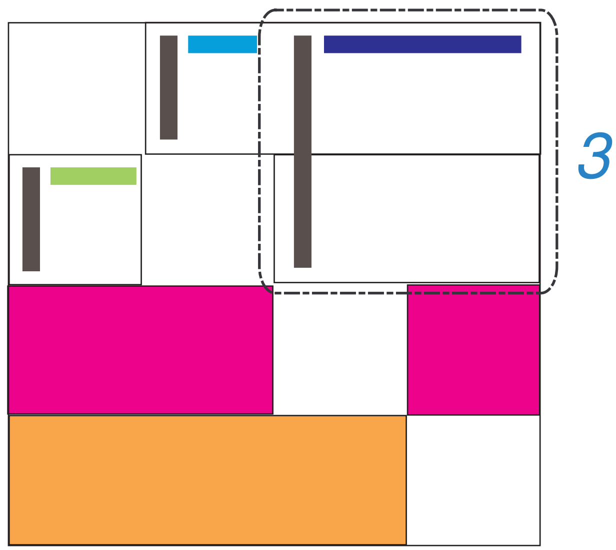
A Numerically Stable Superfast Toeplitz Solver

Ming Gu

Department of Mathematics

University of California, Berkeley

Joint work with J. Xia, Department of Mathematics, Purdue Univ.



Contents

- Introduction: Background and Contributions
- Transform of Toeplitz matrix into Cauchy-like matrix
- Semi-separable matrix Construction for Cauchy-like matrix
- Rank-revealing Factorization for Cauchy-like matrix
- Numerical experiments
- Future work

Introduction: Background

$$T = \begin{pmatrix} t_0 & t_1 & \cdots & t_{n-2} & t_{n-1} \\ t_{-1} & t_0 & t_1 & \ddots & t_{n-2} \\ \vdots & t_{-1} & t_0 & \ddots & \vdots \\ t_{-(n-2)} & \ddots & \ddots & \ddots & t_1 \\ t_{-(n-1)} & t_{-(n-2)} & \cdots & t_{-1} & t_0 \end{pmatrix}.$$

- **Linear system $Tx = b$ involves $O(n)$ parameters.**
- **Applications: Computation of splines; Time series analysis; Markov chains; Queuing theory; Signal and image processing.**
- **First algorithm (Levinson's algorithm) was in 1947.**

Introduction: Algorithm History

METHODS	OPERATIONS	STORAGE
Fast & Stable	$\geq 20n^2$	$\geq n^2/2$
Fast & Unstable	$\geq 3n^2$	$\geq 4n$
"Superfast" & Stable	$O(n^2 + n\log^2 \epsilon)$	$O(n \log^2 \epsilon)$
Superfast & "Unstable"	$O(n \log^2 n)$	$O(n)$
Superfast Preconditioner	$O(n \log n)$	$O(n)$

- **Fast: Levinson-Durbin, Trench ...**
- **Fast stable: Chandrasekaran, Sayed, Gohberg, Kailath, Olshevsky, Gu ...**
- **Superfast: Martinsson, Tygert, Rokhlin, Ammar, Gragg, Stewart, Codevico, van Barel, Heinig, Chandrasekaran, Gu, Xia, Zhu ...**
- **Superfast Preconditioners: Chan, Chan, Strang, Yeung, Di Benedetto, Jin, Kailath, Olshevsky, Ku, Kuo, Strela, Tyrtyshnikov, ...**

Ideal Algorithms: As fast as possible, as reliable as possible.

Introduction: Novel Features of New Algorithm

$C = \left(\dfrac{1}{d_k - f_j}\right)$ **is Cauchy matrix;** $C = \left(\dfrac{u(k,:) \cdot v(j,:)^*}{d_k - f_j}\right)$ **is Cauchy-like.**

- **Transform Toeplitz matrix into Cauchy-like matrix via FFT (not exactly new);**
- **Structured matrix algebra of Cauchy-like matrix**
  - **Compression with Rank-revealing factorization;**
  - **Hierarchical Semi-separable matrix construction;**
- **Additional Structures: Submatrices in all steps are Cauchy-like.**

