Abstract

We introduce Chase And Trim (CAT), an additional deflation mechanism for symmetric tridiagonal QR iteration that creates earlier splits of active blocks. For selected first-subdiagonal couplings, CAT tests whether one local Jacobi similarity rotation
can annihilate the coupling while producing only negligible cross-cut fill. The rotation creates at most two disjoint symmetric fill-in pairs. CAT accepts the split only when both fill magnitudes are no larger than the local tolerance
used for conventional deflation, so the discarded perturbation satisfies the same local spectral-norm bound. A cheap sufficient inequality tests candidates before constructing the rotation. Accepted splits allow subsequent QR iteration to continue independently on smaller child blocks. Candidate cuts are organized into nonintersecting sets, exposing parallelism in testing and accepted rotations.

We tested the eigenvector-accumulating algorithm on eight matrix families at n = 63, 255, 511, 1023, with ε = 1e-14.

Relative reductions in QR sweep positions ranged from 0.8% to 22.5%, and reductions in weighted arithmetic work from 0.2% to 22.4%. The largest reductions occurred for Weak-link symmetric tridiagonal cases; the Dense graded spectrum family also showed a large average reduction of 10.3%,
consistent with the local CAT acceptance criterion. The largest combined cost of CAT testing and accepted rotations was 1.2% of candidate weighted work.

Eigenvalue accuracy relative to LAPACK, eigenvector orthogonality, and eigenpair residuals remained of the same order as for standard QR implementation.

Introduction

Overview

Implicit QR remains an implemented LAPACK method for the symmetric tridiagonal eigenproblem

(LAPACK [2]), alongside divide-and-conquer and other eigensolver routes.

The standard bulge-chasing QR algorithm for a symmetric matrix A proceeds as follows:

① Reduce A to tridiagonal T by Householder similarity transformations.

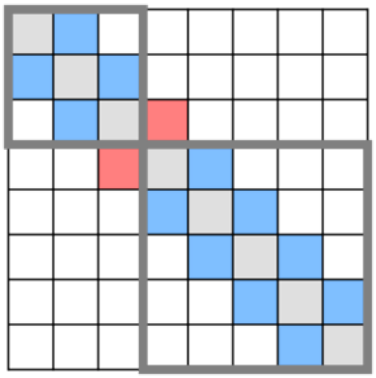
② Choose a shift and apply a sweep of Givens similarity transformations within the active block,
chasing the temporary second-band bulge from top to bottom

and restoring tridiagonal form (Watkins [1]).

③ Scan the off-diagonal band for conventional deflation opportunities between sweeps.
Whenever an off-diagonal entry becomes sufficiently small, typically near the end of the active block,
set it to zero and split the block at that position.

④ Repeat the process until every active block has size at most 2×2 and diagonalize directly.

Splitting is therefore an essential mechanism of convergence.



Additional deflation opportunities

Conventional deflation requires the coupling itself to be numerically negligible.
We investigate whether some entries that fail this criterion can be removed by one Jacobi rotation.
The rotation annihilates the selected first-band entry and creates two new second-band entries,
shown in purple on the image in the next section.
If both new entries are below the accepted tolerance, they may be discarded and allow an early split.

We call this augmented bulge-chasing strategy Chase And Trim (CAT).

A related rotate-then-deflate step occurs in symmetric divide and conquer (Tisseur & Dongarra [12]).

When two poles of the diagonal-plus-rank-one problem are close,
a Givens rotation can zero one component of the update vector,
and the small off-diagonal entry created in the diagonal part may then be discarded.
CAT applies the same broad principle inside tridiagonal QR,
but annihilates the tridiagonal coupling and creates two cross-cut fill pairs,
both of which must satisfy the local deflation tolerance.

To our knowledge, this deflation mechanism has not previously been applied within symmetric tridiagonal QR.

To reduce cost of testing for the additional splits, we use the following strategies:

- Accept every spontaneous conventional split, then test for only non-intersecting active ones
- Retain the local conventional-deflation tolerances τ for reuse in the active-split tests
- Compare the factors first, and form the products only if the comparison does not determine acceptance.

