Parallel, Blocked and Multishift Variants of the QZ Algorithm for Regular Matrix Pairs

Björn Adlerborn and Bo Kågström
{c95ban, bokg}@cs.umu.se
Dept of Computing Science and HPC2N
Umeå University, Sweden

Daniel Kressner
Institut für Mathematik, Technische Universität Berlin
kressner@math.tu-berlin.de

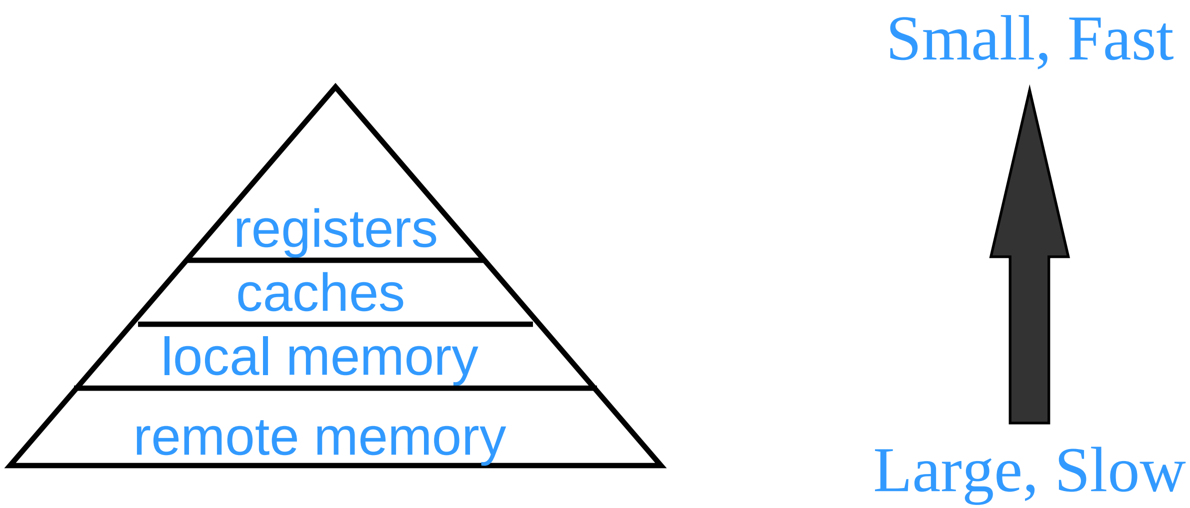
SIAM Applied Linear Algebra '03, July 15-19, Williamsburg

Motivation (1)

- Want to solve large-scale dense nonsymmetric generalized eigenvalue problems

$$Ax = sBx$$

on today's HPC systems with deep memory hierarchies



Motivation (2)

Typically, we want

- all eigenvalues and eigenvectors
- or a subset of the eigenvalues with associated left and right eigenspaces
- high accuracy results

Our approach:

- orthogonal transformation methods
- blocked and parallel (distributed) algorithms ("strive for level 3")

Regular Nonsymmetric $Ax = sBx$

method

Transform (A, B) to generalized Schur form (S, T) using orthogonal equivalence transformations

S quasi-triangular (1x1 and 2x2 blocks on the diagonal)

T upper triangular

- Real eigenvalues are given by the diagonal elements (s_{ii}, t_{ii})
  - Finite eigenvalues s_{ii}/ t_{ii}
  - Infinite eigenvalues represented as $(s_{ii}, 0)$
- 2x2 blocks correspond to complex conjugate eigenvalue pairs

Solve Matrix Pencil Systems for several r.h.s. and s

$(A - s B) x = y$ <==> $(H - s T) u = v$ Hessenberg

Outline

- Three-stage reduction to generalized Schur form

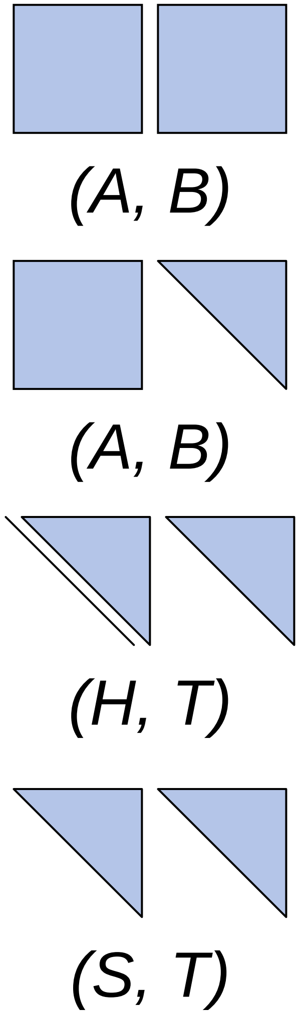
$$(A,B) \xrightarrow{\text{Stage 1}} (H_r,T) \xrightarrow{\text{Stage 2}} (H,T) \xrightarrow{\text{Stage 3}} (S,T)$$

- Stage 1 & 2: Blocked and Parallel Algorithms
- Stage 3: Blocked QZ Algorithm
- Multishift QZ variants
- Aggressive Early Deflation in QZ
- Computational Experiments
- Stage 3: Parallel QZ Algorithm
- Summary and Some References

VETENSKAPSRÅDET
THE SWEDISH RESEARCH COUNCIL
Swedish Foundation for Strategic Research

Unblocked Solution

(presently in LAPACK, based on Moler-Stewart'72)



Original regular $N \times N$ matrix pair (A, B)

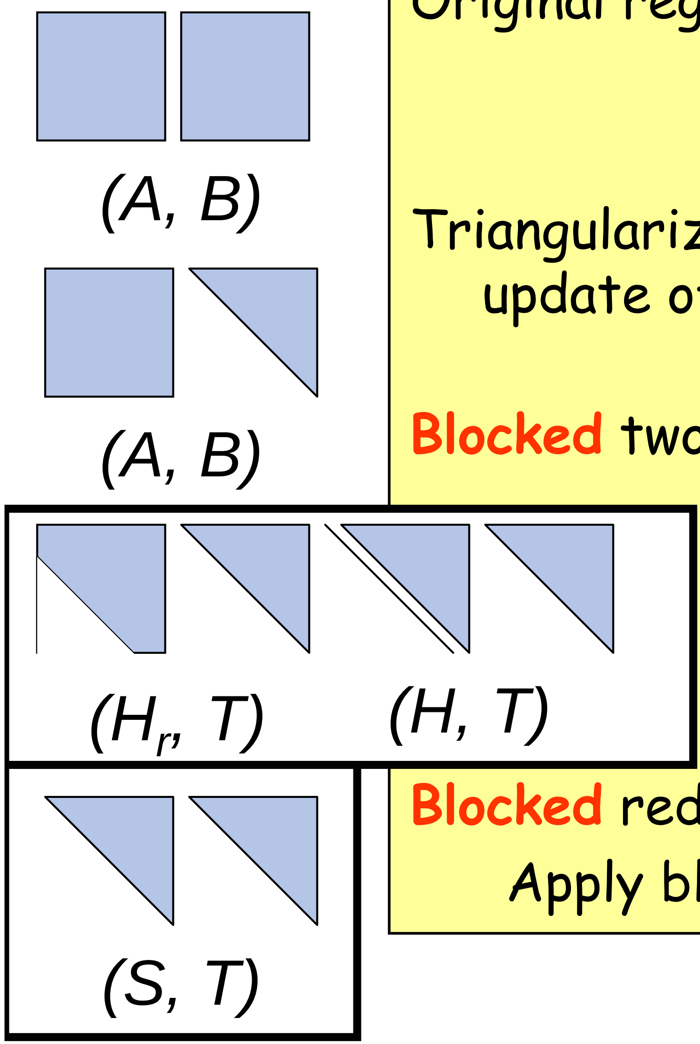
Triangularize B, blocked QR factorization of B and update of A using level 3 operations