Transform Toeplitz Matrix into Cauchy-like (I)

$$Z_1 T - T Z_{-1}^T = h \cdot g^T,$$

where $Z_\delta = \begin{pmatrix} 0 & 0 & 0 & \delta \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$**;** h **and** g **have** 2 **columns. Let**

$$Z_1 = Q_1 D_1 Q_1^*, \quad Z_{-1} = Q_{-1} D_{-1} Q_{-1}^*$$

be the eigendecompositions, and let $\widehat{T} = Q_1^* T Q_{-1}$**. Then**

$$D_1 \widehat{T} - \widehat{T} D_{-1} = \widehat{h} \cdot \widehat{g}^*, \ \textbf{for} \ \widehat{h} = Q_1^* h, \ \widehat{g} = Q_{-1}^* g,$$

where

- $D_{-1} = \omega D_1$ **and** $D_1 = \mathrm{diag}(1, \omega^2, \ldots, \omega^{2(n-1)})$ **for** $\omega = e^{\frac{\pi i}{n}}$**.**
- Q_1 **and** Q_{-1} **are known `fft` matrices.**

Transform Toeplitz Matrix into Cauchy-like (II)

$$D_1 \widehat{T} - \widehat{T} D_{-1} = \widehat{h} \cdot \widehat{g}^*,$$

where $D_{-1} = \omega D_1$ **and** $D_1 = \mathrm{diag}(1, \omega^2, \ldots, \omega^{2(n-1)})$ **for** $\omega = e^{\frac{\pi i}{n}}$**.**
$\widehat{T}$ **is Cauchy-like with**

$$\left(\widehat{T}\right)_{k,j} = \frac{\widehat{h}(k,:) * \widehat{g}(j,:)^*}{d_k - f_j}, \quad \textbf{for} \;\; d_k = \omega^{2k-2}, f_j = \omega^{2j-1}.$$

- $Tx = b$ **can be solved via fft as**

$$\widehat{T}\widehat{x} = Q_1^* b, \qquad x = Q_{-1}\widehat{x}.$$

- **Special Cauchy matrix** C **satisfies**

$$C = \left(\frac{1}{d_k - f_j}\right) = \begin{pmatrix} \frac{1}{1-\omega} & \frac{1}{1-\omega^3} & \cdots & \frac{1}{1-\omega^{2n-1}} \\ \frac{1}{\omega^2-\omega} & \frac{1}{\omega^2-\omega^3} & \cdots & \frac{1}{\omega^2-\omega^{2n-1}} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{1}{\omega^{2n-2}-\omega} & \frac{1}{\omega^{2n-2}-\omega^3} & \cdots & \frac{1}{\omega^{2n-2}-\omega^{2n-1}} \end{pmatrix}.$$

- **Approach: Structural Approximation is done on** C

Numerical Low-rank Structure of Cauchy Matrix

Off-diagonal numerical ranks for $n = 2569, \tau = 10^{-9}$**.**

dims	2240×320	1920×640	1600×960	1280×1280	960×1600	640×1920	320×2240
rank	**26**	**28**	**30**	**31**	**30**	**28**	**26**

- **Off-diagonal submatrices have numerical low ranks;**
- **Higher numerical ranks for middle blocks;**
- **Representable as Hierarchical Semi-Separable Matrices**
- **Related work on such matrix structure: Rokhlin, Paige, Chanderasekaran, Gu, Xia ...**
- **Compression through Rank-Revealing Factorization**
- **Related work on Rank-Revealing: Deterministic, randomized, RRQR, RRLU, ...**

Compression by Rank-Revealing Factorization (II)

Let $C = \left(\dfrac{1}{d_k - f_j}\right)$ **be a Cauchy matrix.**
Compute permutations

$$\Pi_1 C_{p\times q} \Pi_2^T = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \begin{matrix} r \\ p-r \end{matrix}$$

to make $|\det(C_{11})|$ **sufficiently large:**

- **Perform** r **steps of GEPP-like elimination;**
- **Interchanging rows** i **and** $r+j$ **of** $C_{p\times q}$ **[Miranian and Gu, 2003]**

$$\frac{\det(\tilde{C}_{11})}{\det(C_{11})} = \left(C_{21}(C_{11})^{-1}\right)_{k,j} = (N)_{k,j}$$

- **Repeat interchanges until** $|N|_{\max} \leq 2$**;**
- **Repeat similar interchanges on columns.**

Compression by Rank-Revealing Factorization (II)

