior for finite n can be obtained by a more involved analysis.

Consider an arbitrary feasible point $(p_1, ..., p_n)$ of Optimization Problem 3.3. We claim that $(p_1, ..., p_n)$ also satisfies

$$\prod_{i=1}^{k} p_i \le (\tfrac{11}{4}k)^{k/2} p_{k+\ell}^{k-\ell}\, p_k^{\ell} \qquad \text{for} \quad \begin{array}{l} \ell = 1, ..., \min\{k-1, n-k\} \\ k = 2, ..., n-1. \end{array} \tag{3.4}$$

We break our analysis into two cases. If $p_k \le (\sqrt{11}/2)^{k/(k-\ell)} p_{k+\ell}$, then

$$\prod_{i=1}^{k} p_i \le k^{k/2} p_k^k \le (\tfrac{11}{4}k)^{k/2} p_{k+\ell}^{k-\ell}\, p_k^{\ell}.$$

Conversely, if $p_k \ge (\sqrt{11}/2)^{k/(k-\ell)} p_{k+\ell}$, then

$$\prod_{i=1}^{k} p_i \le \frac{k^k p_{k+\ell}^{k-\ell}\, (p_k + p_{k+\ell})^{\ell}}{(k-\ell)^{\frac{k-\ell}{2}} \ell^{\frac{\ell}{2}}} \le k^{k/2} p_{k+\ell}^{k-\ell}\, p_k^{\ell} \left(\frac{k^{k/2}\big(1 + (\frac{2}{\sqrt{11}})^{k/(k-\ell)}\big)^{\ell}}{(k-\ell)^{\frac{k-\ell}{2}} \ell^{\frac{\ell}{2}}} \right) \le (\tfrac{11}{4}k)^{k/2} p_{k+\ell}^{k-\ell} p_k^{\ell},$$

where we have used the fact that

$$\max_{t\in(0,1)} \left(\frac{1}{t}\right)^{\frac{t}{2}} \left(\frac{1}{1-t}\right)^{\frac{1-t}{2}} = \sqrt{2} \quad \text{and} \quad \max_{t\in(0,1)} \left(1 + \left(\frac{2}{\sqrt{11}}\right)^{1/(1-t)}\right)^{t} \approx 1.168 < \sqrt{\frac{11}{8}}.$$

Applying the transformation $q_k = \ln(p_k)$, $k = 1, ..., n$, to Inequality 3.4, we obtain the linear program:

Improved Linear Program

$$\begin{array}{lll} \max & q_1 - q_n & \\ \text{s.t.} & \sum_{i=1}^{k} q_i \le \frac{k}{2}\ln(k) + kq_k & \text{for } k = 1, ..., n \\ & \sum_{i=1}^{k} q_i \le \frac{k}{2}\ln(\frac{11}{4}k) + (k-\ell)q_{k+\ell} + \ell q_k & \text{for } \ell = 1, ..., \min\{k-1, n-k\} \\ & & \quad k = 2, ..., n-1. \end{array} \tag{3.5}$$

and note that the maximum growth factor $g_n(\mathbb{C})$ is upper bounded by e^{OPT}, where OPT is the optimal value of this linear program. Program 3.5 is an improved version of Wilkinson's linear program (Program 2.3), containing all of Wilkinson's constraints as well as additional bounds representing long-range interactions (e.g., bounds relating $A^{(k)}$ and $A^{(k+\ell)}$). In addition, we note that the optimal value of Program 3.5 and the logarithm of the optimal value of Program 3.3 are asymptotically equal up to lower order terms:

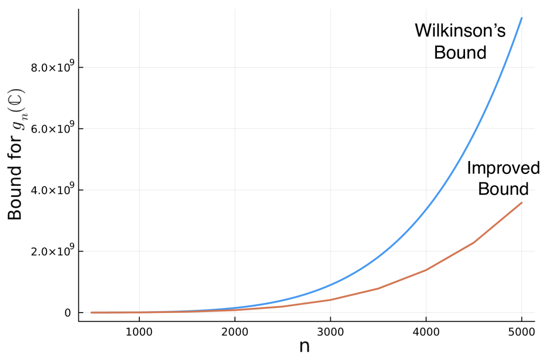
Proposition 3.3. *If OPT is the optimal value of Linear Program 3.5 for n, then the optimal value of Optimization Problem 3.3 for n lies in the interval* $[n^{-3/2}e^{OPT}, e^{OPT}]$.