Unblocked reduction to Hessenberg-triangular form (H, T) by application of Givens rotations to A and B from the left and right hand sides

Unblocked reduction to generalized Schur form (S, T): apply the QZ algorithm iteratively to (H, T) (Givens rotations and Householder transformations of length 3)

Blocked and Parallel Solution

(Dackland-Kågström'99, Adlerborn-D-K'00 and '02)



Original regular $N \times N$ matrix pair (A, B)

Triangularize B, blocked QR factorization of B and update of A using matrix-matrix-based operations

Blocked two-stage reduction to Hessenberg-triangular form (H, T)

1: level 3 based reduction to (H_r, T) form

2: blocked reduction of (H_r, T) to (H, T)

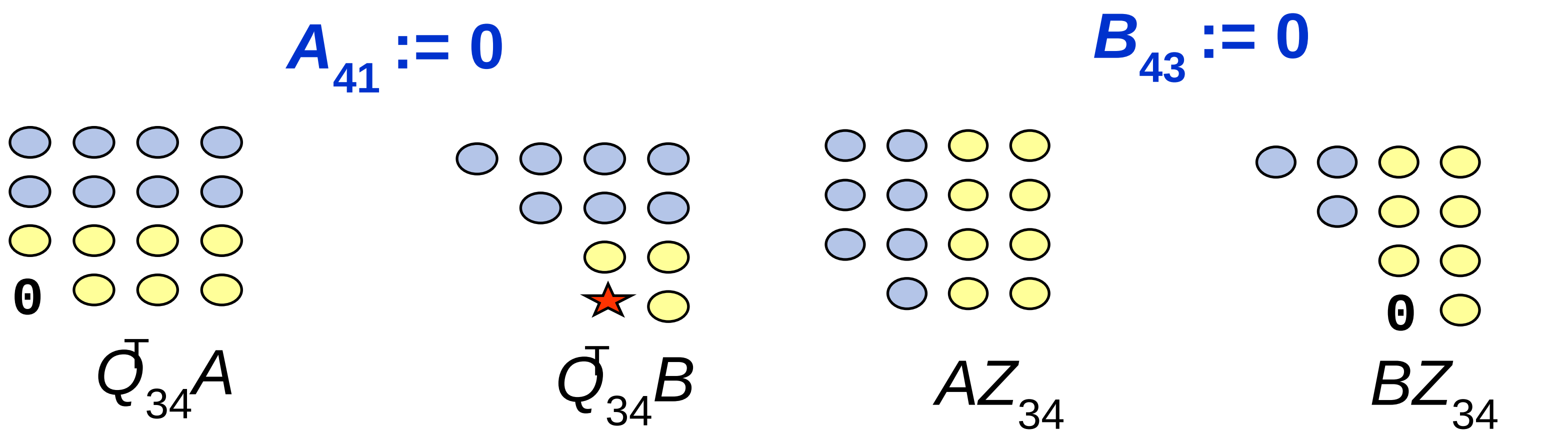
H_r *has r subdiagonals*

Blocked reduction to generalized Schur form (S, T)

Apply blocked QZ sweep iteratively to (H, T)

Basic (H, T) Form Reduction

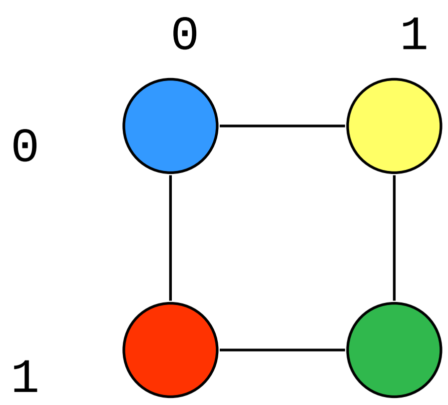
- Two-sided orthogonal transformations to reduce A to **Hessenberg form** while preserving B **triangular**
  - Givens or (2x2) Householder transformations

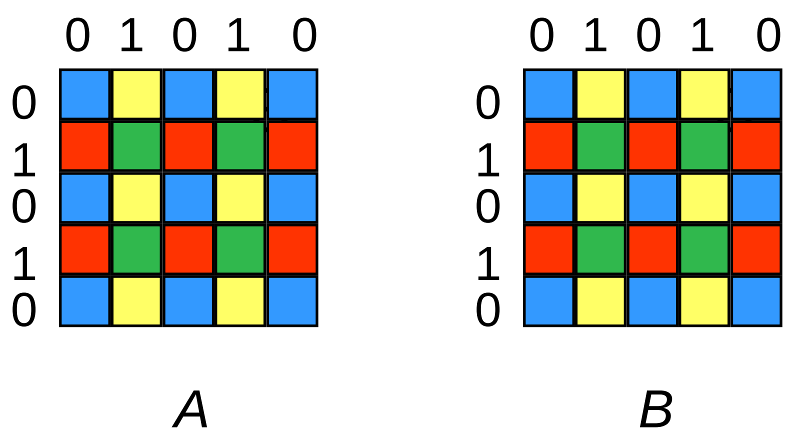


Topology and Data Layout

ScaLAPACK setting: BLAS, LAPACK, BLACS, PBLAS, MPI etc

- Data distribution of *A* and *B* on a *2x2* processor grid

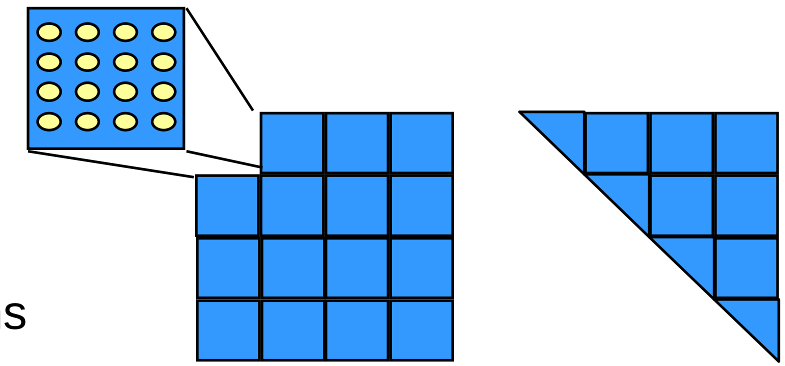




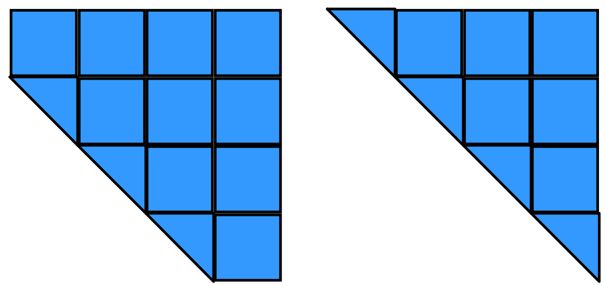
P_r x P_c processor grid   Square block scattered (cyclic, *rxr*)

Blocked Stage 1 Reduction:
From (A, B) to (H_r, T)

- Partition the matrices A and B into $r \times r$ blocks



- Orthogonal matrix U triangularizes B
  - $U^T B = R$, upper triangular,
  - Apply U to A: $A \leftarrow U^T A$
  - Use blocked level 3 QR factorization
  - Updates performed using Level 3 operations

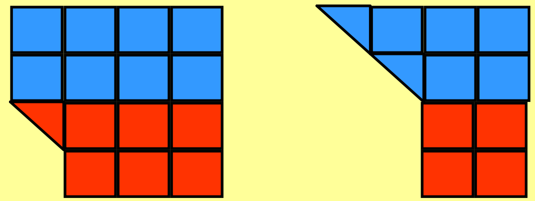


- Orthogonal equivalence transf. reduces A to block upper Hessenberg (H_r) form while preserving B triangular
  - Updates performed using Level 3 operations
  - Annihilations performed using Level 1-2 ops.

