Replacing Pivoting in Distributed Gaussian Elimination with Randomized Techniques

Neil Lindquist, Piotr Luszczek, Jack Dongarra

{nlindqu1,luszczek,dongarra}@icl.utk.edu

Gaussian Elimination

- Solve $Ax = b$
- Factor $PA = LU$, then $x = U^{-1}L^{-1}Pb$
- Pivoting
  - Important for accuracy
  - Introduces expensive data movement

⇒ Replace pivoting with randomization

INNOVATIVE COMPUTING LABORATORY
THE UNIVERSITY OF TENNESSEE KNOXVILLE

Butterfly Matrix

$$B^{\langle n \rangle} = \frac{1}{\sqrt{2}} \begin{bmatrix} R_1 & R_2 \\ R_1 & -R_2 \end{bmatrix}$$

R_1, R_2 - diagonal, nonsingular matrices

ICL INNOVATIVE COMPUTING LABORATORY
THE UNIVERSITY OF TENNESSEE KNOXVILLE

Recursive Butterfly Transform

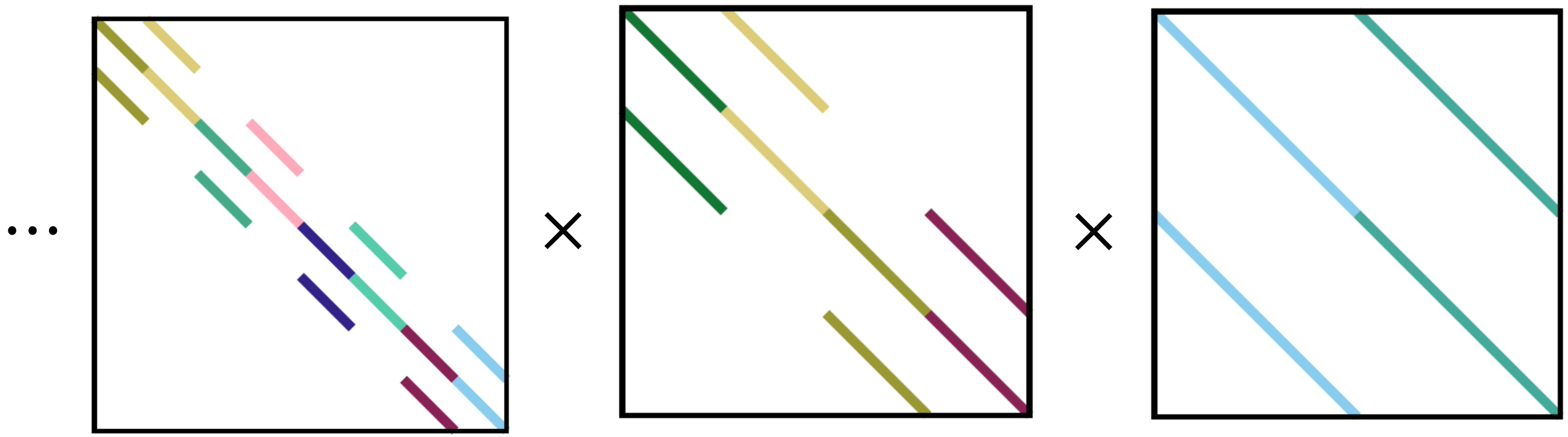
$$\begin{bmatrix} B_1^{\left\langle \frac{n}{2^{d-1}} \right\rangle} & & \\ & \ddots & \\ & & B_{2^{d-1}}^{\left\langle \frac{n}{2^{d-1}} \right\rangle} \end{bmatrix} \cdots \begin{bmatrix} B_1^{\langle n/2 \rangle} & \\ & B_2^{\langle n/2 \rangle} \end{bmatrix} B_1^{\langle n \rangle}$$

INNOVATIVE COMPUTING LABORATORY

THE UNIVERSITY OF TENNESSEE KNOXVILLE

Recursive Butterfly Transform

$$\begin{bmatrix} B_1^{\left\langle \frac{n}{2^{d-1}} \right\rangle} & & \\ & \ddots & \\ & & B_{2^{d-1}}^{\left\langle \frac{n}{2^{d-1}} \right\rangle} \end{bmatrix} \cdots \begin{bmatrix} B_1^{\langle n/2 \rangle} & \\ & B_2^{\langle n/2 \rangle} \end{bmatrix} B_1^{\langle n \rangle}$$