Proof. Let $(q_1, ..., q_n)$ be a feasible point of Linear Program 3.5. It suffices to show that $p_k = k^{3/2}e^{q_k}$, $k = 1, ..., n$, is a feasible point of Optimization Problem 3.3. Considering an arbitrary constraint parameterized by $k > 1$ and ℓ, we have

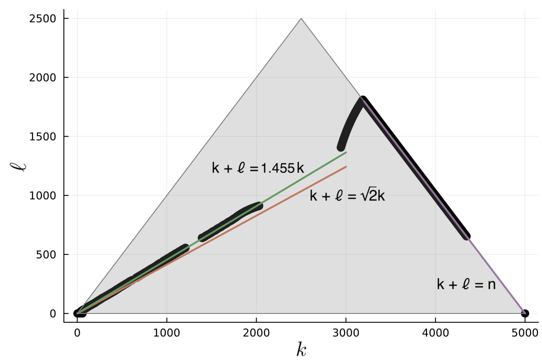
$$\prod_{i=1}^{k} p_i = (k!)^{\frac{3}{2}} \exp\left\{\sum_{i=1}^{k} q_i\right\} \le (k!)^{\frac{3}{2}} \exp\left\{\frac{k}{2}\ln(\tfrac{11}{4}k) + (k-\ell)q_{k+\ell} + \ell q_k\right\}.$$

Rewriting the right-hand side in terms of p_k gives

$$\prod_{i=1}^{k} p_i \le \frac{(k!)^{\frac{3}{2}} (\frac{11}{4}k)^{\frac{k}{2}} p_{k+\ell}^{k-\ell} p_k^{\ell}}{(k+\ell)^{\frac{3}{2}(k-\ell)} k^{\frac{3}{2}\ell}} \le k^{\frac{k}{2}} p_{k+\ell}^{k-\ell} p_k^{\ell} \frac{(k!)^{\frac{3}{2}} (\frac{11}{4})^{\frac{k}{2}}}{k^{\frac{3}{2}k}} \le k^{\frac{k}{2}} p_{k+\ell}^{k-\ell} p_k^{\ell} \le \frac{k^k p_{k+\ell}^{k-\ell}(p_k + p_{k+\ell})^{\ell}}{(k-\ell)^{\frac{k-\ell}{2}} \ell^{\frac{\ell}{2}}},$$



(a) Comparing Wilkinson's Bound to our Improved Bound for $g_n(\mathbb{C})$ for $n \leq 5000$



(b) Active constraints parameterized by (k, ℓ) for the optimal solution to Program 3.5 at $n = 5000$

FIGURE 1. Comparing our Improved Linear Program to Wilkinson's LP: Figure (a) illustrates the difference between Wilkinson's bound for $g_n(\mathbb{C})$ (Inequality 1.1) and the upper bound produced by the optimal value of Program 3.5 for $n \leq 5000$. This illustrates that our improved linear program gives superior estimates even for very small n. Figure (b) is a scatter plot of the pairs (k, ℓ) for which the corresponding inequality in Program 3.5 is tight for a numerically computed optimal solution at $n = 5000$. The grey shaded triangle shows the set of (k, ℓ) corresponding to constraints of Program 3.5, with Wilkinson's constraints parameterized by $(k, 0)$, and the black dots represent the subset of those constraints that are active for the numerically computed optimal solution. For $n = 5000$, almost none of Wilkinson's constraints are active. The red line $k + \ell = \sqrt{2}k$ is the set of constraints used in Section 4 to prove Theorem 1.1, and the green line denotes the asymptotically tight constraints for the feasible point produced in Subsection 4.1. While the points on the purple line $k + \ell = n$ improves the objective value, these constraints do not play a role in the asymptotic leading term of the solution to the linear program.

completing the proof. $\square$

In the following section, we provide nearly matching upper and lower bounds on the optimal value of Program 3.5 for sufficiently large n, thereby proving Theorem 1.1.

