Communication avoiding algorithms for iterative methods

Computational Kernels for Preconditioned Iterative Methods

- Sparse matrix - set of vectors multiplication
- Sparse low rank approximations

Iterative methods

- Enlarge Krylov methods to reduce communication
- Efficiency on selected applications

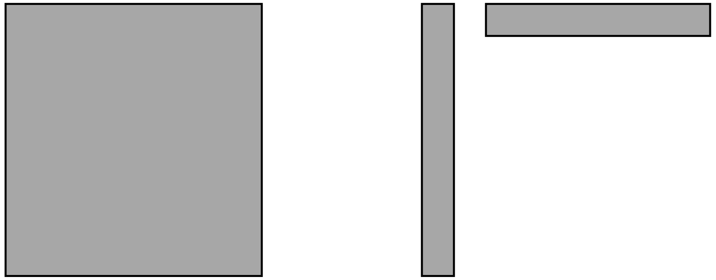
Preconditioners

- Robust multilevel preconditioners based on low rank corrections
- Combine with enlarge Krylov methods

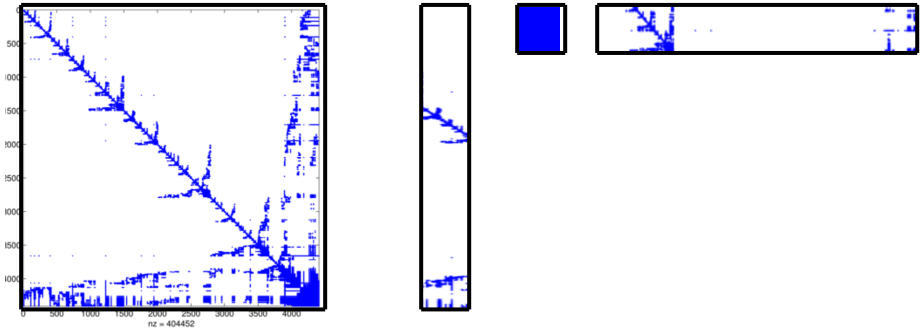
Low rank matrix approximation

- Problem: given m x n matrix A, compute rank-k approximation ZW^T, where Z is m x k and W^T is k x n.

