$$\begin{bmatrix} A00 & A01 & A02 \\ A10 & A11 & A12 \\ A20 & A21 & A22 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ L10 & 1 & 0 \\ L20 & L21 & 1 \end{bmatrix} \begin{bmatrix} U00 & U01 & U02 \\ 0 & U11 & U12 \\ 0 & 0 & U22 \end{bmatrix}$$

GROWTH FACTORS IN GAUSSIAN ELIMINATION

John Urschel, Harvard Society of Fellows

joint work with Alan Edelman

RESEARCH INTERESTS

I work in matrix analysis, primarily focusing on:

- **Spectral Graph Theory**: properties of matrix representations of graphs
- **Numerical Linear Algebra**: matrix computations (linear systems, eigenvalue problems) with a focus on theoretical results
- **Machine Learning/Theoretical Computer Science**: solving theoretical problems that have some linear algebraic formulation

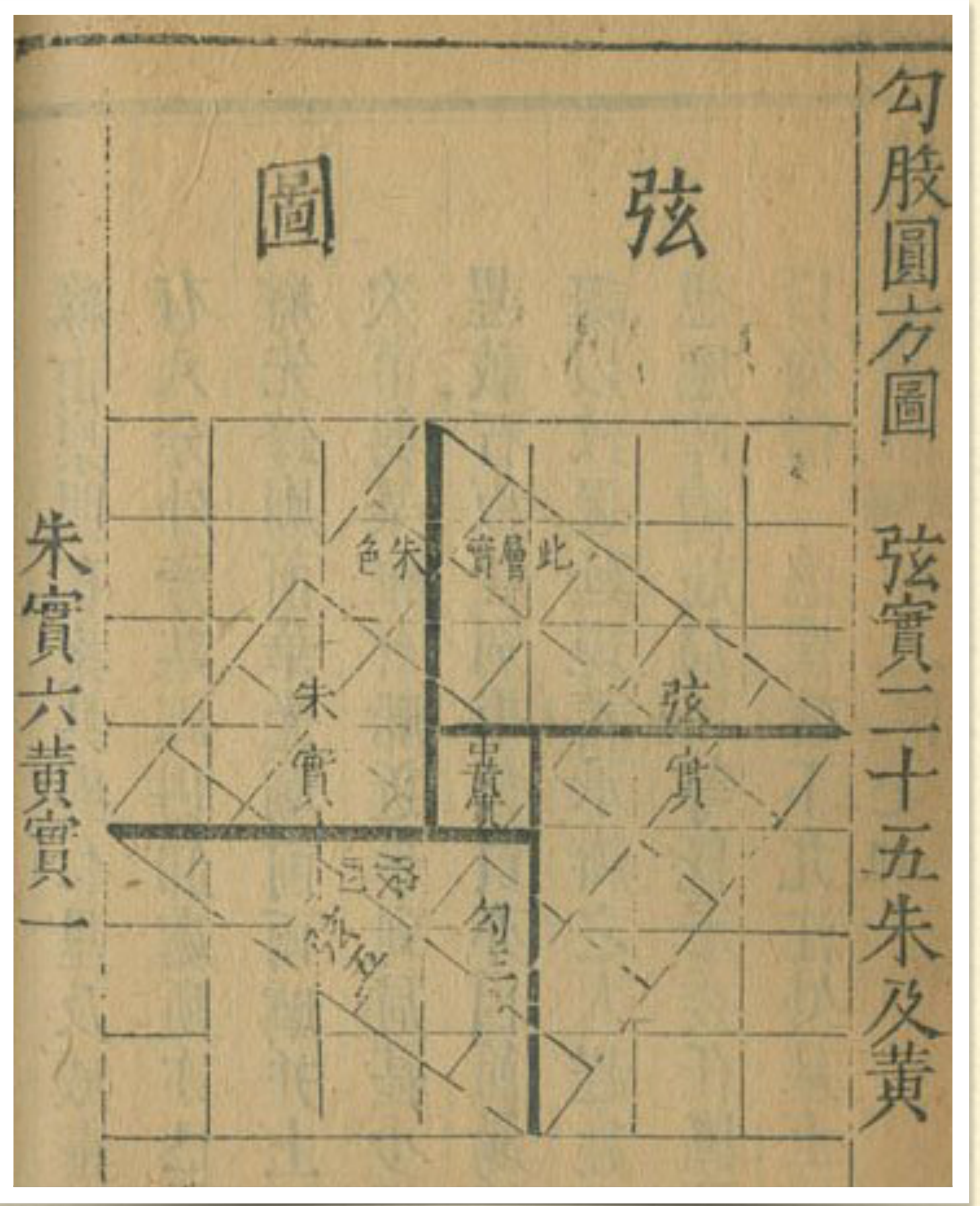
OUTLINE

- Part I: Gaussian Elimination
- Part II: Growth Factor & Why We Should Care
- Part III: History of Question & Recent Results

OUTLINE

- **Part I: Gaussian Elimination**
- Part II: Growth Factor & Why We Should Care
- Part III: History of Question & Recent Results

GAUSSIAN ELIMINATION



Eliminates unknowns by subtracting equations

In 8^{th} chapter of Jiuzhang suanshu (2^{nd} century)

1600 years before Gauss

GAUSSIAN ELIMINATION

sequence of rank one updates:

$$A = A^{(1)} \to A^{(2)} \to \ldots \to A^{(n)}$$

$$n \times n \qquad (n-1) \times (n-1) \qquad 1 \times 1$$

$$A^{(k+1)} := A^{(k)}_{k+1:n,k+1:n} - \frac{1}{A^{(k)}_{k,k}} A^{(k)}_{k+1:n,k} A^{(k)}_{k,k+1:n}$$

GAUSSIAN ELIMINATION

not every matrix has one

e.g. $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