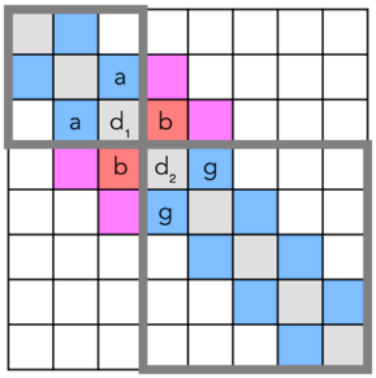
Parallel implementation affords additional benefits:

- The cost of testing can be reduced by testing all non-intersecting pivots simultaneously
- The cost of rotations can be reduced by rotating all non-intersecting pivots simultaneously
- Earlier splits create independent blocks sooner, exposing parallelism earlier

---

Jacobi rotation and fill-in

The rotation that annihilates the pivot b also creates the magenta second-band fill-ins.



The complete local similarity transformation is shown below.

$$T = \begin{bmatrix} \cdot & \bullet & 0 & 0 & 0 & 0 & 0 & 0 \\ \bullet & \cdot & a & 0 & 0 & 0 & 0 & 0 \\ 0 & a & d_1 & b & 0 & 0 & 0 & 0 \\ 0 & 0 & b & d_2 & g & 0 & 0 & 0 \\ 0 & 0 & 0 & g & \cdot & \bullet & 0 & 0 \\ 0 & 0 & 0 & 0 & \bullet & \cdot & \bullet & 0 \\ 0 & 0 & 0 & 0 & 0 & \bullet & \cdot & \bullet \\ 0 & 0 & 0 & 0 & 0 & 0 & \bullet & \cdot \end{bmatrix} \qquad G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & c & -s & 0 & 0 & 0 & 0 \\ 0 & 0 & s & c & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T' = G^T T G = \begin{bmatrix} \cdot & \bullet & 0 & 0 & 0 & 0 & 0 & 0 \\ \bullet & \cdot & ca & -sa & 0 & 0 & 0 & 0 \\ 0 & ca & d_1' & 0 & sg & 0 & 0 & 0 \\ 0 & -sa & 0 & d_2' & cg & 0 & 0 & 0 \\ 0 & 0 & sg & cg & \cdot & \bullet & 0 & 0 \\ 0 & 0 & 0 & 0 & \bullet & \cdot & \bullet & 0 \\ 0 & 0 & 0 & 0 & 0 & \bullet & \cdot & \bullet \\ 0 & 0 & 0 & 0 & 0 & 0 & \bullet & \cdot \end{bmatrix}$$

• the pivot positions become zero
• the neighboring first-band entries become (c a) and (c g)
• the second-band fill-in entries have magnitudes | s a | and | s g |

To construct the rotation, define $\Delta = d_1 - d_2$ and $t = \frac{s}{c}$

The condition for the rotated pivot to vanish is
$b\,(1 - t^2) = \Delta\, t.$

Choosing the shorter rotation and Jacobi root with $|t| \le 1$ gives

$$t = \frac{2b}{\Delta + \operatorname{sign}(\Delta)\sqrt{\Delta^2 + 4b^2}}$$

with $|t| = 1$ when $\Delta = 0$ and $t = 0$ when $b = 0$.
The rotation coefficients are

$$c = \frac{1}{\sqrt{1 + t^2}}, \quad s = t\,c.$$

We define that the rotation produces a valid split when both fill-in magnitudes are under the accepted tolerance τ:
$|s\,a| \le \tau$ and $|s\,g| \le \tau$.

We use the same local threshold as for conventional deflation:
$\tau = \varepsilon\,(|d_1| + |d_2|)$,
where ε is the relative tolerance chosen for the calculation.

Defining $M = \max(|a|, |g|)$, these two requirements combine into the single condition
$|s|\,M \le \tau$.

For completeness, we include a boundary example where d_1 is the first diagonal entry.

The left neighboring coupling a is absent, so the fill-in (−s a) is not produced, and only the single fill-in (s g) appears.