3.4. **Bounding the Growth Factor in Practice.** While the proof of Theorem 1.1 focuses on the behavior for large n, we note that an improvement in exponential constant exists in practice for reasonably sized matrices as well. We provide a comparison of the optimal value of Program 3.5 to the optimal value of Wilkinson's LP in Figure 1 for $n \leq 5000$. The numerically computed solutions to Program 3.5 were obtained using the Gurobi Optimizer [10] called through the JuMP package for mathematical optimization [12] in the Julia programming language [1]. We stress that numerically computed solutions to a linear program can be converted into mathematical bounds via a dual feasible point verified in exact arithmetic. In addition, Program 3.5 can be adapted in a number of ways for computational efficiency. For instance, the linear transformation $Q(k) = \sum_{i=1}^{k} q_i$ produces a linear program with a simple objective and sparse constraints (at most four variables in each). Furthermore, as the analysis in Section 4 suggests, only a linear number of constraints are required to produce a reasonable upper bound for the optimal value. One natural choice would consist of Wilkinson's original constraints, and all additional constraints of the form $k+\ell = n$ and $k+\ell \in [\sqrt{2}k-1, \sqrt{2}k+C]$ for some constant C (Theorem 1.1 is proved using only constraints of the form $k+\ell = \lceil \sqrt{2}k \rceil$).

Finally, we stress that the techniques used in this paper to produce improved estimates can be further optimized to obtain even better bounds in both theory and practice. We hope that the interested reader will do so.

4. Bounding the Optimal Value of our Linear Program

Finally, we prove that the objective of Program 3.5 satisfies the bound

$$\max\ q_1 - q_n \ \le\ \alpha \ln^2 n + (\beta + 1/2)\ln n, \qquad \text{where} \quad \alpha = \frac{1}{2(2+(2-\sqrt{2})\ln 2)}$$

and $\beta = 0.41$, thus completing the proof of Theorem 1.1. We do so via a duality argument, making use of the constraints for k and ℓ satisfying $k + \ell \approx \sqrt{2}k$. Before proving the above bound, we first illustrate why $[2(2+(2-\sqrt{2})\ln 2)]^{-1}$ is the correct choice of α for constraints of the form $k + \ell \approx \sqrt{2}k$, and show that this choice is within 0.00024 of the exact asymptotic constant of Program 3.5.

4.1. On the Choice and Optimality of the Constant $\alpha = [2(2+(2-\sqrt{2})\ln 2)]^{-1}$. Suppose that $q_x - q_1 = -\gamma \ln^2 x + O(1)$. Then, for the constraint

$$\sum_{i=1}^{k} (q_i - q_1) \le \frac{k}{2}\ln(\tfrac{11}{4}k) + (k-\ell)(q_{k+\ell} - q_1) + \ell(q_k - q_1),$$

the left-hand side equals

$$\int_1^k -\gamma \ln^2 x \, \mathrm{d}x + O(k) = -\gamma k \ln^2 k + 2\gamma k \ln k + O(k)$$

and the right-hand side equals

$$-\gamma k \ln^2 k + \big[k/2 - 2\gamma(k-\ell)\ln(1+\ell/k)\big]\ln k + O(k).$$