$$L(i,j) = \frac{A^{(j)}_{i,j}}{A^{(j)}_{j,j}} \quad i \geq j \quad \text{and} \quad U(i,j) = A^{(i)}_{i,j} \quad j \geq i$$

GAUSSIAN ELIMINATION

Given $A = LU$, can solve $Ax = b$ quickly:

Solve $Ly = b$ and $Ux = x$

forward substitution

backward substitution

OUTLINE

- Part I: Gaussian Elimination
- **Part II: Growth Factor & Why Care**
- Part III: History of Question & Recent Results

GROWTH FACTOR

Growth factor g[A] is largest magnitude entry encountered during Gaussian elimination:

$$g[A] := \frac{\max_{i,j,k} |A^{(k)}_{i,j}|}{\max_{i,j} |A_{i,j}|}$$

Idea: three factors control ability to solve $Ax = b$

1) g[A]

2) condition # of A

3) # of bits of precision

SOLVING A SYSTEM

Gaussian elimination produces $(A+E)\hat{x} = b$

with $|E| \le n\mathrm{u}\left(3|A| + 5|\hat{L}||\hat{U}|\right) + O(\mathrm{u}^2)$

$|\hat{L}|$ and $|\hat{U}|$ can be large even if A is well-conditioned:

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

PARTIAL PIVOTING

Partial Pivoting: swap rows so pivot is largest entry in first row

Used by most packages
e.g., MATLAB '\' performs GE with partial pivoting

$$g[A] \leq 2^{n-1}$$

tight for

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

COMPLETE PIVOTING

Complete Pivoting: swap rows & columns so pivot is largest entry in matrix

$$g[A] \le 2\, n^{\ln(n)/4 + 1/2}$$

Certainly not tight.
How close to the truth?

Of great theoretical interest, due to improved growth factor in theory + practice

Is $g[A] = \text{Poly}(n)$?
$g[A] \le n$?

OUTLINE

- Part I: Gaussian Elimination
- Part II: Growth Factor & Why We Should Care
- **Part III: History & Recent Results**

GOLDSTINE & VON NEUMANN

Thought about GE in terms of matrices.