$$T = \begin{bmatrix} d_1 & b & 0 & 0 & 0 & 0 & 0 & 0 \\ b & d_2 & g & 0 & 0 & 0 & 0 & 0 \\ 0 & g & \bullet & \bullet & 0 & 0 & 0 & 0 \\ 0 & 0 & \bullet & \bullet & \bullet & 0 & 0 & 0 \\ 0 & 0 & 0 & \bullet & \bullet & \bullet & 0 & 0 \\ 0 & 0 & 0 & 0 & \bullet & \bullet & \bullet & 0 \\ 0 & 0 & 0 & 0 & 0 & \bullet & \bullet & \bullet \\ 0 & 0 & 0 & 0 & 0 & 0 & \bullet & \bullet \end{bmatrix} \qquad G = \begin{bmatrix} c & -s & 0 & 0 & 0 & 0 & 0 & 0 \\ s & c & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T' = G^{\mathsf{T}} T G = \begin{bmatrix} d_1' & 0 & sg & 0 & 0 & 0 & 0 & 0 \\ 0 & d_2' & cg & 0 & 0 & 0 & 0 & 0 \\ sg & cg & \bullet & \bullet & 0 & 0 & 0 & 0 \\ 0 & 0 & \bullet & \bullet & \bullet & 0 & 0 & 0 \\ 0 & 0 & 0 & \bullet & \bullet & \bullet & 0 & 0 \\ 0 & 0 & 0 & 0 & \bullet & \bullet & \bullet & 0 \\ 0 & 0 & 0 & 0 & 0 & \bullet & \bullet & \bullet \\ 0 & 0 & 0 & 0 & 0 & 0 & \bullet & \bullet \end{bmatrix}$$

Since any symmetric 2×2 block is diagonalized directly by one Jacobi rotation, we use the recursive stopping criterion of active-block size no larger than 2.

Why two fill-ins do not increase the accepted perturbation

We compare the spectral norms of the perturbation matrices discarded by conventional and active splitting (Golub & Van Loan [9]).

This measures the largest factor by which a matrix can increase the Euclidean length of a vector.

After a Jacobi rotation, the local submatrix $T' =$

$$\begin{bmatrix} \bullet & ca & -sa & 0 \\ ca & d_1' & 0 & sg \\ -sa & 0 & d_2' & cg \\ 0 & sg & cg & \bullet \end{bmatrix}$$

An active split discards $-sa$ and sg, giving $\tilde{T} =$

$$\begin{bmatrix} \bullet & ca & 0 & 0 \\ ca & d_1' & 0 & 0 \\ 0 & 0 & d_2' & cg \\ 0 & 0 & cg & \bullet \end{bmatrix}$$

and $T' = \tilde{T} + E$

where $E = \begin{bmatrix} 0 & 0 & -sa & 0 \\ 0 & 0 & 0 & sg \\ -sa & 0 & 0 & 0 \\ 0 & sg & 0 & 0 \end{bmatrix}$ is the discarded perturbation

For any unit vector $\hat{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}$ $E\,\hat{v} = \begin{bmatrix} -sa\ v_3 \\ sg\ v_4 \\ -sa\ v_1 \\ sg\ v_2 \end{bmatrix}$

Therefore

$$\| E\,\hat{v} \|_2^2 = (s\,a)^2 (v_1^2 + v_3^2) + (s\,g)^2 (v_2^2 + v_4^2).$$

Denote $\eta = \max(|s\,a|, |s\,g|)$

$$\downarrow$$

$$\| E\,\hat{v} \|_2^2 \le \eta^2 (v_1^2 + v_2^2 + v_3^2 + v_4^2) = \eta^2 \| \hat{v} \|_2^2$$

$$\downarrow$$

$$\| E \|_2 \le \eta.$$

Even though the active split discards two symmetric fill-in couplings, they occupy disjoint coordinate pairs, and the spectral norm is determined by the larger entry.

The active-split acceptance conditions
$|s\,a| \le \tau$ and $|s\,g| \le \tau$
therefore imply $\| E \|_2 \le \tau$.

Conventional deflation discards only b: (E passive) = $\begin{bmatrix} 0 & b \\ b & 0 \end{bmatrix}$

For $\hat{\mathbf{v}} = [\, v_1, v_2 \,]^T$, (E passive) $\hat{\mathbf{v}} = [\, b\, v_2, b\, v_1 \,]^T$,

↓

$\| \text{(E passive) } \hat{\mathbf{v}} \|_2^2 = b^2 (v_1^2 + v_2^2) = b^2 \| \hat{\mathbf{v}} \|_2^2$,

↓

$\| \text{E passive} \|_2 = | b |$.