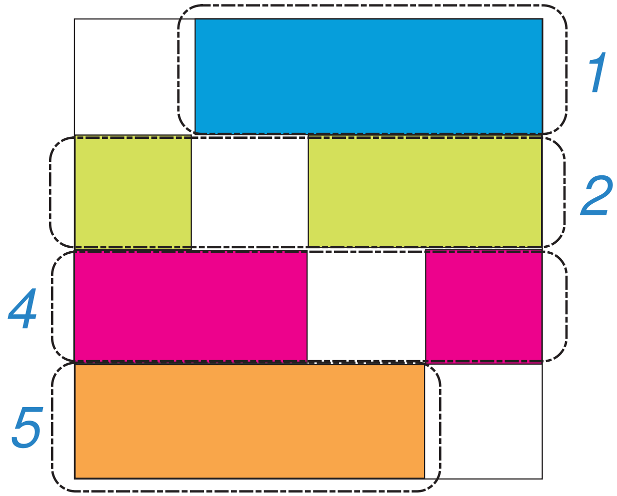
$C = \left(\frac{1}{d_k - f_j}\right)$ **is Cauchy matrix.**

$$\begin{aligned}\Pi_1 C_{p\times q} \Pi_2^T &= \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \begin{matrix} r \\ p-r \end{matrix} \approx \begin{pmatrix} I \\ N \end{pmatrix} \begin{pmatrix} C_{11} & C_{12} \end{pmatrix} \\ &\approx \begin{pmatrix} C_{11} \\ C_{21} \end{pmatrix} \begin{pmatrix} I & \widehat{N} \end{pmatrix}.\end{aligned}$$

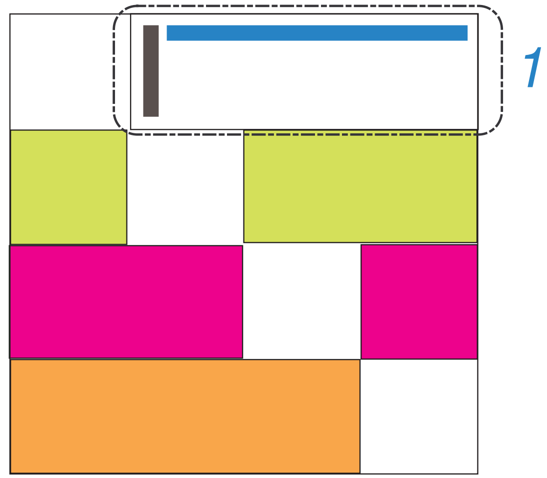
- **Representative cols** $\begin{pmatrix} C_{11} \\ C_{21} \end{pmatrix}$**, representative rows** $\begin{pmatrix} C_{11} & C_{12} \end{pmatrix}$
- $N = \left(\frac{z_i w_j}{d_{r+i} - d_j}\right)_{(p-r)\times r}$ **remains Cauchy-like, so does** $\widehat{N}$**;**
- $\|N\|_{\max} \le 2$ **and** $\left\|\widehat{N}\right\|_{\max} \le 2$**;**
- **Every elimination and interchange step costs** $O(n)$**;**
- **Full relative accuracy in whole Factorization;**
- **Approximation accurate for any tolerance.**

Hierarchical Semi-Separable Matrix (I)

4-blockrow Partition

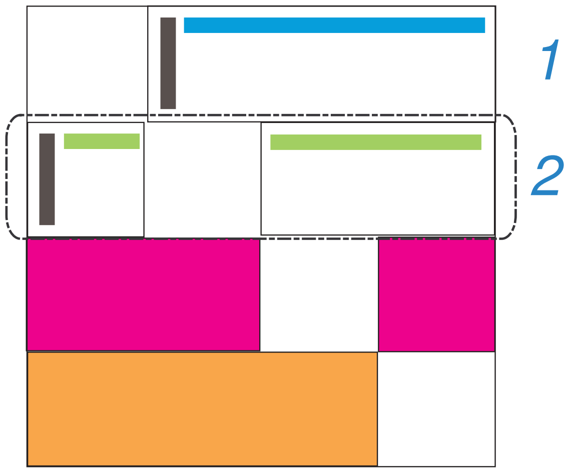


Bottom-level Compression

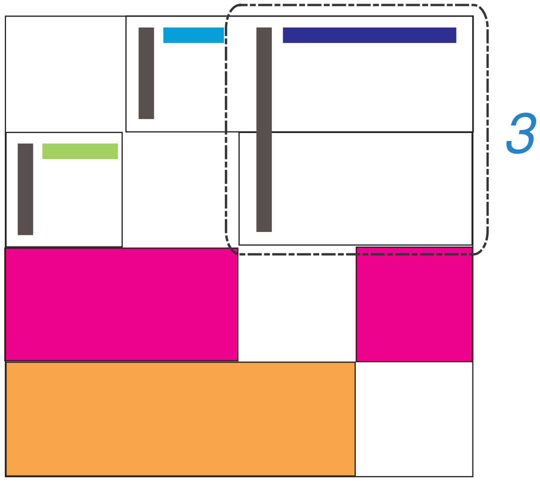


Hierarchical Semi-Separable Matrix (II)