Letting $t = \ell/k$, the right-hand side is asymptotically larger than the left-hand side if

$$\gamma \le \frac{1}{4(1+(1-t)\ln(1+t))}.$$

The values $t = 0$ and $t = 1$ (e.g., when $\ell = 0$ or $\ell = k$) correspond to the constraints of Wilkinson's linear program, and for $t = 0$ and $t = 1$, we obtain $\gamma \le 1/4$ (e.g., Wilkinson's bound). The value $t = \sqrt{2} - 1$ produces the upper bound $1/[2(2+(2-\sqrt{2})\ln 2)] \approx 0.20781$ of Theorem 1.1. The quantity $[4(1+(1-t)\log(1+t))]^{-1}$ on the interval $[0, 1]$ is minimized by $t = \exp\{W(2e) - 1\} - 1 \approx 0.4547$, where $W(x)$ is the Lambert W function, with a minimum value of

$$\frac{1}{4\big(1 + (2 - e^{W(2e)-1})(W(2e)-1)\big)} \approx 0.207576.$$

This implies the existence of a solution to Program 3.5 with $q_1 - q_n = 0.207575 \ln^2 n - O(\ln n)$, thus illustrating that our upper bound of $\alpha = [2(2+(2-\sqrt{2})\ln 2)]^{-1} \approx 0.207811$ is within 0.00024 of the optimal value of the linear program. We do not pursue further improvement on this constant.

4.2. Reducing Theorem 1.1 to Geometric Mean Growth. For ease of analysis, we consider a continuous version of our variables $q = (q_1, ..., q_n)$. Let

$$f(x) = q_{\lceil x \rceil} - q_1 \qquad \text{and} \qquad F(x) = \frac{1}{x}\int_0^x f(t)\,\mathrm{d}t \qquad \text{for } x > 0,$$

where $\{q_k\}_{k=1}^{\infty}$ is any sequence such that $(q_1, ..., q_n)$ is a feasible point of Program 3.5 for all $n \in \mathbb{N}$. Any optimal solution $(q_1, ..., q_n)$ for the n-dimensional linear program can be converted into such a sequence by simply setting $q_k = q_n$ for all $k > n$. The constraint of Program 3.5 with $k = \lceil x \rceil$ and $\ell = \lceil \sqrt{2}x \rceil - \lceil x \rceil$ implies that for all $x > 0$,

$$\begin{aligned} F(\lceil x \rceil) &\le \frac{\ln(\frac{11}{4}\lceil x \rceil)}{2} + \left(\frac{2\lceil x \rceil - \lceil \sqrt{2}x \rceil}{\lceil x \rceil} \right) f(\sqrt{2}x) + \left(\frac{\lceil \sqrt{2}x \rceil - \lceil x \rceil}{\lceil x \rceil} \right) f(x) \\ &\le \frac{\ln(\frac{11}{4}x)}{2} + \frac{1}{2x} + \left(\sqrt{2} - 1 - \frac{\sqrt{2}}{x} \right) \left(\sqrt{2} f(\sqrt{2}x) + f(x) \right). \end{aligned} \tag{4.1}$$

We make the following claim regarding $F(x)$.

Lemma 4.1. *$F(x) > -\alpha \ln^2 x - \beta \ln x$ for all $x > 100$.*

Lemma 4.1 implies our desired result, as

$$F(n) = \frac{1}{n} \sum_{i=1}^{n} (q_i - q_1) \le \frac{1}{n} \left(\frac{n}{2} \ln n + n q_n - n q_1 \right),$$

and $\alpha \ln^2 n + (\beta + 1/2) \ln n$ is larger than Wilkinson's bound for $x \le 100$. A tighter bound may be obtained by adding together constraints of the form $k + \ell = n$ for $k \ge n/(8\alpha)$ (e.g., the constraints appearing in Figure 1(b)). However, the analysis is involved and the improvement on the $1/2 \ln n$ term produced by the argument above is minor (≈ 0.046 improvement, at the cost of lower-order terms).

4.3. **Proof of Lemma 4.1: Base Case.** The proof of Lemma 4.1 is, in spirit, by "induction on x" via a duality argument. Clearly the assertion holds for $x \in (100, 1700]$ for β sufficiently large. However, verifying the base case of $x \in (100, 1700]$ for $\beta = 0.41$ requires some analysis, as the quantity $\alpha \ln^2 n + \beta \ln n$ is strictly less than Wilkinson's bound. We have

$$F(x) = \frac{1}{x} \int_0^x q_{\lceil t \rceil} - q_1 \, \mathrm{d}t = \frac{x - \lfloor x \rfloor}{x} (q_{\lceil x \rceil} - q_1) + \frac{1}{x} \sum_{k=1}^{\lfloor x \rfloor} (q_k - q_1).$$