Parallel Stage 1: Level 3 Alg. (r, p)

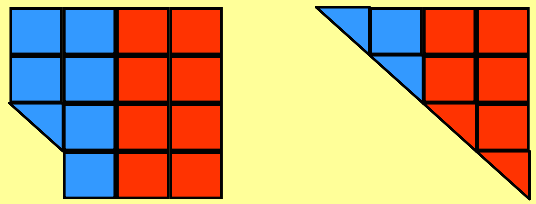
Annihilate $A_{4,1}$ and triangularize $A_{3,1}$

- QR fact $A_{3:4,1}$ <-- $U^T A_{3:4,1}$
- Update $A_{3:4,2:4}$ <-- $U^T A_{3:4,2:4}$
- Update $B_{3:4,3:4}$ <-- $U^T B_{3:4,3:4}$



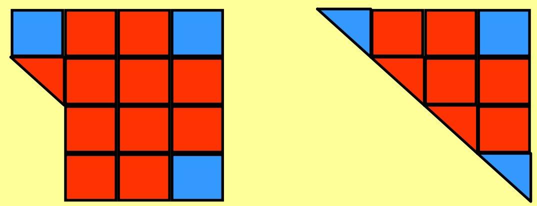
Re-triangularization of $B_{3:4,3:4}$

- RQ fact $B_{3:4,3:4}$ <-- $B_{3:4,3:4} V$
- Update $B_{1:2,3:4}$ <-- $B_{1:2,3:4} V$
- Update $A_{:,3:4}$ <-- $A_{:3:,4} V$



Annihilate $A_{3,1}$ and triangularize $A_{2,1}$

Re-triangularization of $B_{2:3,2:3}$

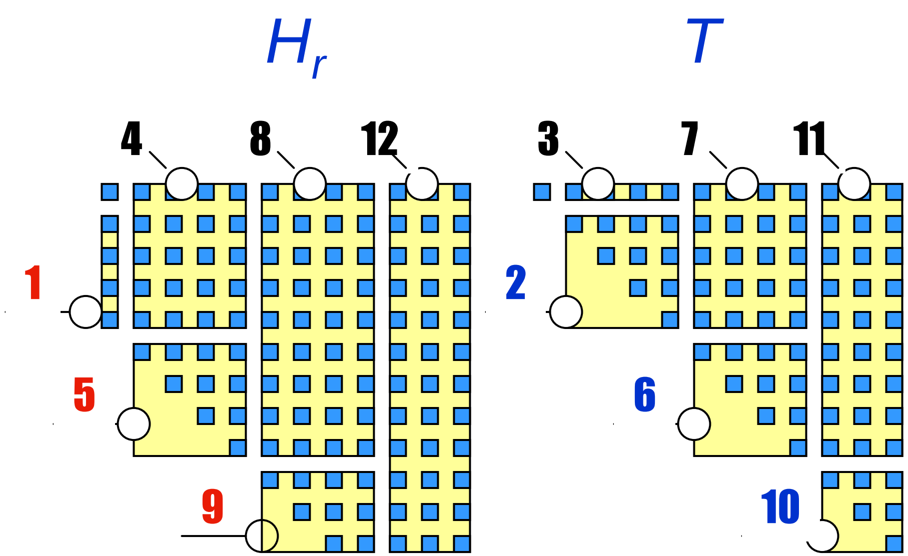


- First block column of A reduced to block Hessenberg form
- *The remaining columns of A reduced in a similar way*
- ***r = block size p = #blocks reduced per iteration (p=2 min fill-in)***

Blocked Stage 2 Reduction:
From (H_r, T) to (H, T)

Blocked Reduction using Givens rotations

Illustration of blocked sweep reducing 3 elements ($r = 4$) in column 1



"Spiral Blocked Reference Pattern"

Reduce

1. Annihilate 3 elements: **row1**
2. Update w.r.t (1); zero fill-in: **col1**
3. Update w.r.t (2)

Chase

4. Update w.r.t (1, 2)
5. Update w.r.t (2); zero fill-in: **row2**
6. Update w.r.t (5); zero fill-in: **col2**
7. Update w.r.t (1, 6)
8. Update w.r.t (1, 5, 6)
9. Update w.r.t (6); zero fill-in: **row3**
10. Update w.r.t (9); zero fill-in: **col3**
11. Update w.r.t (1, 5, 10)
12. Update w.r.t (1, 5, 9, 10)

Blocked Stage 2 Reduction:
From (H_r, T) to (H, T)

Generalization to ***Super-Sweep*** - reducing m columns

Reduction:

Reduce m columns of H_r ; updates of H_r and T are restricted to r consecutive columns at a time

Store **row** and **column** rotations to enable delayed updates

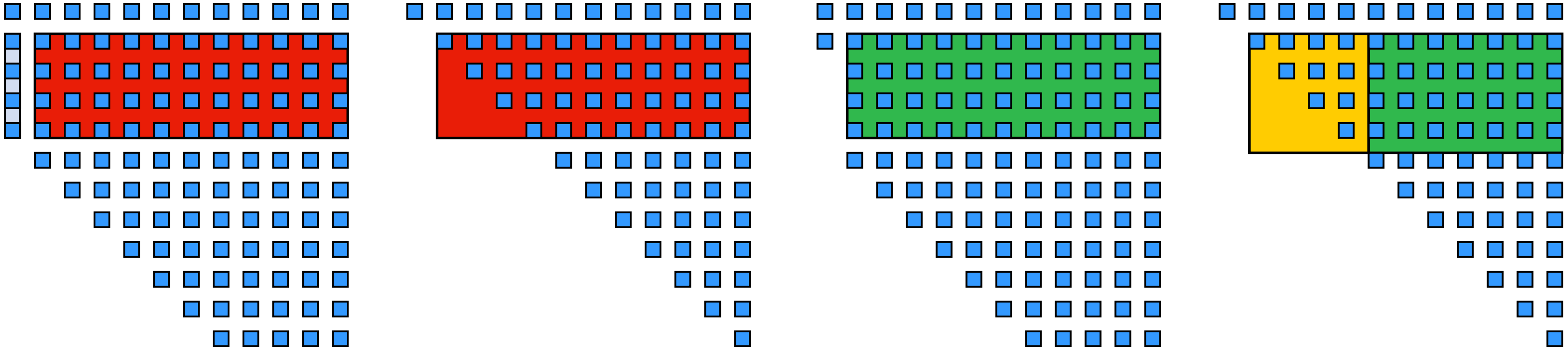
Chasing:

A super-sweep advances the m sweeps of one block column (r cols) ahead in a pipelined fashion, starting with the leading block