Referred to the use of complete pivoting in solving a linear system as ``customary procedure”

We may therefore interpret the elimination method as … the combination of two tricks: First, it decomposes A into a product of two semi-diagonal matrices … [and second] it forms their inverses by a simple, explicit, inductive process.

— von Neumann and Goldstine [1947, 1053][74]

The onset of World War II accelerated the interest and development of electronic digital computers. In the early days, the computations of artillery-trajectory tables was a laborious process accomplished by many women using manual mechanical calculators. In 1938, the United States established the Ballistics Research Laboratory and John von Neumann (1903-1957) was brought together with Herman Goldstine (1913-) and a few others. Together with their colleagues, von Neumann and Goldstine developed the first digital electronic computer in which both the program and the data resided in the computer's memory.

In 1949 the first new stored-program digital computer went into operation and von Neumann and Goldstine (along with other mathematicians) directed their attention toward understanding the cumulative effect of rounding in computations carried out on their new machines. Attention focused on solving a square system of linear equations using Gaussian elimination. Herman Goldstine later said

> "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."

"History of Gaussian Elimination", Meyer

WILKINSON

Started a rigorous analysis of error in Gaussian elimination, pivoting strategies, and of the growth factor

1961 **Bound**: $g[A] \leq \sqrt{n}\left(2\,3^{1/2}\ldots n^{1/(n-1)}\right)^{1/2} \leq 2\,n^{\ln(n)/4+1/2}$

pessimistic

1965 **Conjecture**: $g[A] \leq n$ for all $n \times n$ real matrices

$g[A] \geq n$ for Hadamard matrices

TREFETHEN & OTHERS

- 1985 "Three mysteries of Gaussian elimination" — Trefethen

  Nonlinear Programming, Stanford Optimization Lab
- 1988 Numerically searching for large growth with NPSOL Library — Day & Peterson
- 1990 Average case analysis of growth —Trefethen & Schreiber
- 1991 Floating point counterexample for $n = 13$ with LANCELOT ($g[A] = 13.0205$) — Gould
- 1992 Counterexample in exact arithmetic for $n = 13$ — Edelman

RECENT PROGRESS

We (Alan Edelman & I) have made progress on three fronts:

- Wilkinson's conjecture is false for all $n \geq 11$ & off by multiplicative constant
- Complexity of growth factor universal over arbitrary entry restrictions
- Growth factor in floating point & exact arithmetic "almost" the same

TECHNICAL LEMMA

For $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_{n-1}) \in \mathbb{R}^{n-1}, \varepsilon_i > -1$, let

$\mathbf{CP}_n^{\varepsilon}(\mathbb{R}) = \{A \in \mathrm{GL}_n(\mathbb{R}) \,|\, |a_{i,j}^{(k)}| \leq (1 + \varepsilon_k) |a_{k,k}^{(k)}|$ for all $i, j \geq k\}$,

e.g., ``almost" completely pivoted matrices
(or, for $\varepsilon_k < 0$, ``overly" completely pivoted)
up to a multiplicative error of ε_k at the k^{th} step of GE