By Inequalities 1.1 and 2.5,

$$q_1 - q_{\lceil x \rceil} \le \frac{\ln^2 \lceil x \rceil}{4} + \frac{\ln \lceil x \rceil}{2} + \ln 2 \qquad \text{and} \qquad \frac{1}{\lfloor x \rfloor} \sum_{k=1}^{\lfloor x \rfloor} (q_1 - q_k) \le \frac{\ln^2 \lfloor x \rfloor}{4} + \ln 2.$$

Altogether, we obtain the lower bound

$$\begin{aligned} F(x) &\ge -\frac{1}{x} \left(\frac{\ln^2 \lceil x \rceil}{4} + \frac{\ln \lceil x \rceil}{2} + \ln 2 \right) - \left(\frac{\ln^2 \lfloor x \rfloor}{4} + \ln 2 \right) \\ &\ge -\frac{1}{x} \left(\frac{(\ln x + \frac{1}{x})^2}{4} + \frac{\ln x + \frac{1}{x}}{2} + \ln 2 \right) - \left(\frac{\ln^2 x}{4} + \ln 2 \right). \end{aligned}$$

By inspection, the right-hand side of the above inequality is strictly greater than $-(\alpha \ln^2 x + \beta \ln x)$ for our interval of interest $x \in [100, 1700]$.

4.4. **Proof of Lemma 4.1: Inductive Step.** In order to verify the claim for some $y > 1700$, we integrate over $x \in \left[\frac{y}{2}, \frac{y}{\sqrt{2}}\right]$ to obtain a lower bound for $F(y)$ in terms of $F(x)$ for $x < y$. In particular, by integrating Inequality 4.1 over $x \in \left[\frac{y}{2}, \frac{y}{\sqrt{2}}\right]$ we have

$$\begin{aligned}\frac{1}{\frac{y}{\sqrt{2}} - \frac{y}{2}} \int_{\frac{y}{2}}^{\frac{y}{\sqrt{2}}} F(\lceil x \rceil)\,\mathrm{d}x \;&\leq\; \frac{1}{\frac{y}{\sqrt{2}} - \frac{y}{2}} \left[\left(\sqrt{2} - 1 - \frac{2\sqrt{2}}{y}\right) \int_{\frac{y}{2}}^{y} f(x)\,\mathrm{d}x + \int_{\frac{y}{2}}^{\frac{y}{\sqrt{2}}} \frac{\ln(\frac{11}{4}x)}{2} + \frac{1}{2x}\,\mathrm{d}x\right] \\ &= \left(1 - \frac{4 + 2\sqrt{2}}{y}\right)\left(2F(y) - F(\tfrac{y}{2})\right) + \frac{\ln y}{2} \\ &\quad + \frac{\ln 2 + \sqrt{2}\ln\frac{11}{4} - \sqrt{2}}{2\sqrt{2}} + \frac{(\sqrt{2}+1)\ln 2}{2y}.\end{aligned}$$

Rearranging the above inequality allows us to lower bound $F(y)$ by a positive linear combination of $F(x)$ for $x \in \left[\frac{y}{2}, \frac{y}{\sqrt{2}}\right]$. We note that this is the reason for the choice of $k + \ell \approx \sqrt{2}k$, as this approach does not give us such a bound if $\sqrt{2}$ is replaced by a larger constant. Now, suppose our claim is false, and let $y > 1700$ be the smallest value such that $F(y) \leq -\alpha \ln^2 y - \beta \ln y$. We aim to show that this contradicts the above lower bound for $F(y)$. By assumption,

$$\begin{aligned}F(\lceil x \rceil) &> -\alpha \ln^2(x+1) - \beta \ln(x+1) \\ &> -\alpha \ln^2 x - \beta \ln x - \frac{2\alpha \ln x}{x} - \frac{\beta}{x} - \frac{\alpha}{x^2} \qquad \text{for } x \in \left[\tfrac{y}{2}, \tfrac{y}{\sqrt{2}}\right],\end{aligned}$$