$$A \approx Z \; W^T$$



$$A \approx C \; U \; R$$

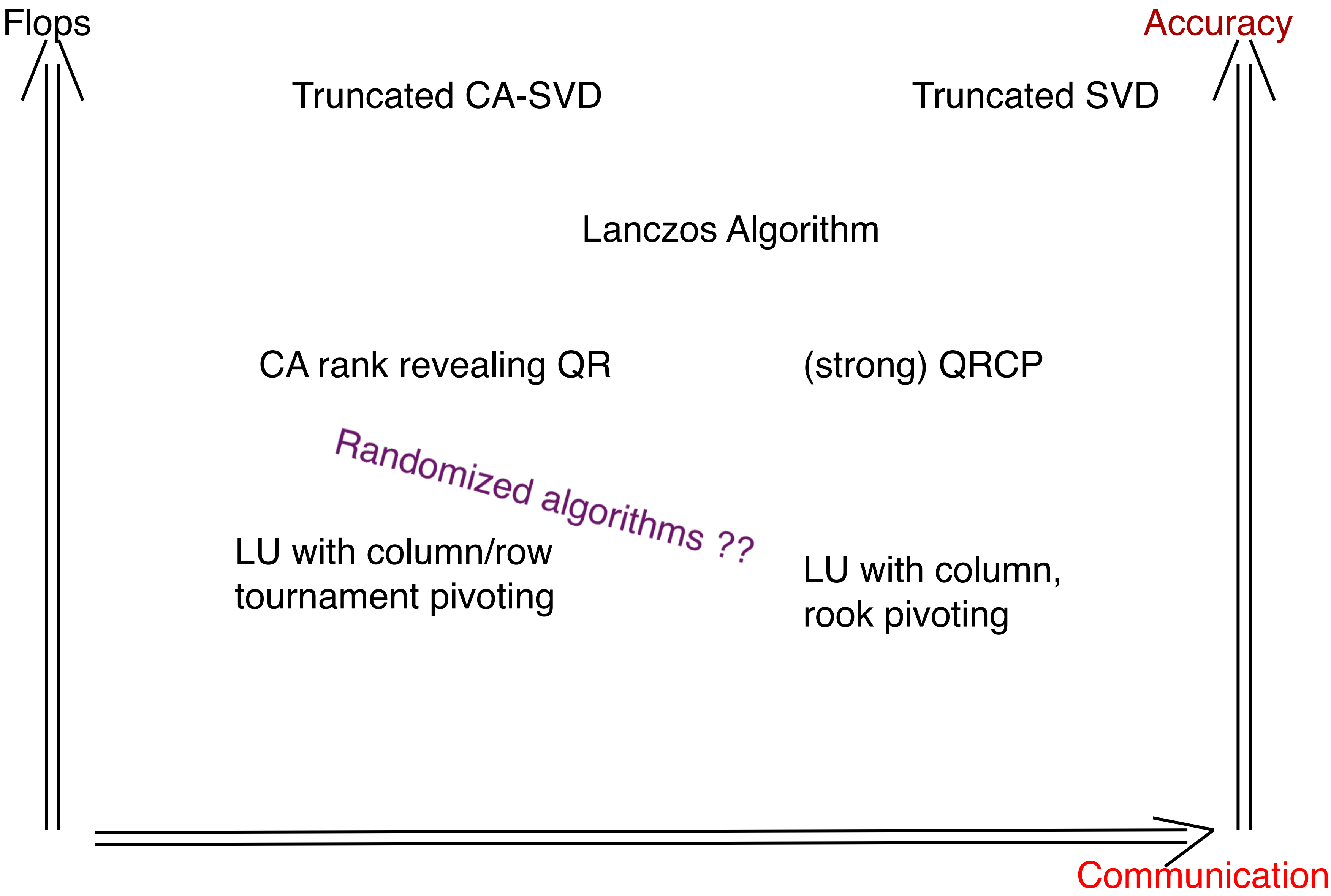


- Problem with diverse applications
  - from scientific computing: fast solvers for integral equations, H-matrices
  - to data analytics: principal component analysis, image processing, …

- Used in iterative process by multiplication with a set of vectors

$$\begin{aligned} & Ax && \rightarrow && ZW^T x \\ \text{Flops:}\quad & 2mn && \rightarrow && 2(m+n)k \end{aligned}$$

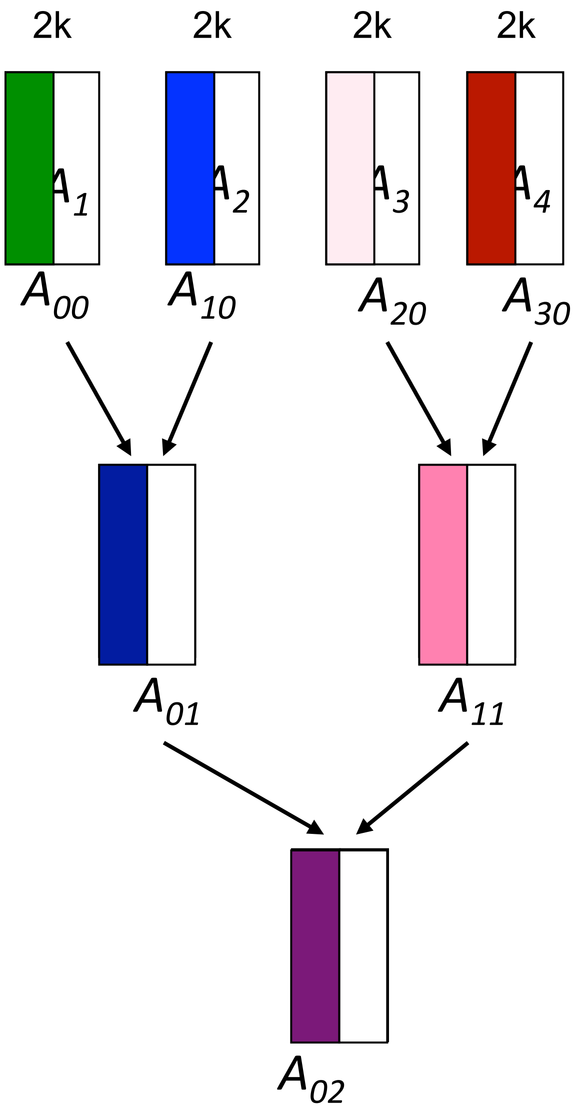
Low rank matrix approximation: trade-offs