⇒ Reduced # loads; improved data reuse

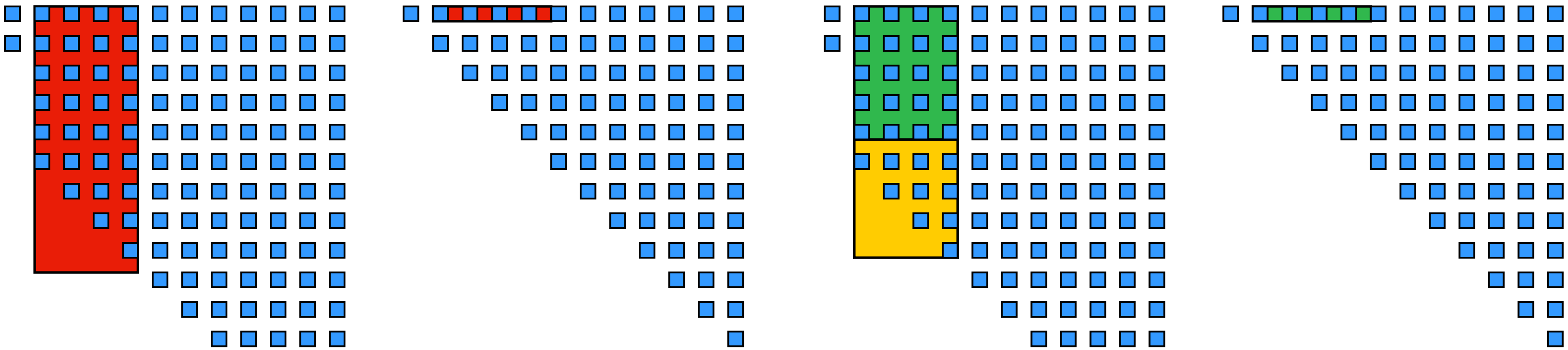
Parallel Stage 2

partial supersweep ($m = 1$, $r = 4$)



1: Zero el's (in *A*) and broadcast

2: Update and zero el's (in *B*)

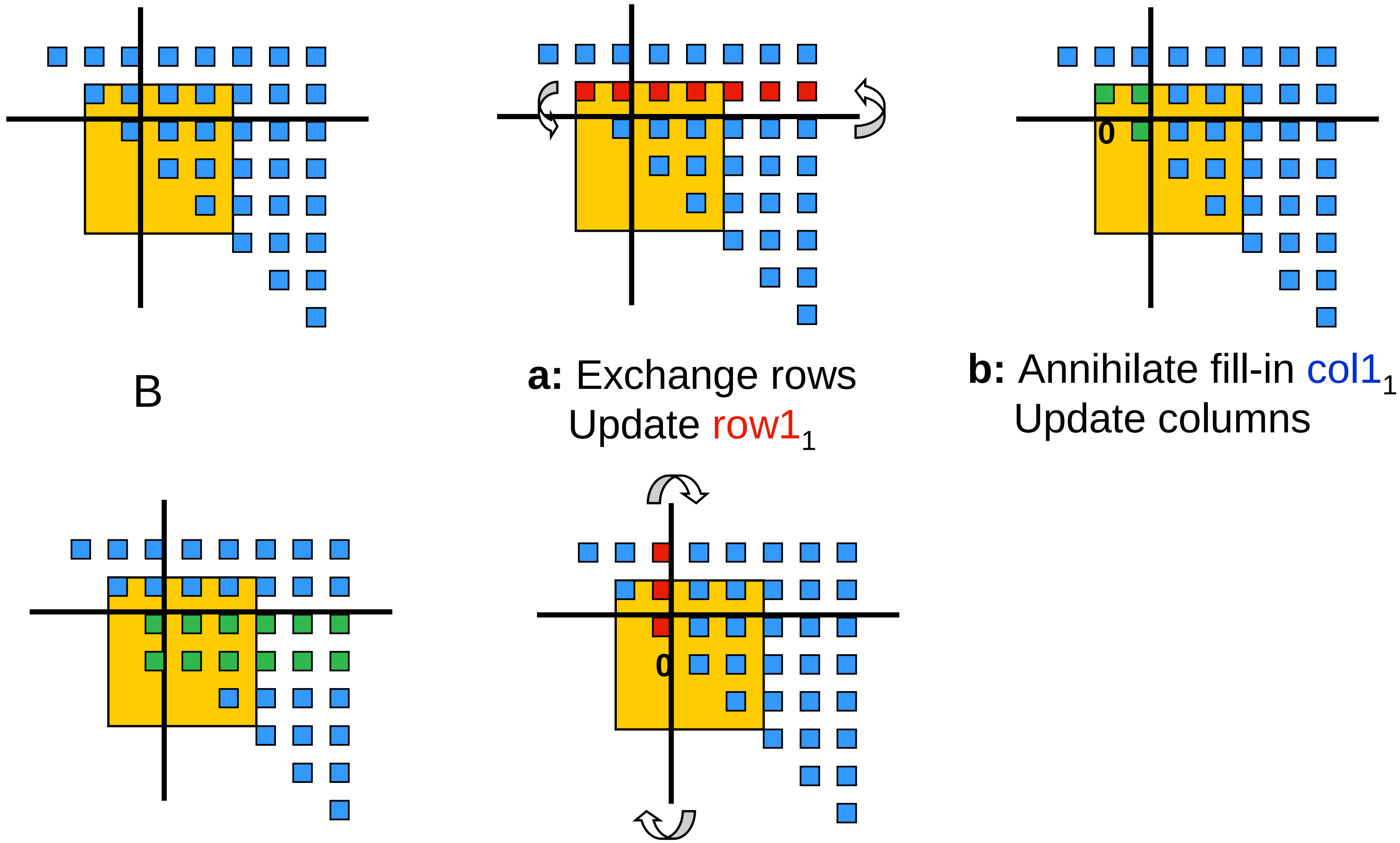


3: Broadcast

4: Update and zero el's (in *A*)

Apply r -1 row rotations to B (in step2)

Zero fill-in across processor borders



a: Exchange rows
Update row1_1

b: Annihilate fill-in col1_1
Update columns

c: Next row rotation
applied row1_2

d: Exchange columns
Annihilate fill-in col1_2

Parallel Performance-Scaled Speedup Configuration

- Fixed local matrix size (1024 x 1024 entries)
- Vary block size $r = NB$
- Vary grid configuration ($P = P_r \times P_c$)
- $p = \max(2, P_r)$ (# blocks reduced/iter. Stage 1)
- Fixed $m = 2$ (supersweep - Stage 2)

Stage 1 + Stage 2 = P**DGGHRD** - reduction to (H, T) form

Parallel Stage 1 + Stage 2 Implementation - Summary

- New ScaLAPACK-style algorithms for
  $(A, B) \rightarrow (H, T)$
- **Stage 1** speedup ≈ $k \times \mathrm{sqrt}(P)$ (= $P\_r \times P\_c$)
  - Mainly Level 3 operations
- **Stage 2** speedup ≈ $\mathrm{sqrt}(P)$
  - Level 1 - 2.5 operations ("fewer flops")
  - More costly communication
- **Stage 1 + 2** speedup > $\mathrm{sqrt}(P)$
  - Overall performance limited by Stage 2
- Performance increases with problem and grid sizes

==> Can solve large scale problems!

QZ Algorithm Overview

- A sequence of single or double shift QZ steps applied to matrix pair in (H,T) form:

$$(H,T) \longleftarrow Q^T(H,T)Z$$

- Implicit shifts are used for accelerating convergence
  - Eigenvalues of the trailing 2×2 block
- QZ step:
  - Introducing an implicit shift (single or double)
  - Bulge chasing of unwanted elements along subdiagonals of H and T
- Deflation: Problem splits into two subproblems - one with converged eigenvalues

Monitoring and decoupling assoc. with deflation surrounds a QZ step.

Deflation Strategies

Small (sub)diagonal convergence criterion:
If h(i+1,i) or t(i,i) gets tiny, it is set to 0!