implying that

$$\begin{aligned}\frac{1}{\frac{y}{\sqrt{2}} - \frac{y}{2}} \int_{\frac{y}{2}}^{\frac{y}{\sqrt{2}}} F(\lceil x \rceil)\,\mathrm{d}x \;&>\; -\alpha \ln^2 y - \left((\sqrt{2}\ln 2 - 2)\alpha + \beta\right)\ln y - \left(\frac{\ln 2}{\sqrt{2}} - 1\right)\beta \\ &\quad - \left(2 - \frac{(3+\sqrt{2})\ln^2 2}{2\sqrt{2}} - \sqrt{2}\ln 2\right)\alpha - \frac{2(\sqrt{2}+1)\alpha \ln 2 \ln y}{y} \\ &\quad - \frac{(\sqrt{2}+1)(\beta \ln 2 - \frac{3}{2}\alpha \ln^2 2)}{y} - \frac{2\sqrt{2}\alpha}{y^2}.\end{aligned}$$

In addition,

$$2F(y) - F(\tfrac{y}{2}) < -\alpha \ln^2 y - (2\alpha \ln 2 + \beta)\ln y + \alpha \ln^2 2 - \beta \ln 2.$$

Combining our upper and lower bounds, we observe that the terms containing $\ln^2 y$ are equal, and the terms containing $\ln y$ are equal

$$-\left((\sqrt{2}\ln 2 - 2)\alpha + \beta\right) = \frac{1}{2} - (2\alpha \ln 2 + \beta)$$

due to the value of α. We are left with the inequality

$$\frac{(\sqrt{2}-1)\ln 2 + \sqrt{2}}{\sqrt{2}}\beta + \frac{(2-\sqrt{2})\ln^2 2 - 4(2-\sqrt{2})(\ln\frac{11}{4} - 1)\ln 2 - 8\ln\frac{11}{4}}{8(2 + (2-\sqrt{2})\ln 2)} + g(\beta, y) < 0,$$

where $g(\beta, y)$ is a linear function of β of order $O(\ln^2(y)/y)$. The left-hand side is strictly greater than zero for a sufficiently large choice of β. However, verifying that our choice of $\beta = 0.41$ is sufficiently large requires an explicit analysis of $g(\beta, y)$ for $\beta = 0.41$ and $y > 1700$. The function $g(\beta, y)$ is given by

$$\begin{aligned}g(\beta, y) = &-\frac{2+\sqrt{2}}{2 + (2-\sqrt{2})\ln 2}\,\frac{\ln^2 y}{y} - \left((4+2\sqrt{2})\beta + \frac{(5+3\sqrt{2})\ln 2}{2 + (2-\sqrt{2})\ln 2}\right)\frac{\ln y}{y} \\ &+ \left(\frac{(11+4\sqrt{2})\ln^2 2}{4(2 + (2-\sqrt{2})\ln 2)} - (5+3\sqrt{2})\beta \ln 2 - \frac{(\sqrt{2}+1)\ln 2}{2}\right)\frac{1}{y} \\ &- \frac{\sqrt{2}}{2 + (2-\sqrt{2})\ln 2}\,\frac{1}{y^2}.\end{aligned}$$

When $\beta = 0.41$ and $y > 1700$,

$$\frac{(\sqrt{2}-1)\ln 2 + \sqrt{2}}{\sqrt{2}}\beta + \frac{(2-\sqrt{2})\ln^2 2 - 4(2-\sqrt{2})(\ln\frac{11}{4} - 1)\ln 2 - 8\ln\frac{11}{4}}{8(2+(2-\sqrt{2})\ln 2)} > 0.086$$

and

$$g(0.41, y) > -\frac{\frac{3}{2}\ln^2 y}{y} - \frac{6\ln y}{y} - \frac{3}{y} - \frac{1}{y^2} > -\frac{\frac{3}{2}\ln^2 1700}{1700} - \frac{6\ln 1700}{1700} - \frac{3}{1700} - \frac{1}{1700^2} > -0.08,$$

thus obtaining our desired contradiction. This completes the proof of Theorem 1.1.