Bottom-level Compression

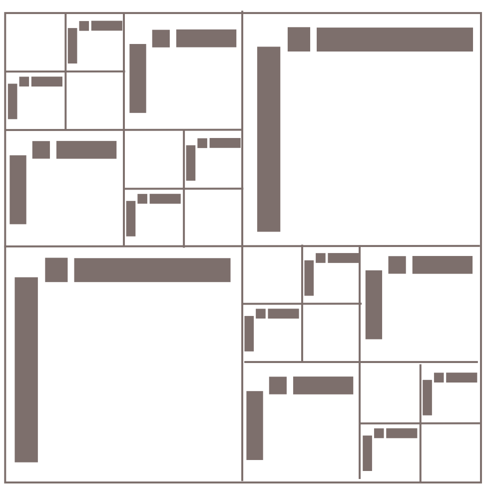


Merge and Compress



Hierarchical Semi-Separable Matrix (III)

HSS Structure



Hierarchical Semi-Separable Matrix (IV)

$$C = \left(\frac{1}{d_k - f_j}\right) = \begin{pmatrix} \frac{1}{1-\omega} & \frac{1}{1-\omega^3} & \cdots & \frac{1}{1-\omega^{2n-1}} \\ \frac{1}{\omega^2-\omega} & \frac{1}{\omega^2-\omega^3} & \cdots & \frac{1}{\omega^2-\omega^{2n-1}} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{1}{\omega^{2n-2}-\omega} & \frac{1}{\omega^{2n-2}-\omega^3} & \cdots & \frac{1}{\omega^{2n-2}-\omega^{2n-1}} \end{pmatrix},$$

where $\omega = e^{\frac{\pi i}{n}}$. Compressions on off-diagonals of C.

- **There are 2^{k-1} block rows at level k; they are multiples of each other;**
- **Total cost for $\log n$ levels: $O(nr \log n)$, where r is numerical rank.**
- **All submatrices from construction are Cauchy-like.**
- **Computations only depend on n.**

Off-diag Compression of Cauchy-like $\widehat{T}$

$$\widehat{T} = \left(\frac{\widehat{h}(k,1:2) * \widehat{g}(j,1:2)^*}{d_k - f_j}\right) \textbf{ and } C = \left(\frac{1}{d_k - f_j}\right).$$

$$\begin{aligned}
C_{p\times q} &= \left(\frac{1}{d_i - f_j}\right)_{p\times q} \approx \begin{pmatrix} I \\ N \end{pmatrix} \begin{pmatrix} C_{11} & C_{12} \end{pmatrix} \quad \textbf{permutations ignored} \\
\widehat{T}_{p\times q} &= \left(\frac{\widehat{h}(k,1:2) * \widehat{g}(j,1:2)^*}{d_k - f_j}\right)_{p\times q} \\
&= \operatorname{diag}(\widehat{h}_{:,1}) C_{p\times q} \operatorname{diag}(\widehat{g}^*_{:,1}) + \operatorname{diag}(\widehat{h}_{:,2}) C_{p\times q} \operatorname{diag}(\widehat{g}^*_{:,2}).
\end{aligned}$$

$\widehat{T}$ has exactly the same semi-separable structure as C, but up to twice the numerical ranks.

Structured Eliminations in Off-diagonal

A block row of $\widehat{T}$

$$\widehat{T}_{1:m,1:n} = \underbrace{\begin{pmatrix} \widehat{T}_{1:r,1:r} & \widehat{T}_{1:r,r+1:m} \\ \widehat{T}_{r+1:m,1:r} & \widehat{T}_{r+1:m,r+1:m} \end{pmatrix}}_{\textbf{diagonal}} \underbrace{\begin{pmatrix} I \\ \widetilde{N} \end{pmatrix} \begin{pmatrix} \widehat{T}_{1:r,m+1:m+r} & \widehat{T}_{1:r,m+r+1:n} \end{pmatrix}}_{\textbf{off-diagonal}}$$