TECHNICAL LEMMA

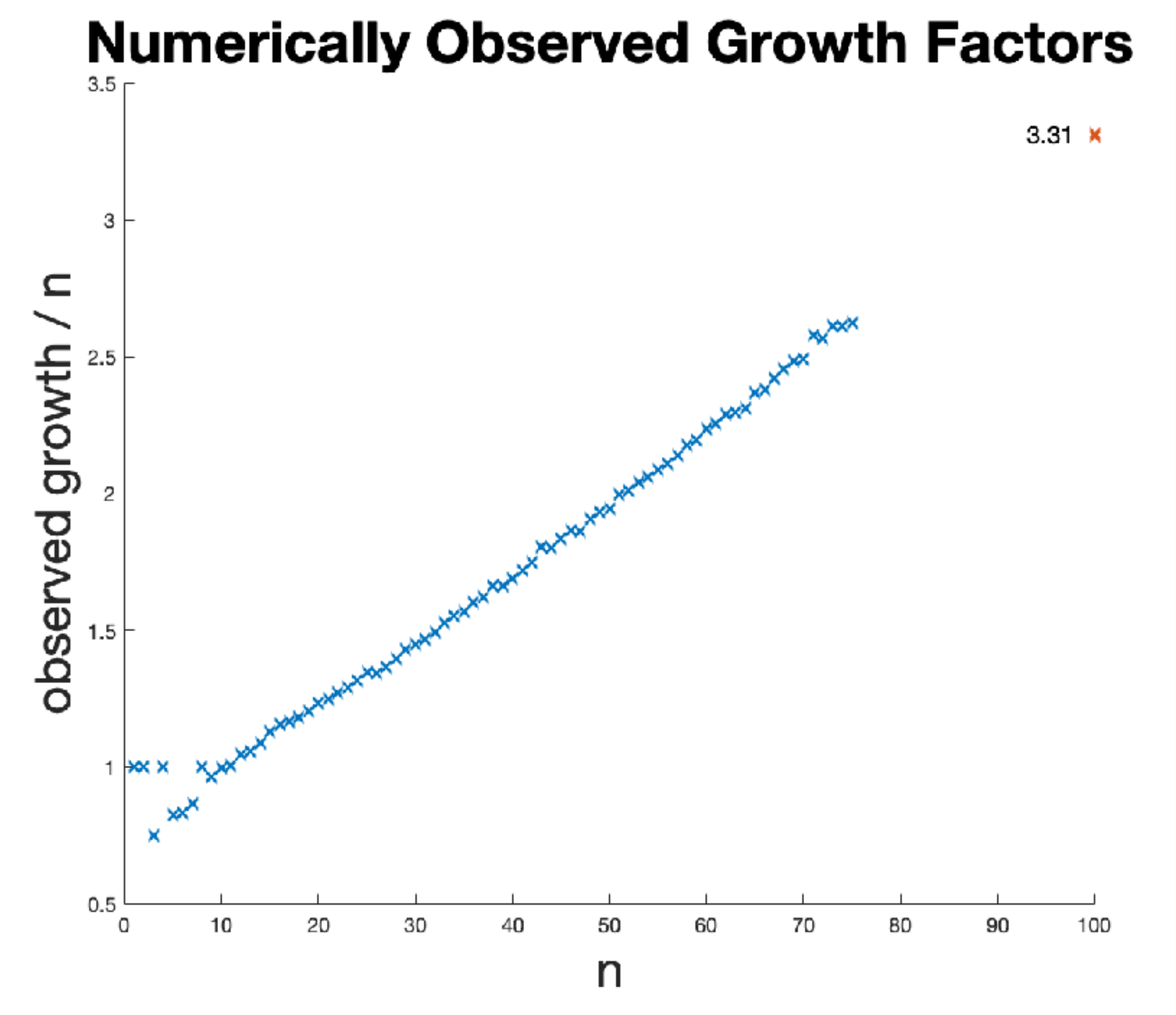
For every $A \in \mathbf{CP}_n^{\varepsilon}(\mathbb{R})$ and $\delta = (\delta_1, \ldots \delta_{n-1})$ satisfying $-1 < \delta_i \leq 0 \leq \varepsilon_i$, exists a matrix $B \in \mathbf{CP}_n^{\delta}(S)$ such that $b_{n,n}^{(k)} = a_{n,n}^{(k)}$ for all $k = 1,\ldots,n$, and:

$$\left| b_{i,j}^{(k)} - a_{i,j}^{(k)} \right| \leq \max_{\min\{i,j\} \leq \ell \leq n-1} \frac{\left[\left(\dfrac{1+\varepsilon_\ell}{1+\delta_\ell} \right)^2 - 1 \right] \left| a_{\ell,\ell}^{(\ell)} \right|}{\prod_{p=\min\{i,j\}}^{\ell-1} 1 + \delta_p} + \sum_{m=\min\{i,j\}}^{\ell-1} \frac{(\varepsilon_m - \delta_m) \left| a_{m,m}^{(m)} \right|}{\prod_{p=\min\{i,j\}}^{m} 1 + \delta_p}.$$