- Norm-wise: $|h_{i+1,i}| \leq \text{eps} \cdot \|H\|_F$

  Used by EISPACK (QZIT), LAPACK (DHGEQZ), Dackland-Kågström

- Neighbor-wise: $|h_{i+1,i}| \leq \text{eps} \cdot (|h_{ii}| + |h_{i+1,i+1}|)$

  Used in state-of-the-art QR implementations.
  Our current Multishift QZ implementation uses nearby-diagonal deflation strategy.

Graded Matrix Pair

$$(H, T) = \left(\begin{bmatrix} 10^{0} & 10^{-3} & 0 & 0 \\ 10^{-3} & 10^{-7} & 10^{-10} & 0 \\ 0 & 10^{-10} & 10^{-14} & 10^{-17} \\ 0 & 0 & 10^{-17} & 10^{-21} \end{bmatrix}, I_4 \right)$$

Exact eigenvalues	Norm-wise deflation	Neighbor-wise deflation
1.0000009999991000	$1.000000999999910\underline{0}1$	$1.00000099999909\underline{9}$
$-.8999991111128208\times10^{-06}$	$-.89999911111128\underline{2}12\times10^{-06}$	$-.89999911111128\underline{2}13\times10^{-06}$
$.2111111558732113\times10^{-13}$	$.211111\underline{1}085047986\times10^{-13}$	$.21111115587321\underline{1}4\times10^{-13}$
$-.3736841266803067\times10^{-20}$	$0.09999999999999999\times10^{-20}$	$-.37368412668030\underline{6}8\times10^{-20}$

Results from Kressner's Diploma Thesis'01

- Nearby-diagonal deflation more accurate for graded matrices.

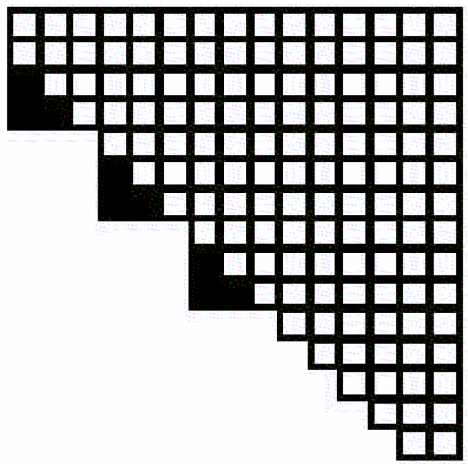
Multishift QZ

- Compute the eigenvalues of trailing $M \times M$ submatrix pair of (H, T).
- Choose MS of the M eigenvalues as shifts.
  - Choosing $M > MS$ will probably give better shifts but the computation is more expensive.
- Use the shifts in a pipelined fashion
- Avoid shift blurring by restricting the bulge size.
  - Due to rounding errors the eigenvalues of a certain bulge pencil that carries the shift information tend to be ill-conditioned for a large degree multishift (Watkins-Elsner'94, Watkins'96, '98)
- $BS = 2, 4$ up to 8 can be okay!

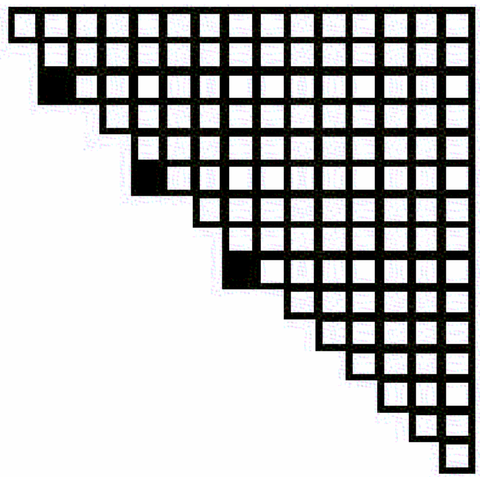
Pipelined Bulge Chasing

- Introduce $S = MS/2$ bulges of degree two and chase them down the diagonals of H and T

H



T



Bottom shifts are applied first - most likely to be closest to the eigenvalues that will deflate next.

Multishift Variants

- Pipelining of independent bulges is earlier proposed in the context of parallel QR algorithms (Van de Geijn'93, Watkins'94, Henry-Van de Geijn'97, Lang'98).

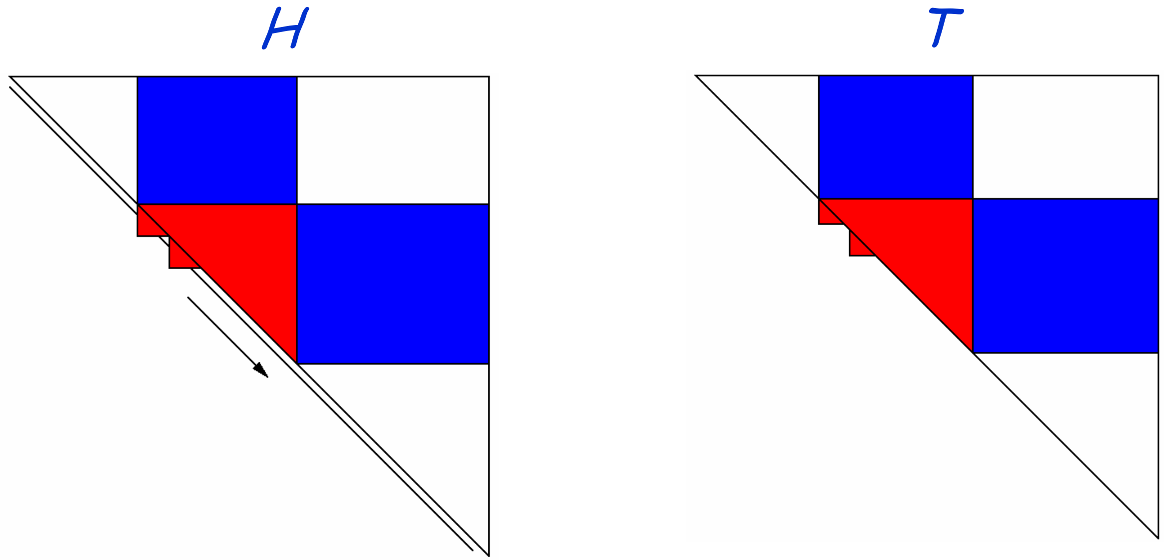
- Pipelining of tightly coupled clusters of bulges
  - Recently and successfully applied to the QR algorithm (Braman-Byers-Mathias'02a).

- Aggressive early deflation strategy for speeding up the convergence of the multishift QR algorithm (Braman-Byers-Mathias'02b).
  - Take advantage of perturbations outside the subdiagonal entries of the Hessenberg matrix in the QR iteration.

Variants generalize to matrix pencils and the QZ algorithm.

Chasing of Tightly Coupled Bulges

- $MS = 4$, $BS = 2$, $S = MS/2 = 2$, bulges of degree two are chased down within the "red window".



All transformations created are bundled and applied to "blue areas" using level 3 operations.

Deflating Perturbation Pair

- (H, T) in unreduced HT-form, Q and Z unitary
- P_H and P_T perturbation matrices such that

$$(\hat{H}, \hat{T}) \equiv Q^H (H + P_H, T + P_T) Z$$

is in reduced HT-form:

$$\hat{H} = \begin{bmatrix} \hat{H}_{11} & \hat{H}_{12} \\ 0 & \hat{H}_{22} \end{bmatrix} \qquad \hat{T} = \begin{bmatrix} \hat{T}_{11} & \hat{T}_{12} \\ 0 & \hat{T}_{22} \end{bmatrix}$$

If norm of (P_H, P_T) is tiny, problem split in two (or more) subproblems.

Ideally, we want the (2,2)-pair in (S, T) form.