Select k cols using tournament pivoting

Partition $A=(A_1, A_2, A_3, A_4)$.
Select k cols from each column block,
by using QR with column pivoting

At each level i of the tree
At each node j do in parallel
Let $A_{v,i-1}$, $A_{w,i-1}$ be the cols selected by
the children of node j
Select b cols from $(A_{v,i-1}, A_{w,i-1})$,
by using QR with column pivoting
Return columns in A_{ji}



LU_CRTP

- Given LU_CRTP factorization

$$P_r A P_c = \begin{pmatrix} \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{22} \end{pmatrix} = \begin{pmatrix} I & \\ \bar{A}_{21}\bar{A}_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} \bar{A}_{11} & \bar{A}_{12} \\ & S(\bar{A}_{11}) \end{pmatrix},$$

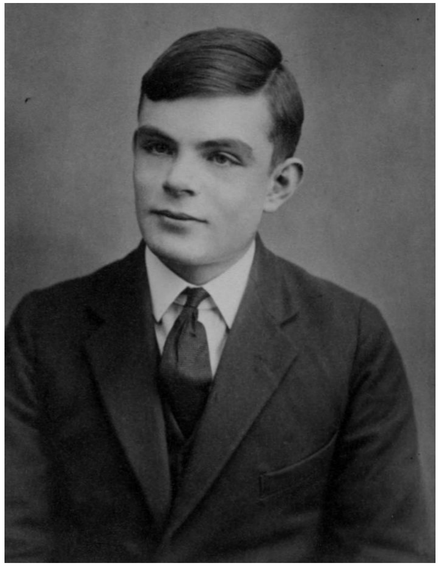
the rank - k CUR approximation is

$$\tilde{A}_k = \begin{pmatrix} I \\ \bar{A}_{21}\bar{A}_{11}^{-1} \end{pmatrix} \begin{pmatrix} \bar{A}_{11} & \bar{A}_{12} \end{pmatrix} = \begin{pmatrix} \bar{A}_{11} \\ \bar{A}_{21} \end{pmatrix} \bar{A}_{11}^{-1} \begin{pmatrix} \bar{A}_{11} & \bar{A}_{12} \end{pmatrix}$$

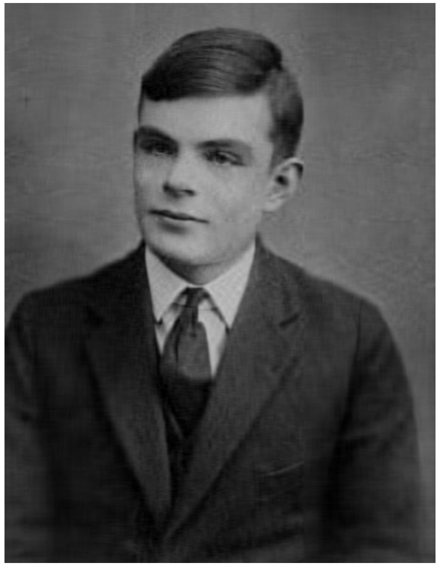
- $\bar{A}_{11}^{-1}$ is never formed, its factorization is used when $\tilde{A}_k$ is applied to a vector
- In randomized algorithms, U = $C^+ A R^+$, where C^+, R^+ are Moore-Penrose generalized inverses