Acknowledgements

This material is based upon work supported by the National Science Foundation under grant no. OAC-1835443, grant no. SII-2029670, grant no. ECCS-2029670, grant no. OAC-2103804, and grant no. PHY-2021825. We also gratefully acknowledge the U.S. Agency for International Development through Penn State for grant no. S002283-USAID. The information, data, or work presented herein was funded in part by the Advanced Research Projects Agency-Energy (ARPA-E), U.S. Department of Energy, under Award Number DE-AR0001211 and DE-AR0001222. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. This material was supported by The Research Council of Norway and Equinor ASA through Research Council project "308817 - Digital wells for optimal production and drainage". Research was sponsored by the United States Air Force Research Laboratory and the United States Air Force Artificial Intelligence Accelerator and was accomplished under Cooperative Agreement Number FA8750-19-2-1000. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the United States Air Force or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein. The third author thanks Mehtaab Sawhney for interesting conversions regarding linear programming. The authors thank Louisa Thomas for improving the style of presentation.

References

[1] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B Shah. Julia: A fresh approach to numerical computing. *SIAM review*, 59(1):65–98, 2017.
[2] Rajendra Bhatia. *Matrix analysis*, volume 169. Springer Science & Business Media, 2013.
[3] AM Cohen. A note on pivot size in Gaussian elimination. *Linear Algebra and its Applications*, 8(4):361–368, 1974.
[4] Colin W Cryer. Pivot size in Gaussian elimination. *Numerische Mathematik*, 12(4):335–345, 1968.
[5] Alan Edelman. The complete pivoting conjecture for Gaussian elimination is false. 1992.
[6] Alan Edelman and John Urschel. Some new results on the maximum growth factor in Gaussian elimination. *arXiv preprint arXiv:2303.04892*, 2023.
[7] Herman H Goldstine. *The computer from Pascal to von Neumann*. Princeton University Press, 1993.
[8] Gene H Golub and Charles F Van Loan. *Matrix computations*. JHU press, 2013.
[9] Nick Gould. On growth in Gaussian elimination with complete pivoting. *SIAM Journal on Matrix Analysis and Applications*, 12(2):354–361, 1991.
[10] Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2023.
[11] Chi-Kwong Li and Roy Mathias. The determinant of the sum of two matrices. *Bulletin of the Australian Mathematical Society*, 52(3):425–429, 1995.
[12] Miles Lubin, Oscar Dowson, Joaquim Dias Garcia, Joey Huchette, Benoît Legat, and Juan Pablo Vielma. JuMP 1.0: Recent improvements to a modeling language for mathematical optimization. *Mathematical Programming Computation*, 2023.
[13] Carl Meyer. History of Gaussian elimination. `http://carlmeyer.com/pdfFiles/GaussianEliminationHistory.pdf`.
[14] Erich Strohmaier, Jack Dongarra, Horst Simon, and Martin Meuer. Top 500 list: November 2023. `https://www.top500.org/lists/top500/2023/11/`.
[15] Leonard Tornheim. Pivot size in Gauss reduction. *Tech Report, Chevron Research Co., Richmond CA*, 1964.
[16] Leonard Tornheim. Maximum third pivot for Gaussian reduction. In *Tech. Report*. Calif. Res. Corp Richmond, Calif, 1965.
[17] Leonard Tornheim. A bound for the fifth pivot in Gaussian elimination. *Tech Report, Chevron Research Co., Richmond CA*, 1969.
[18] Leonard Tornheim. Maximum pivot size in Gaussian elimination with complete pivoting. *Tech Report, Chevron Research Co., Richmond CA*, 10, 1970.
[19] John von Neumann and Herman H Goldstine. Numerical inverting of matrices of high order. 1947.
[20] James Hardy Wilkinson. Error analysis of direct methods of matrix inversion. *Journal of the ACM (JACM)*, 8(3):281–330, 1961.