Restricting the Perturbations?!

(P_H, P_T) only nonzero in last k rows and $k+1$ columns:

$$P_H = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & P_{32}^{(H)} & P_{33}^{(H)} \end{bmatrix} \qquad P_T = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & P_{33}^{(T)} \end{bmatrix}$$

Block rows and columns are of size $n-k-1$, 1, and k

The problem to find the minimal perturbation (P_H, P_T) is related to the distance to uncontrollability for the descriptor system

$$T_{33}\dot{x}(t) = H_{33}x(t) + H_{32}u(t)$$

$$\text{rank}([H_{32}, \beta H_{33} - \alpha T_{33}]) = n \text{ for all } (\alpha, \beta) \in C^2 \quad |\alpha|^2 + |\beta|^2 = 1$$

⇨ system is completely controllable

Aggressive Early Deflation in QZ (1)

Consider (H,T) in unreduced HT-form:

$$H = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ 0 & H_{32} & H_{33} \end{bmatrix} \qquad T = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ 0 & T_{22} & T_{23} \\ 0 & 0 & T_{33} \end{bmatrix}$$

Block rows and columns are of size n-k-1, 1, and k

$$\begin{bmatrix} I & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & Q^H \end{bmatrix} (H,T) \begin{bmatrix} I & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & Z \end{bmatrix} = (\hat{H}, \hat{T})$$

Generalized Schur decompostion of $k \times k$ (H_{33}, T_{33})

$$(\hat{S}_{33}, \hat{T}_{33}) = Q^H (H_{33}, T_{33}) Z$$

Aggressive Early Deflation in QZ (2)

Equivalence transformation of (H, T):

$$\begin{bmatrix} I & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & Q^H \end{bmatrix} (H, T) \begin{bmatrix} I & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & Z \end{bmatrix} = (\hat{H}, \hat{T})$$

where

$$\hat{H} = \begin{bmatrix} H_{11} & H_{12} & H_{13}Z \\ H_{21} & H_{22} & H_{23}Z \\ 0 & s & \hat{S}_{33} \end{bmatrix} \qquad \hat{T} = \begin{bmatrix} T_{11} & T_{12} & T_{13}Z \\ 0 & T_{22} & T_{23}Z \\ 0 & 0 & \hat{T}_{33} \end{bmatrix}$$

$$s = Q^H H_{32} \; (k \times 1)$$ "the k-spike"

Early Deflation?

- If m of the trailing components of the vector

$$s = Q^H H_{32} \ (k \times 1)$$

are tiny, we can make a deflation with respect to the trailing $m \times m$ matrix pair:

$$\tilde{H} = \begin{bmatrix} H_{11} & H_{12} & \hat{H}_{13} & \hat{H}_{14} \\ H_{21} & H_{22} & \hat{H}_{23} & \hat{H}_{24} \\ 0 & \hat{s} & \hat{H}_{33} & \hat{H}_{34} \\ 0 & 0 & 0 & \hat{H}_{44} \end{bmatrix} \qquad \tilde{T} = \begin{bmatrix} T_{11} & T_{12} & \hat{T}_{13} & \hat{T}_{14} \\ 0 & T_{22} & \hat{T}_{23} & \hat{T}_{24} \\ 0 & 0 & \hat{T}_{33} & \hat{T}_{34} \\ 0 & 0 & 0 & \hat{T}_{44} \end{bmatrix}$$

Block rows and columns are of size *n-k-1, 1, k-m,* and *m*, respectively.

Example of Early Deflation

$$H_6 = \begin{bmatrix} 6 & 5 & 4 & 3 & 2 & 1 \\ 0.001 & 1 & 0 & 0 & 0 & 0 \\ & 0.001 & 2 & 0 & 0 & 0 \\ & & 0.001 & 3 & 0 & 0 \\ & & & 0.001 & 4 & 0 \\ & & & & 0.001 & 5 \end{bmatrix} \quad T_6 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ & 1 & 0 & 0 & 0 & 0 \\ & & 1 & 0 & 0 & 0 \\ & & & 1 & 0 & 0 \\ & & & & 1 & 0 \\ & & & & & 1 \end{bmatrix}$$

Adlerborn-Dackland-Kågström'02

- Estimates of the distances ($||.||_F$) between (H_6, T_6) and a matrix pair with eigenvalues equal to (h_{ii}, t_{ii}) of (H_6, T_6) for $k = 5$.
- A matrix pair with …
  - 5 as eigenvalue is within distance 10^{-17}
  - 5 and 4 as eigenvalues is within distance 10^{-13}
  - 5, 4 and 3 as eigenvalues is within distance 10^{-10}
- Set the trailing m (= 1,2,3) components of s to zero.

Retransform to HT-form

$$\tilde{H} = \begin{bmatrix} H_{11} & H_{12} & \hat{H}_{13} & \hat{H}_{14} \\ H_{21} & H_{22} & \hat{H}_{23} & \hat{H}_{24} \\ 0 & \hat{s} & \hat{H}_{33} & \hat{H}_{34} \\ 0 & 0 & 0 & \hat{H}_{44} \end{bmatrix} \qquad \tilde{T} = \begin{bmatrix} T_{11} & T_{12} & \hat{T}_{13} & \hat{T}_{14} \\ 0 & T_{22} & \hat{T}_{23} & \hat{T}_{24} \\ 0 & 0 & \hat{T}_{33} & \hat{T}_{34} \\ 0 & 0 & 0 & \hat{T}_{44} \end{bmatrix}$$

- Construct $Q = I - \beta v v^T$ such that $Q^T \hat{s} = c e_1$
- Transform $(Q^T \hat{H}_{33}, Q^T \hat{T}_{33})$ to HT-form:

$$Q^T \hat{T}_{33} = (I - \beta v v^T)\hat{T}_{33} = \hat{T}_{33} - \beta v (\hat{T}_{33}^T v)^T$$

Rank-1 perturbation of an upper triangular matrix

-Apply RQ-updating, $2(k - m - 1)$ rotations -> Z

-Transform $(Q^T \hat{H}_{33} Z, Q^T \hat{T}_{33} Z)$ to HT-form using standard algorithm

Computational Experiments

$$(H,T) \xrightarrow{\text{Stage 3}} (S,T)$$

Report on results from testing of Multishift variants (prel.) of the QZ algorithm

Parameters in Multishift QZ

- *MS* = multishift size
- *BS* = # shifts/bulge
  (used in pipelined chasing)
- *NS* = # steps the bulges are chased along the block diagonals = 3**MS*/2 - 2
- *WS* = window size for aggressive deflation = 3**MS*/2

Comparisons with
- LAPACK
- Blocked QZ
(Dackland-Kågström'99)

Computer: HPC2N chips
- 4 Power3 (375MHz)
- 4GB of memory

Ex1: Random (A,B) red. to HT-form QZ variants w/o Aggressive Deflation

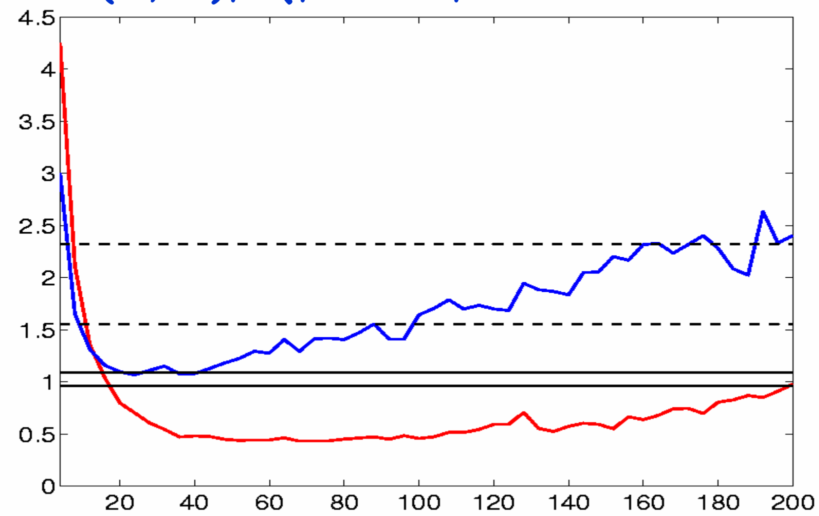
1. LAPACK "dashed line"
(upper - without, lower - with)