Introducing zeros into the off-diag (applying $\begin{pmatrix} -\tilde{N} & I \\ I & 0 \end{pmatrix}$ on the left)

$$\begin{pmatrix} -\tilde{N} & I \\ I & 0 \end{pmatrix} \widehat{T}_{1:m,1:n}$$

$$= \begin{pmatrix} \widehat{T}_{r+1:m,1:m} - \tilde{N}\widehat{T}_{1:r,1:m} \\ \begin{pmatrix} \widehat{T}_{1:r,1:r} & \widehat{T}_{1:r,r+1:m} \end{pmatrix} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ \widehat{T}_{1:r,m+1:m+r} & \widehat{T}_{1:r,m+r+1:n} \end{pmatrix}$$

Factorization Preserves Cauchy-like Structures

Let the displacement equations for $\tilde{N}$ and $\widehat{T}_{1:r,1:m}$ be

$$\begin{aligned} \operatorname{diag}(D_2)\tilde{N} - \tilde{N}\operatorname{diag}(D_1) &= u_2 w^T, \\ \operatorname{diag}(D_1)\widehat{T}_{1:r,1:m} - \widehat{T}_{1:r,1:m}\operatorname{diag}(F) &= u_1 v^T, \end{aligned}$$

Then $\tilde{N}\widehat{T}_{1:r,1:m}$ satisfies

$$\operatorname{diag}(D_2)(\tilde{N}\widehat{T}_{1:r,1:m}) - (\tilde{N}\widehat{T}_{1:r,1:m})\operatorname{diag}(F) = \begin{pmatrix} u_2 & N u_1 \end{pmatrix} \begin{pmatrix} w^T \widehat{T}_{1:r,1:m} \\ v^T \end{pmatrix}$$

Similar displacement equation exists for $\begin{pmatrix} \widehat{T}_{r+1:m,1:m} - \tilde{N}\widehat{T}_{1:r,1:m} \\ \begin{pmatrix} \widehat{T}_{1:r,1:r} & \widehat{T}_{1:r,r+1:m} \end{pmatrix} \end{pmatrix}$**.**

Numerical Experiments (I)

T**: random; relative tolerance** $\tau = 10^{-9}$

n	$2^2 \times 80$	$2^3 \times 80$	$2^4 \times 80$	$2^5 \times 80$	$2^6 \times 80$	$2^7 \times 80$	$2^8 \times 80$	$2^9 \times 80$
HSS	$1.44e6$	$3.63e6$	$8.02e6$	$1.75e7$	$3.59e7$	$7.34e7$	$1.49e8$	$2.98e8$
HSS$/1000n$	4.50	5.67	6.27	6.84	7.01	7.17	7.08	7.28

Structure construction flops

n	$2^2 \times 80$	$2^3 \times 80$	$2^4 \times 80$	$2^5 \times 80$	$2^6 \times 80$	$2^7 \times 80$	$2^8 \times 80$	$2^9 \times 80$
HSS	$1.46e7$	$3.96e7$	$9.19e7$	$2.07e8$	$4.30e8$	$8.93e8$	$1.83e9$	$3.68e9$
HSS$/10000n$	4.56	6.19	7.18	8.08	8.40	8.72	8.93	8.98

Solution flops. The HSS version ignores Cauchy-like structure for now

Numerical Experiments: Mildly Ill-conditioned

n	$2^2 \times 80$	$2^3 \times 80$	$2^4 \times 80$	$2^5 \times 80$	$2^6 \times 80$	$2^7 \times 80$	$2^8 \times 80$
$\kappa_2(T)$	$2.05e6$	$3.11e6$	$5.32e6$	$3.02e7$	$1.40e9$	$5.79e9$	
Soln accuracy	$3.03e{-14}$	$5.82e{-14}$	$1.24e{-11}$	$2.45e{-11}$	$1.49e{-10}$	$2.99e{-10}$	$5.98e{-10}$
# of it. ref. steps	1	1	1	1	1	1	1
New accuracy	$2.38e{-16}$	$3.32e{-16}$	$5.05e{-16}$	$6.84e{-16}$	$9.83e{-16}$	$1.40e{-15}$	$1.94e{-15}$