RECENT PROGRESS

- **Wilkinson's conjecture is false for all $n \geq 11$ & off by multiplicative constant**
- Complexity of growth factor universal over arbitrary entry restrictions
- Growth factor in floating point & exact arithmetic "almost" the same

FINDING LARGE GROWTH



Theorem:

$$g\left[\mathbf{CP}_n(\mathbb{R})\right] \geq 1.0045\, n \;\forall\, n \geq 11,$$

$$\limsup_n \left(g\left[\mathbf{CP}_n(\mathbb{R})\right]/n\right) \geq 3.317.$$

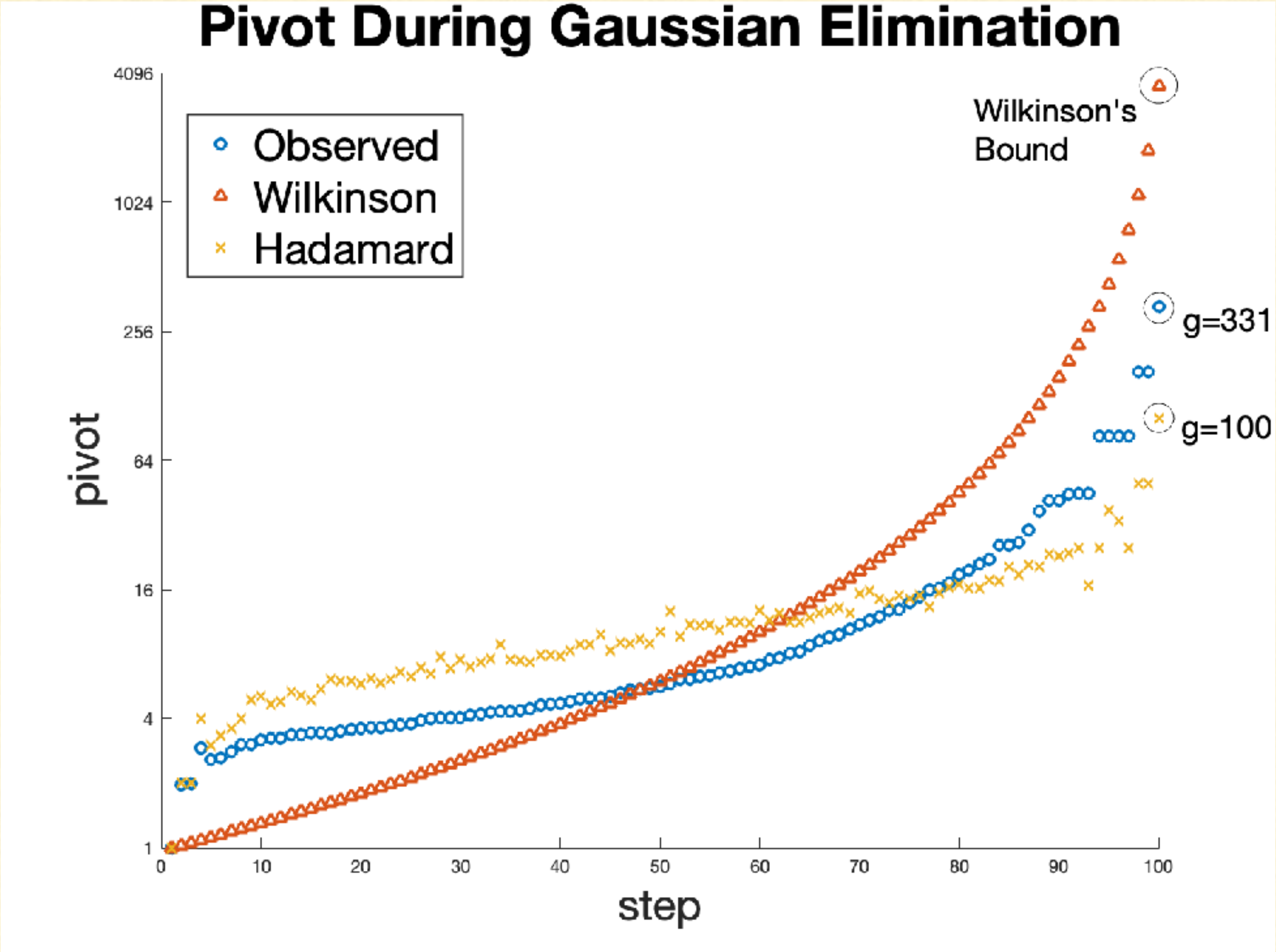
Idea:

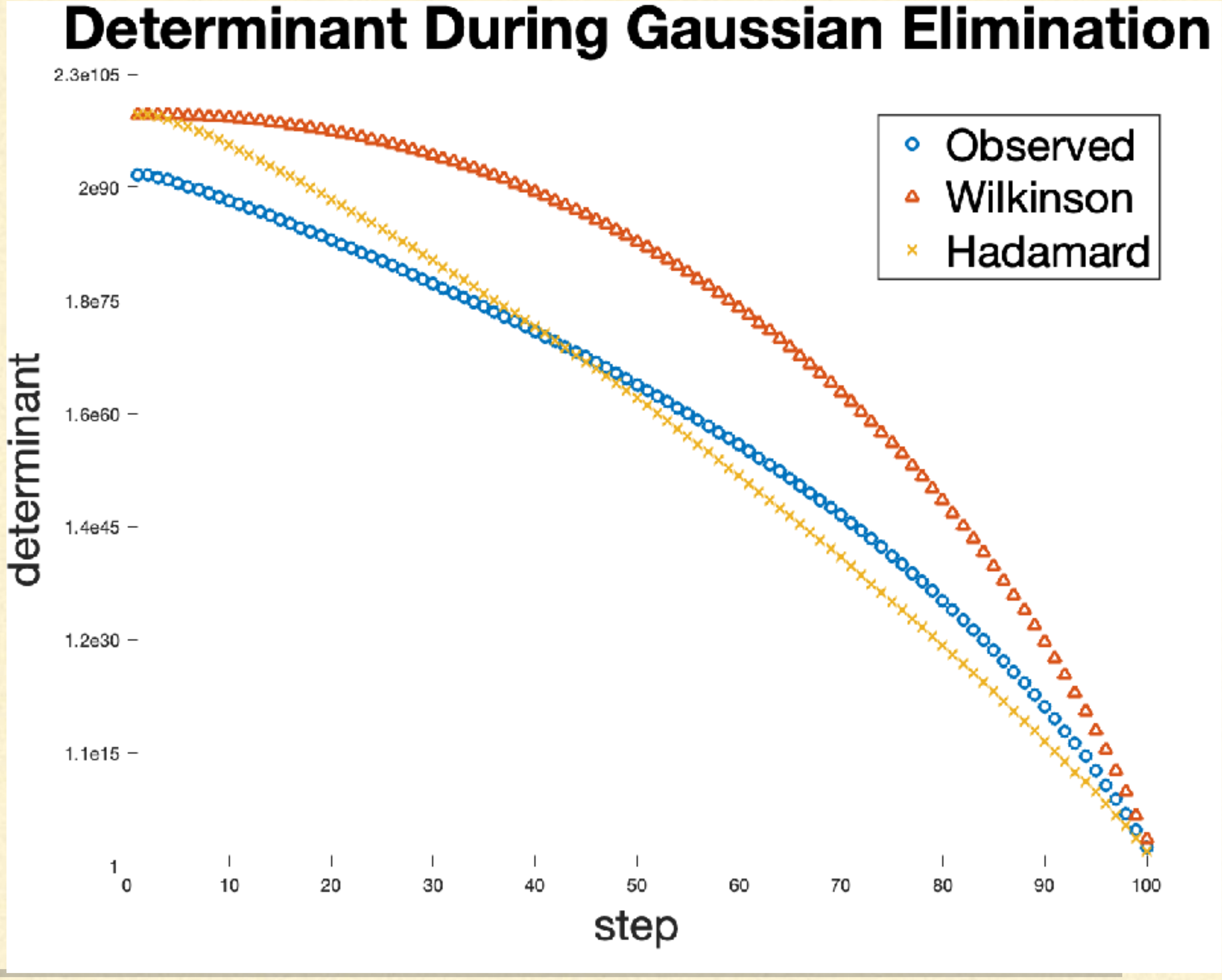
- Technical Lemma
- NL Optimization Software (JuMP + IPOPT)
- Alan's Cluster
- Extrapolation Lemma