2. Blocked (Dackland-Kågström, TOMS'99) "solid line"

3. Multishift - tightly coupled 2-by-2 bulges

4. Multishift + Aggressive Early Deflation

Time: (S,T), Q, and Z, $N = 1000$



Multishift size
$MS = 4:4:200$

AD-windows:
1. 40
2. 30 ($NB = 80$)
4. $3*MS/2$

3. Best $MS = 20, 32, 36$ 4. Best $MS = 68, 72, 76$

Ex1: Random (A, B) red. to HT-form QZ variants w/o Aggressive Deflation

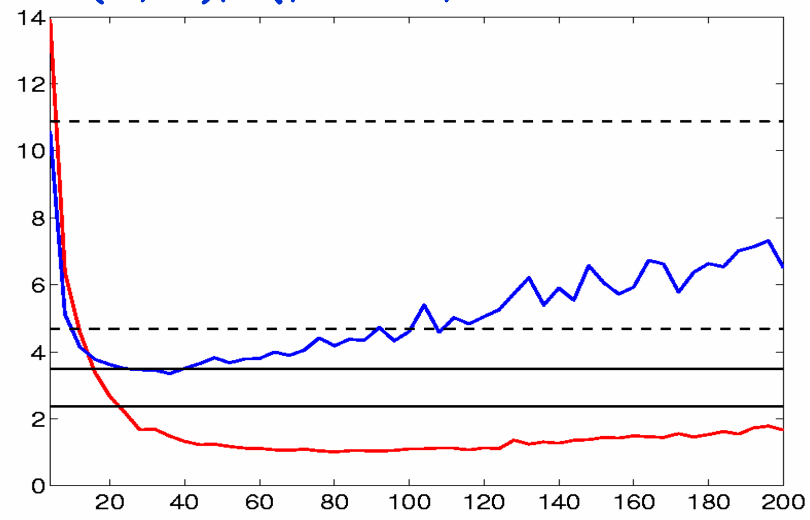
1. LAPACK "dashed line"
(upper - without, lower - with)

2. Blocked (Dackland-Kågström, TOMS'99) "solid line"

3. Multishift - tightly coupled 2-by-2 bulges

4. Multishift + Aggressive Early Deflation

Time: (S, T), Q, and Z, $N = 1500$



Multishift size
$MS = 4{:}4{:}200$

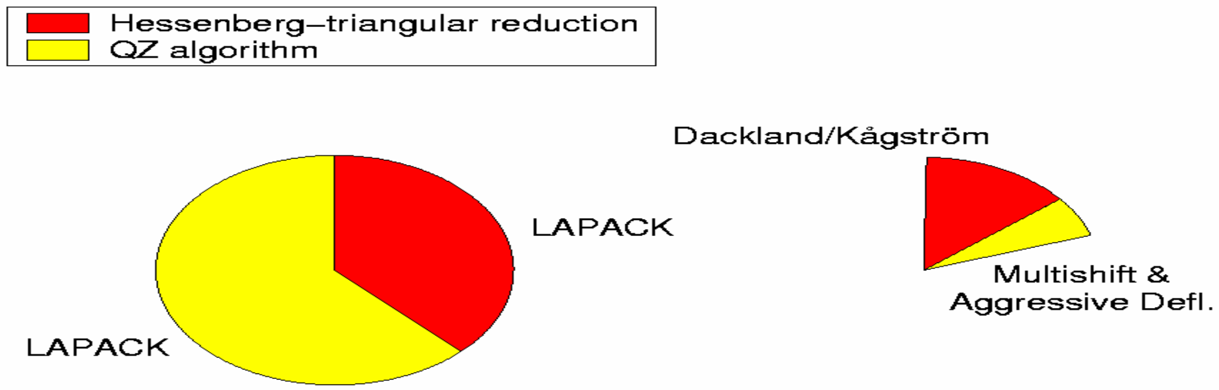
AD-windows:
1. 50
2. 30 ($NB = 72$)
4. $3*MS/2$

3. Best $MS = 28, 32, 36$

4. Best $MS = 80, 88, 92$

Cost for HT- and ST-reductions

Ex 1: Random (A,B) $N = 1500$



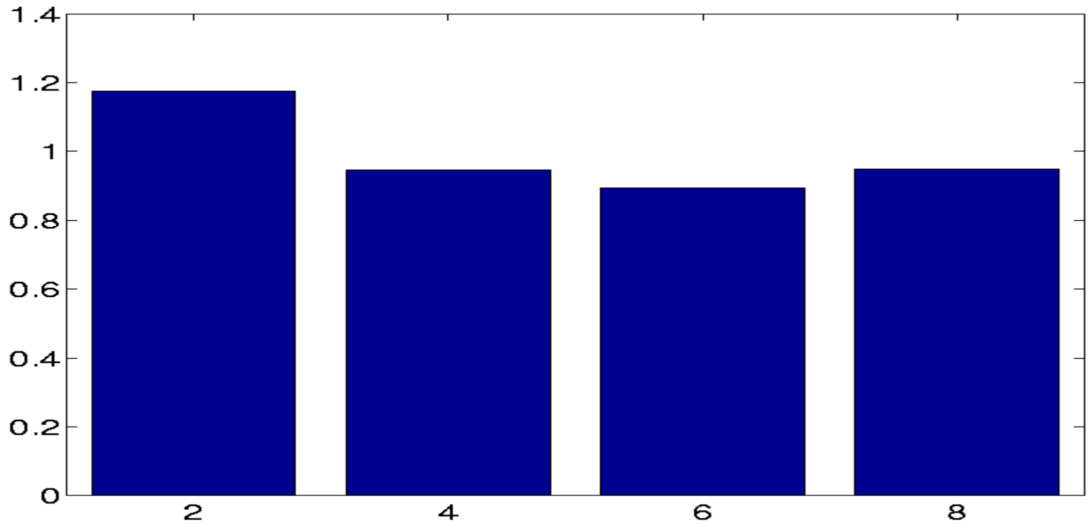
Stage 1 & Stage 2: Dackland-Kågström blocked

Stage 3: Multishift + Aggressive Early Deflation

Varying the Bulge Size in the Pipeline Chasing

Ex 1: Random (A,B) $N = 1500$ $WS = 90$

Time: (S,T), Q, and Z



Bulge Size BS

Risk for shift blurring!
(Watkins-Elsner'94, Watkins'96)

Ex2: (H_6, T_6) for varying N

$$H_6 = \begin{bmatrix} 6 & 5 & 4 & 3 & 2 & 1 \\ 0.001 & 1 & 0 & 0 & 0 & 0 \\ & 0.001 & 2 & 0 & 0 & 0 \\ & & 0.001 & 3 & 0 & 0 \\ & & & 0.001 & 4 & 0 \\ & & & & 0.001 & 5 \end{bmatrix} \qquad T_6 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ & 1 & 0 & 0 & 0 & 0 \\ & & 1 & 0 & 0 & 0 \\ & & & 1 & 0 & 0 \\ & & & & 1 & 0 \\ & & & & & 1 \end{bmatrix}$$