ICL INNOVATIVE COMPUTING LABORATORY
THE UNIVERSITY OF TENNESSEE KNOXVILLE

Relation to the Fast Fourier Transform

- FFT is a RBT followed by a permutation

$$B^{\langle n \rangle} = \frac{1}{\sqrt{2}} \begin{bmatrix} I & \Omega \\ I & -\Omega \end{bmatrix}$$

$$\Omega = \mathrm{diag}(1, \omega_{2n}, \omega_{2n}^2, \ldots, \omega_{2n}^{n-1})$$

INNOVATIVE COMPUTING LABORATORY

THE UNIVERSITY OF TENNESSEE KNOXVILLE

Overheads

- $4dn^2$ FLOP to apply a 2-sided RBT
- $2dn$ FLOP to apply an RBT to a vector

- dn elements to store

INNOVATIVE COMPUTING LABORATORY

THE UNIVERSITY OF TENNESSEE KNOXVILLE

RBT-based Solver

$\mathcal{U}, \mathcal{V}$ – recursive butterfly transforms

Write $Ax = b$ as $(\mathcal{U}^T A \mathcal{V})(\mathcal{V}^{-1} x) = (\mathcal{U}^T b)$

1. $A' = \mathcal{U}^T A \mathcal{V}$
2. $b' = \mathcal{U}^T b$
3. $LU = A'$
4. $x' = U^{-1} L^{-1} b'$
5. $x = \mathcal{V} x'$

INNOVATIVE COMPUTING LABORATORY

THE UNIVERSITY OF TENNESSEE KNOXVILLE

Previous Work

- Proposed: Parker, 1995
- Practical experiments: Baboulin et al, 2012-2016
  - GPU-accelerated, symmetric/non-symmetric
  - Distributed, symmetric
  - Sparse

ICL INNOVATIVE COMPUTING LABORATORY
THE UNIVERSITY OF TENNESSEE KNOXVILLE

Implementation Details

- Using SLATE (Software for Linear Algebra Targeting Exascale)
  - Distributed, GPU-accelerated, dense linear algebra
  - Part of DoE Exascale Computing Project
- Recursive transform depth of 2
  - Matrix size must be divisible by 4
- 1 step iterative refinement

INNOVATIVE COMPUTING LABORATORY
THE UNIVERSITY OF TENNESSEE KNOXVILLE

Experiment Configuration

- 8 nodes of Summit
- Each node
  - 2 processes
  - 2 22-core IBM POWER 9 CPUs
  - 6 NVIDIA Volta V100 GPUs
- Double precision reals
- Spectrum MPI 10.3.1.2, ESSL 6.1.0-2
- GCC 8.1.1, CUDA 10.1.243

ICL INNOVATIVE COMPUTING LABORATORY
THE UNIVERSITY OF TENNESSEE KNOXVILLE

Accuracy results

$$n = 100\,000, \ \text{error} = \frac{\|r\|_1}{\|A\|_1 \cdot \|x\|_1}$$

	Partial pivoting	RBT Solver Refined	RBT Solver	No pivoting Refined	No pivoting
Random [0,1]	1.23×10^{-15}	2.97×10^{-17}	6.43×10^{-12}	2.67×10^{-17}	4.10×10^{-12}
Random [-1, 1]	2.39×10^{-15}	2.93×10^{-17}	1.53×10^{-11}	1.19×10^{-15}	8.76×10^{-11}
Random Normal	1.77×10^{-15}	3.29×10^{-17}	6.68×10^{-12}	3.23×10^{-17}	1.71×10^{-11}
Random {0,1}	1.82×10^{-15}	2.25×10^{-17}	6.15×10^{-12}	NA	NA
fielder	3.03×10^{-18}	2.92×10^{-19}	1.73×10^{-17}	NA	NA
orthog	2.29×10^{-16}	9.19×10^{-3}	1.00×10^{-2}	1.21×10^{-1}	1.30×10^{-1}
gfpp	NA	2.79×10^{-19}	5.06×10^{-18}	NA	NA