Conventional deflation requires $| b | \leq \tau$,
so both rules impose the same spectral-norm bound, $\| E \|_2 \leq \tau$:

- conventional split: $\| \text{E passive} \|_2 = | b | \leq \tau$
- active split: $\| E \|_2 = \max(| s\, a |, | s\, g |) \leq \tau$

Two related observations are worth noting:

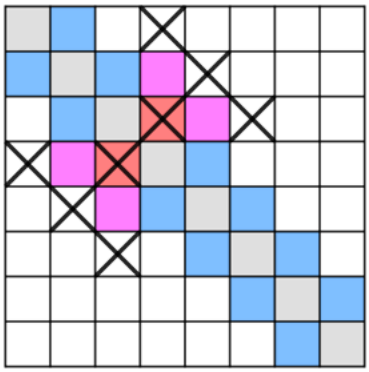
① Suppose we require a more permissive condition than b = 0:
$| b | \leq \tau$, and $\max(| s\, a |, | s\, g |) \leq \tau$

The discarded perturbation will contain b, as well as the two fill-ins,
so its spectral norm can exceed τ.
We tested this relaxed rule, and it produced very few additional split opportunities.

② A second chance for the discarded fill-ins

One may ask whether the discarded magenta fill-ins should themselves be rotated.
The crossed entries in the image below are the ones affected by the second set of

rotations.



The additional cleanup was tested.
It benefited graded-spectrum cases, but otherwise did not materially change the results,
while increasing the accepted perturbation.
We therefore simply mention it here as a possible extension.

From the exact condition to an acceptance test

The exact numerical-split condition was derived to be
$| s | M \leq \tau$,
where $M = \max(| a |, | g |)$

τ is normally computed to evaluate conventional deflation opportunities as
$\tau = \varepsilon (| d_1 | + | d_2 |)$
and retained.

Since the cost of computing t as well as s from t is substantial, we look for

acceptance inequalities to determine whether a rotation can produce a valid split.

Since $|s| \le 1$, any candidate with
$M \le \tau$
automatically satisfies $|s| M \le \tau$ and is accepted.

For the remaining candidates:

The Jacobi rotation equation is
$b(1 - t^2) = \Delta t$, where
- $\Delta = d_1 - d_2$
- from choosing the shorter rotation: $|t| \le 1$, so $0 \le (1 - t^2) \le 1$

We seek an upper bound on $|s|$ for $|s| M \le \tau$
from $s = t c$ and the Jacobi rotation equation:

$$|s| = |t||c| = \frac{|b|(1 - t^2)|c|}{|\Delta|}$$

(the acceptance test is used only for $\Delta \neq 0$)

Since $0 \le (1 - t^2) \le 1$ and $0 < |c| \le 1$,

$$|s| \le \frac{|b|}{|\Delta|}$$

Substitute the upper bound of $|s|$ into the exact numerical-split condition:

$$\text{if } \left(\frac{|b|}{|\Delta|} \right) M \le \tau,$$

the exact condition $|s| M \le \tau$ is satisfied, and the rotation produces a valid split.

Equivalently,
$|b| M \le |\Delta| \tau$
is a sufficient acceptance condition for a valid split, although

some candidates failing this acceptance condition may still satisfy the exact condition.

Summary: a one-rotation split is favored by
• a small pivot $|b|$ allows a smaller rotation
• a large local diagonal separation $|\Delta| = |d_1 - d_2|$ also allows a smaller rotation
• small neighboring couplings $|a|$ and $|g|$, so the rotation creates only small fill-ins

These opportunities tend to appear as ordinary bulge chasing progresses and
• off-diagonal entries get smaller
• diagonal entries converge on distinct eigenvalues

Why were so few exact opportunities missed?

The algorithm accepts a candidate when the sufficient condition
$|b| M \le |\Delta| \tau$
is satisfied.

The exact one-rotation split condition is
$|s| M \le \tau$,
where s is the sine of the Jacobi rotation and M is the larger neighboring coupling.

$$\text{Because } |s| \le \frac{|b|}{|\Delta|},$$

the acceptance condition satisfies the exact inequality, but the converse may not hold.

When the additional diagnostic found only a few missed opportunities over the full test set,
we initially searched for an error and then examined the gap algebraically.

A false negative requires
$|s| M \le \tau$
and

$$\frac{|b| M}{|\Delta|} > \tau$$

Its possibility is therefore determined by the gap between $|s|$ and $\frac{|b|}{|\Delta|}$

For $\Delta \neq 0$, with $t = \tan\theta$ and $|t| < 1$,

$$|s| = \frac{|t|}{\sqrt{1 + t^2}},$$

and the exact Jacobi rotation equation gives

$$\frac{|b|}{|\Delta|} = \frac{|t|}{1 - t^2}$$

Usually $M \gg \tau$, so the admissible rotation and $|t|$ are small, and
both quantities approximate the numerator $|t|$,
so the acceptance condition for $|s|$ lies very close to its exact value.