1. LAPACK "dashed line" (without aggressive deflation)

2. Blocked (Dackland-Kågström, TOMS'99) "solid line"

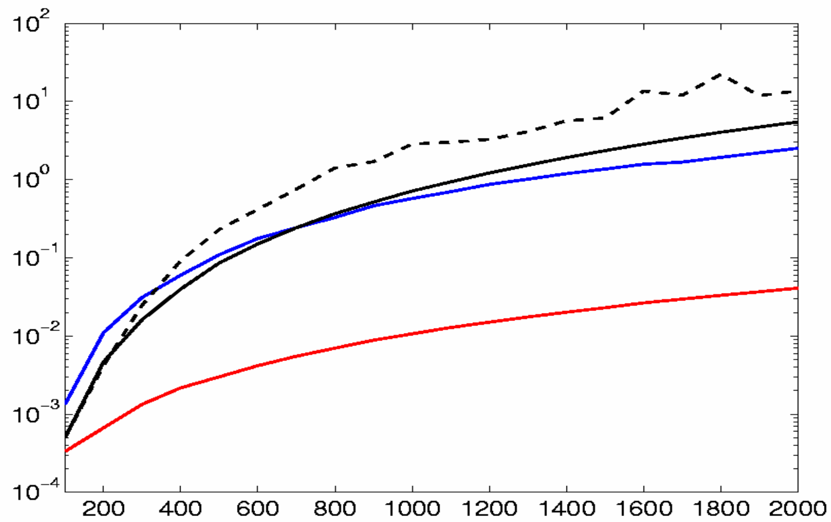
3. Multishift - tightly coupled 2-by-2 bulges

4. Multishift + Aggressive Early Deflation

AD-windows:

1. -
2. - (NB = 48)
4. 3*MS/2 (MS = 40)

Time: (S, T), Q, and Z



Problem size

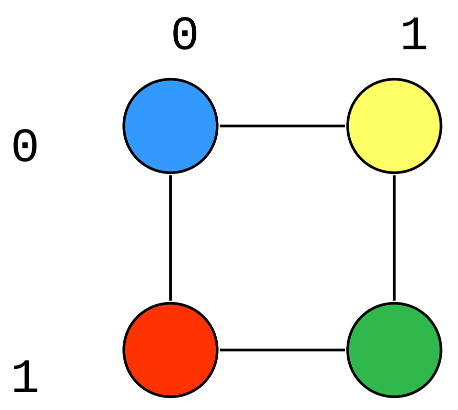
N = 100:100:2000

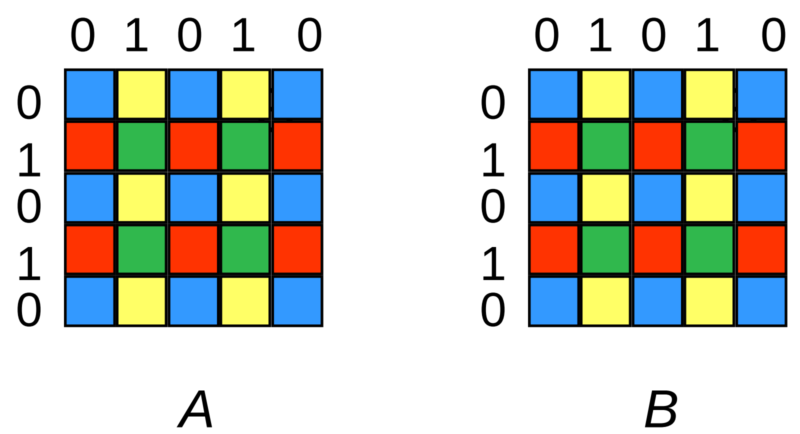
4. Two orders of magnitude better!

Topology and Data Layout

ScaLAPACK setting: BLAS, LAPACK, BLACS, PBLAS, MPI etc

- Data distribution of *A* and *B* on a *2x2* processor grid



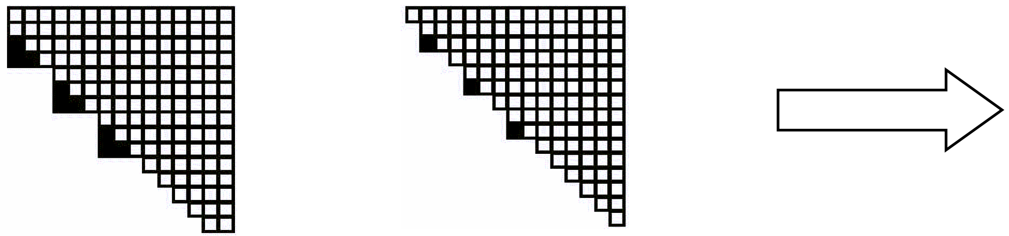


P_r x P_c processor grid   Square block scattered (cyclic, *r x r*)

P00: **(A_{11}, B_{11})**, (A_{13}, B_{13}), (A_{15}, B_{15}), (A_{31}, B_{31}), **(A_{33}, B_{33})**, (A_{35}, B_{35}), (A_{51}, B_{51}), (A_{53}, B_{53}), **(A_{55}, B_{55})**

Stage 3 in Parallel Setting

- Choose #shifts *MS* large enough, and introduce the bulges with enough space between them



workload can be evenly distributed across the 2D processor grid.

Currently: One bulge per diagonal block
To come:

- One tightly cluster of bulges per diagonal block
- *Aggressive early deflation*

Parallel Multishift QZ prototype works!

- *Accurate results*
- Shows speedup with increasing #shifts

Summary and Ongoing Work

$$(A,B) \xrightarrow{\text{Stage 1}} (H_r,T) \xrightarrow{\text{Stage 2}} (H,T) \xrightarrow{\text{Stage 3}} (S,T)$$

- Efficient blocked and parallel algorithms and software for Stage 1 and Stage 2 completed. Blocked Stage 3 as well.
- Stage 3: New Multishift QZ with aggressive early deflation. Variants are implemented and tested - good results!
- Parallel Multishift QZ prototype implemented!
- Combining the parallel and blocked Stage 1-3 reductions enable effective computation of the generalized Schur form of a regular matrix pair (A,B).
- This work is essential since the reduction to generalized Schur form is a fundamental operation in large-scale control applications.
- Our software is designed for integration in state-of-the-art libraries such as LAPACK, ScaLAPACK and SLICOT.

Some of Our References

K. Dackland and B. Kågström (1999) Blocked Algorithms and Software for Reduction of a Regular Matrix Pair to Generalized Schur Form. ACM Trans. Math. Software, Vol. 25, No. 4, 425-454, 1999.

B. Adlerborn, K. Dackland, and B. Kågström (2001) Parallel Two-Stage Reduction of a Regular Matrix Pair to Hessenberg-Triangular Form, PARA2000, Springer-Verlag, Lecture Notes in Computer Science (LNCS), Vol. 1947, pp 92-102.

B. Adlerborn, K. Dackland, and B. Kågström (2002) Parallel and Blocked Algorithms for Reduction of a Regular Matrix Pair to Hessenberg-Triangular and Generalized Schur Forms, PARA2002, Springer-Verlag, LNCS, Vol. 2367, pp 319-328.

B. Kågström and D. Kressner (2003) Multishift Variants of the QZ Algorithm with Aggressive Early Deflation (in prep.)

B. Adlerborn, B. Kågström and D. Kressner (2003) Parallelizing multishift QZ algorithms for the generalized eigenvalue problem (in prep.)