ICL INNOVATIVE COMPUTING LABORATORY

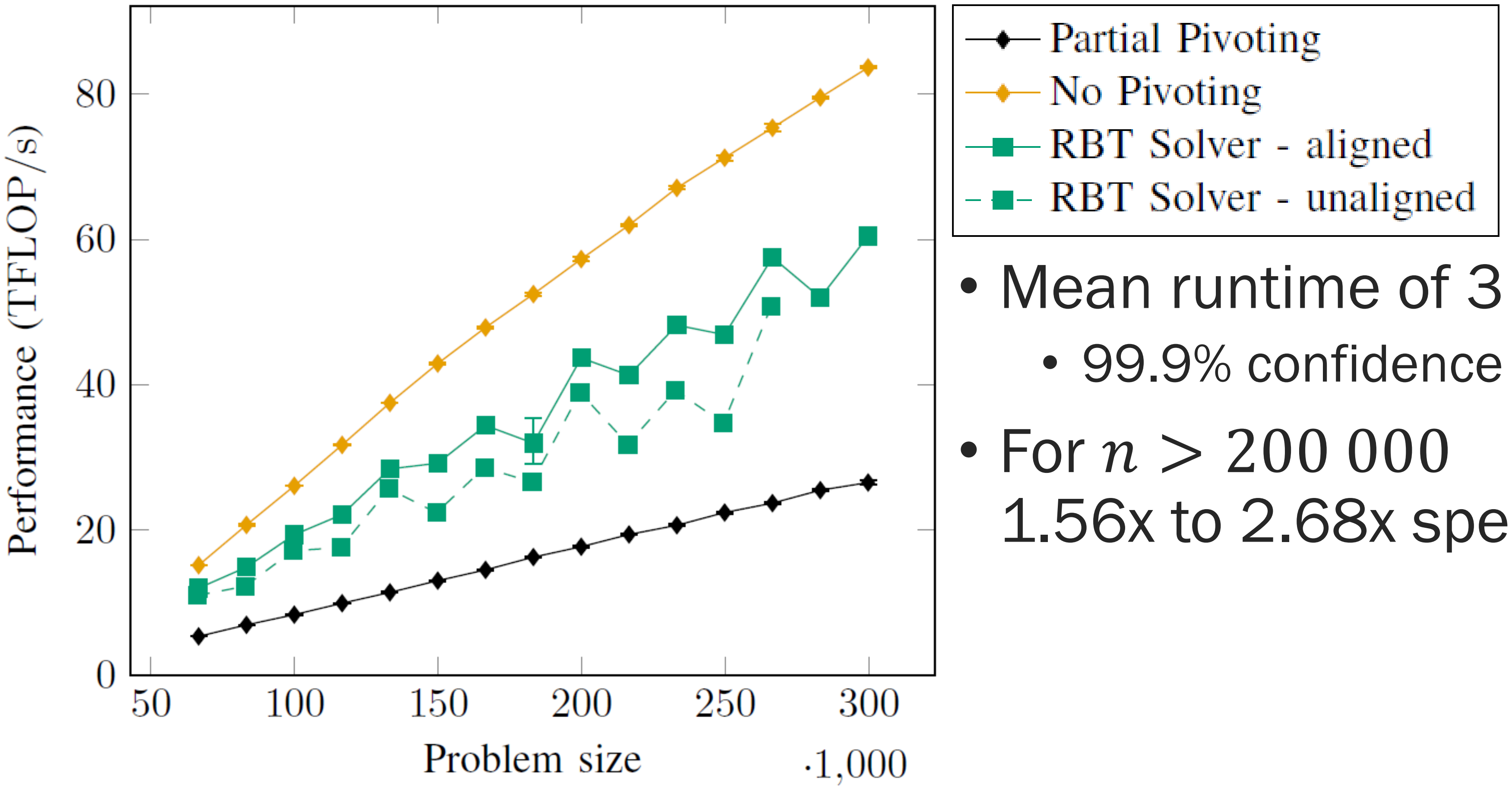
THE UNIVERSITY OF TENNESSEE KNOXVILLE

Pivoting fallback

- Revert to pivoting if not pivoting fails
- Either
  - Check backwards error at end of solve
  - Monitor growth factor while factoring

ICL INNOVATIVE COMPUTING LABORATORY
THE UNIVERSITY OF TENNESSEE KNOXVILLE

Performance results



- Mean runtime of 3 tests
  - 99.9% confidence interval
- For $n > 200\,000$
  1.56x to 2.68x speedup

Acknowledgements

- This material is based upon work partially supported by the National Science Foundation under OAC Grant No. 2004541.
- This research used the computational resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725, provided by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration.

ICL INNOVATIVE COMPUTING LABORATORY

THE UNIVERSITY OF TENNESSEE KNOXVILLE

ICL is hiring!

- ICL is offering exciting opportunities for Research Scientists, Post Docs, and Software Developers. Join us!
- Scan the QR code below or go to https://www.icl.utk.edu/jobs for more information.