TABLE 4. GECP Data computed by JuMP for $n = 1 : 75$ and 100

$n =$	$g \geq \downarrow$	$n =$	$g \geq \downarrow$	$n =$	$g \geq \downarrow$	$n =$	$g \geq \downarrow$	$n =$	$g \geq \downarrow$
1	1	16	18.46	31	45.43	46	85.85	61	137.55
2	2	17	19.86	32	47.74	47	87.54	62	141.83
3	9/4	18	21.25	33	50.36	48	91.44	63	144.72
4	4	19	22.85	34	52.78	49	94.72	64	148.05
5	4.13	20	24.71	35	54.84	50	97.24	65	153.98
6	5	21	26.21	36	57.66	51	101.82	66	157.05
7	6.05	22	28.01	37	59.91	52	104.61	67	162.20
8	8	23	29.72	38	63.18	53	108.09	68	166.89
9	8.69	24	31.63	39	64.87	54	111.19	69	171.33
10	9.96	25	33.67	40	67.52	55	114.76	70	174.45
11	11.05	26	34.96	41	70.44	56	118.18	71	182.98
12	12.55	27	36.88	42	73.49	57	121.90	72	184.91
13	13.76	28	39.05	43	77.68	58	126.23	73	190.57
14	15.25	29	41.46	44	79.25	59	129.42	74	193.28
15	16.92	30	43.40	45	82.56	60	134.27	75	196.79
									⋮
								100	331.71

FINDING LARGE GROWTH





RECENT PROGRESS

- Wilkinson's conjecture is false for all $n \geq 11$ & off by multiplicative constant
- **Complexity of growth factor universal over arbitrary entry restrictions**
- Growth factor in floating point & exact arithmetic "almost" the same

GROWTH FOR RESTRICTED ENTRIES

Theorem: For any $S \subset \mathbb{R}$,
$$g\left[\mathbf{CP}_{14n^2}(S)\right] \geq \left(\mathrm{diam}(S)/2\max(S)\right) g\left[\mathbf{CP}_n(\mathbb{R})\right].$$

meaning: understanding $g[A]$ for any restricted set, e.g., binary matrices, is equivalent (up to poly factor) to understanding $g[A]$ for all matrices

surprising within small poly. factors to find such sweeping results.

RECENT PROGRESS

- Wilkinson's conjecture is false for all $n \geq 11$ & off by multiplicative constant
- Complexity of growth factor universal over arbitrary entry restrictions
- **Growth factor in floating point & exact arithmetic "almost" the same**

EXACT & FLOATING POINT GROWTH

Theorem: Maximum growth factor for real $n \times n$ matrix under floating point arithmetic with base β and mantissa length $t \geq 1 + \log_\beta\left[5n^3 g^2\left[\mathbf{CP}_n(\mathbb{R})\right]\right]$ is at most $(1 + 1/n)\, g\left[\mathbf{CP}_n(\mathbb{R})\right]$.

meaning: $\log^2 n$ bits enough for difference to be negligible, only $\log n$ needed if $g[A] = \text{Poly}(n)$