Condition number and accuracy $\frac{||Tx-b||_2}{|||T||x|+|b|||_2}$

Numerical Experiments: More Ill-Conditioned

T: **generated with** $1 + \texttt{randn}(n, 1) \times 10^{-9}$; **relative tolerance** $\tau = 10^{-16}$

n	$2^2 \times 80$	$2^3 \times 80$	$2^4 \times 80$	$2^5 \times 80$	$2^6 \times 80$	$2^7 \times 80$
$\kappa_2(T)$	$1.99e12$	$1.42e13$	$1.17e13$	$2.25e13$	$2.59e13$	$8.87e13$
Soln accuracy	$3.05e{-}14$	$5.77e{-}14$	$1.24e{-}11$	$2.45e{-}11$	$1.49e{-}10$	$2.99e{-}10$
# of it. ref. steps	1	1	1	1	1	2
New accuracy	$2.38e{-}16$	$3.36e{-}16$	$5.00e{-}16$	$8.74e{-}16$	$2.12e{-}15$	$7.35e{-}15$

T: **generated with** $1 + \texttt{randn}(n, 1) \times 10^{-15}$; **relative tolerance** $\tau = 10^{-16}$

n	$2^2 \times 80$	$2^3 \times 80$	$2^4 \times 80$	$2^5 \times 80$	$2^6 \times 80$	$2^7 \times 80$
$\kappa_2(T)$	$1.29e19$	$1.46e20$	$5.01e19$	$4.89e19$	$3.56e20$	$6.89e20$
Soln accuracy	$1.33e{-}14$	$4.04e{-}14$	$1.15e{-}12$	$1.99e{-}11$	$8.08e{-}11$	$1.01e{-}10$
# of it. ref. steps	1	1	1	3	3	1
New accuracy	$5.33e{-}16$	$3.07e{-}15$	$7.86e{-}13$	$7.29e{-}13$	$3.79e{-}12$	$4.83e{-}11$

Numerical Experiments: Prolate Matrices

T: **Prolate matrix, generated with** $\frac{\sin(2k\omega\pi)}{k\pi}$, $\omega = 0.25$; **relative** $\tau = 10^{-16}$

n	$2^2 \times 80$	$2^3 \times 80$	$2^4 \times 80$	$2^5 \times 80$	$2^6 \times 80$	$2^7 \times 80$
$\kappa_2(T)$	$8.76e17$	$3.23e17$	$2.87e18$	$9.60e17$	$1.04e18$	
Soln accuracy	$2.17e-15$	$4.17e-15$	$6.54e-12$	$1.12e-11$	$5.95e-11$	$8.32e-11$
# of it. ref. steps	1	1	1	1	1	1
New accuracy	$3.98e-16$	$4.42e-16$	$1.95e-16$	$1.29e-15$	$2.639e-12$	$8.22e-11$

Condition number and accuracy $\frac{||Tx-b||_2}{|||T||x|+|b|||_2}$

Numerical Experiments: Highly Ill-conditioned

T **generated by (RBF with uniform mesh)**

$$\phi(x;\alpha,h) = \mathrm{sech}(\frac{\alpha}{h}x), \alpha = \frac{1}{16}, \ \ \kappa_\infty(T) = \frac{1}{4}\exp(\frac{\pi^2}{2\alpha}) = 4.88 \times 10^{33}$$

n	$2^2 \times 80$	$2^3 \times 80$	$2^4 \times 80$	$2^5 \times 80$	$2^6 \times 80$	$2^7 \times 80$
Factorization	$2.85e6$	$7.70e6$	$1.76e7$	$3.87e7$	$8.15e7$	$1.66e8$
Solution	$3.79e7$	$1.19e8$	$2.89e8$	$6.80e8$	$1.49e9$	$3.08e9$

Flop counts for structure construction and system solution, $\tau = 10^{-16}$

n	$2^2 \times 80$	$2^3 \times 80$	$2^4 \times 80$	$2^5 \times 80$	$2^6 \times 80$	$2^7 \times 80$
Soln accuracy	$2.82e{-16}$	$1.97e{-16}$	$7.43e{-12}$	$8.82e{-12}$	$8.35e{-12}$	$4.35e{-11}$
# of it. ref. steps			1	1	1	1
New accuracy			$3.24e{-16}$	$3.10e{-16}$	$3.81e{-15}$	$1.57e{-15}$

Accuracy ($\frac{\|Ax-b\|}{\||A||x|+|b|\|}$) *and the accuracy after iterative refinement*

Concluding Remarks

- **Structured rank-revealing LU**
- **All compressions and factorizations in terms of Cauchy-like blocks**
- **(Theoretical) worst-case element growth factor** $= O\left(n^{O(\log r)}\right)$**.**
- **More improvements in the works.**
- **Toeplitz least squares problems to be done**
- **Similar schemes for block Toeplitz/Toeplitz-like problems**