Results for image of size 256x707

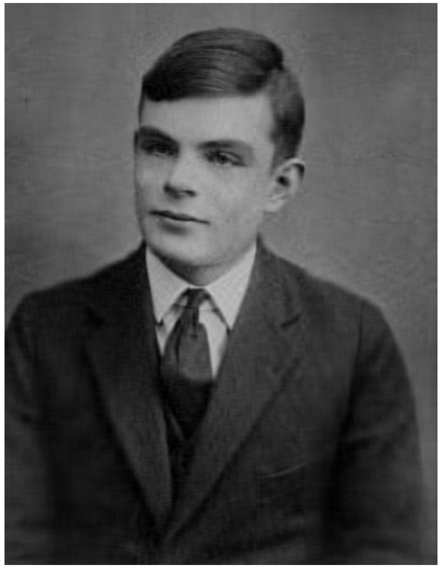
Original image, 256x707



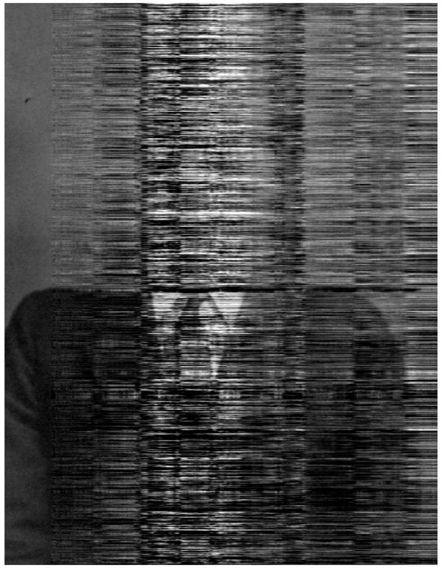
SVD: Rank-38 approximation



Rank-75 approximation, SVD

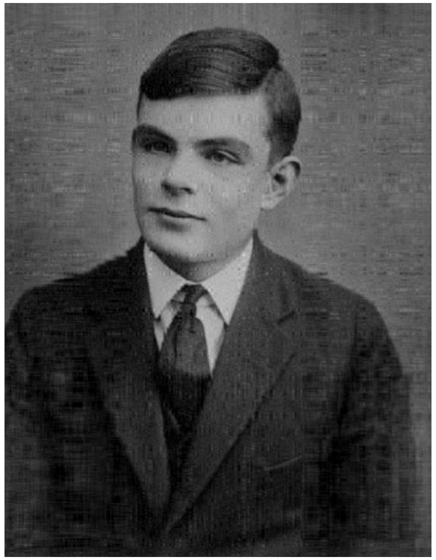


LUPP: Rank-75 approximation



LU_CRTP: Rank-38 approx.

LU_CRTP: Rank-75 approx.



Performance results

Selection of 256 columns by tournament pivoting

Edison, Cray XC30 (NERSC) – 2x12-core Intel Ivy Bridge (2.4 GHz)

Tournament pivoting uses SPQR (T. Davis) + dGEQP3 (Lapack), time in secs

Matrices:	n x n	n x n/32
• Parab_fem:	528825 x 528825	528825 x 16432
• Mac_econ:	206500 x 206500	206500 x 6453

	Time n x 2k	Time n x n/32 SPQR+GEQP3	Number of MPI processes 16	32	64	128	256	512	1024
Parab_fem	0.26	0.26+1129	46.7	24.5	13.7	8.4	5.9	4.8	4.4
Mac_econ	0.46	25.4+510	132.7	86.3	111.4	59.6	27.2	-	-