ICL INNOVATIVE COMPUTING LABORATORY

THE UNIVERSITY OF TENNESSEE KNOXVILLE

Conclusions

- Recursive Butterfly Transforms can replace pivoting in Gaussian Elimination
  - Often as accurate
  - 1.56x to 2.68x speedup

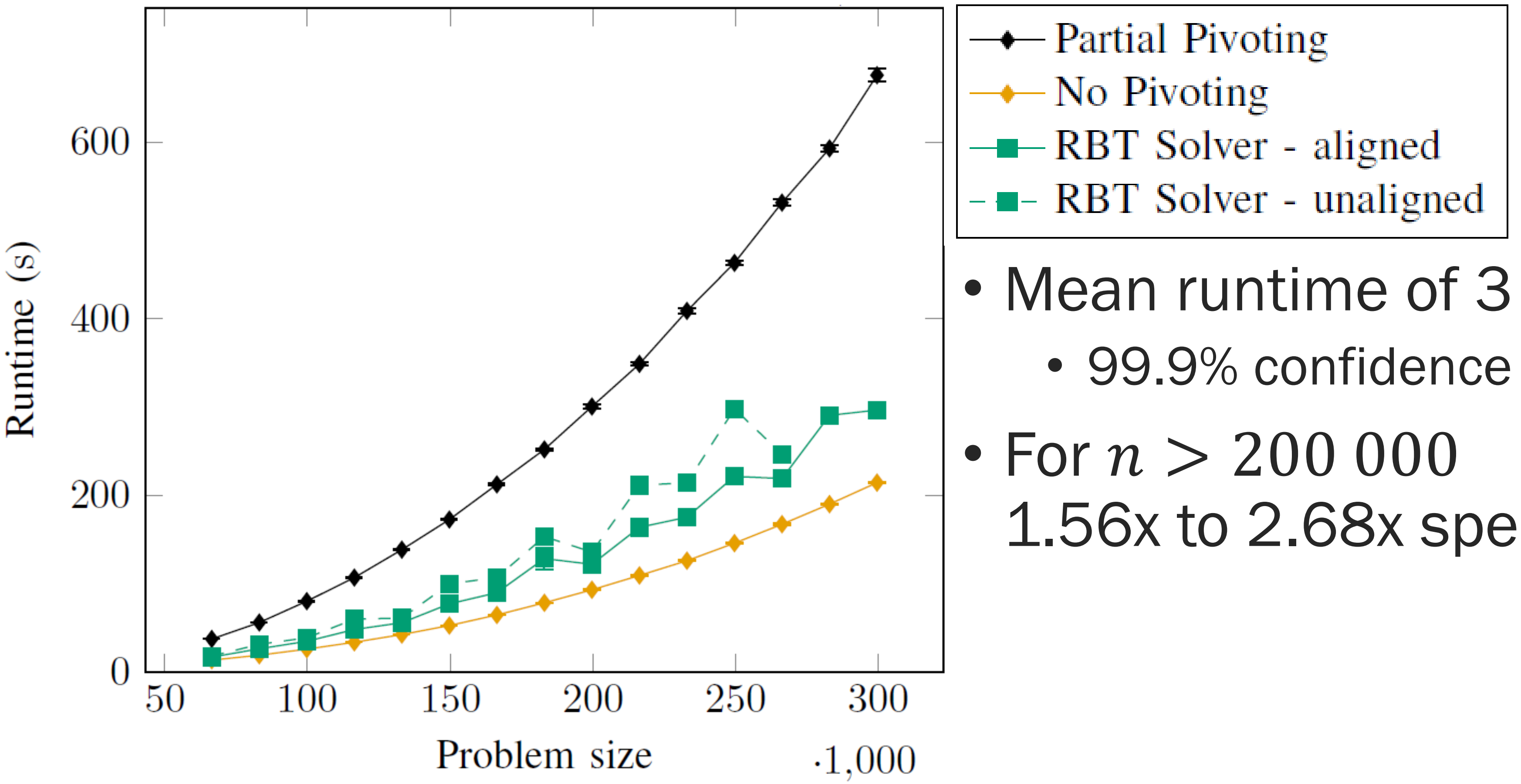
INNOVATIVE COMPUTING LABORATORY

THE UNIVERSITY OF TENNESSEE KNOXVILLE

ICL
INNOVATIVE
COMPUTING LABORATORY

THE UNIVERSITY OF
TENNESSEE
KNOXVILLE

Performance results



- Mean runtime of 3 tests
    - 99.9% confidence interval
- For $n > 200\,000$
  1.56x to 2.68x speedup