Eigenvectors

The eigenvector accounting below covers only the tridiagonal QR phase, so both the tridiagonalization and the subsequent application of its eigenvectors are excluded.

The tridiagonal eigensolver starts its eigenvector accumulator at the identity.

During the algorithm, V is updated as follows:

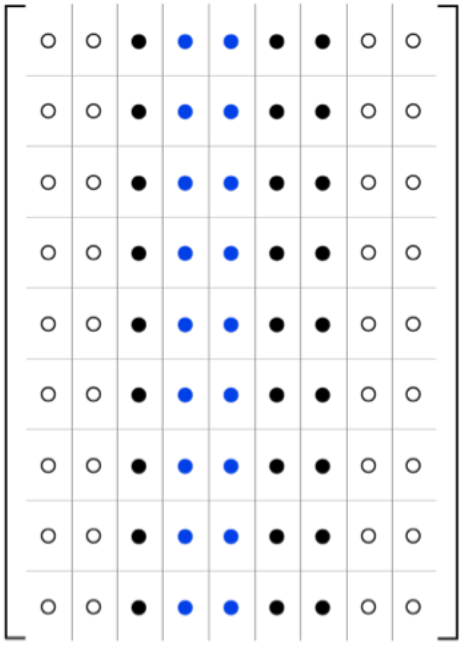
① With every Givens rotation used in bulge chasing
- The Givens parameters c and s are stored during the sweep.
- After the sweep, the rotations are applied to V by updating only the two affected columns.

② With every Jacobi rotation of an accepted active CAT split,
which is also done by updating the two corresponding columns of V.

③ With every Jacobi rotation of a terminal 2 × 2 block

④ Conventional deflation requires no additional eigenvector update

Suppose
- n = 9 is the size of the original matrix
- m = 5 is the size of the current active tridiagonal block
- p = 3 is the start index of the block

$V' = V\,G$ updates the columns k and k+1  for every row i as

- $v'_{i,k} = c\,v_{i,k} + s\,v_{i,k+1}$
- $v'_{i,k+1} = -s\,v_{i,k} + c\,v_{i,k+1}$

Each eigenvector rotation updates two columns of the full n × n matrix V.
For each of its n rows, this requires four multiplications and two additions or subtractions,
so one rotation costs 6n arithmetic operations.

An accepted active CAT split adds one Jacobi rotation and therefore 6n eigenvector operations.

A QR sweep on an unreduced block of order m uses m − 1 Givens rotations.
Its eigenvector accumulation therefore costs 6n(m − 1) arithmetic operations.

Matrix families used for testing

Matrix family	Construction for the experiment	Source if applicable
General dense symmetric No imposed structure	• Diagonal entries sampled uniformly from [−2.0, 2.0] • Off-diagonal entries sampled uniformly from [−2.0, 2.0], mirrored across the diagonal	
Dense correlation Positive-definite with unit diagonal	• Eigenvalues sampled uniformly from [0.5, 1.5] • D is the diagonal eigenvalue matrix • Q is generated as a product of randomized plane rotations • A is formed as $Q\ D\ Q^T$ • Rows and columns rescaled so that all diagonal entries equal 1	(Davies & Higham [5])
Dense clustered spectrum Eigenvalues concentrated in 4 narrow clusters	• Eigenvalues generated in 4 clusters centered at -1.5, -0.5, 0.5, 1.5 with independent jitter in [−0.001, 0.001] • D is the diagonal eigenvalue matrix • Q is generated as a product of randomized plane rotations • A is formed as $Q\ D\ Q^T$	(LAPACK [4])
Dense graded spectrum Positive-definite with a geometrically graded spectrum	• Eigenvalue exponents evenly spaced from -3.0 to 3.0 ↓ Eigenvalues range from 0.001 to 1000.0 • D is the diagonal eigenvalue matrix • Q is generated as a product of randomized plane rotations • A is formed as $Q\ D\ Q^T$	(LAPACK [4])
General symmetric tridiagonal Symmetric tridiagonal matrix	• Diagonal entries sampled uniformly from [−2.0, 2.0] • First off-diagonal entries	

without imposed weak links	sampled uniformly from [−2.0, 2.0] and mirrored	
Weak-link symmetric tridiagonal Deliberately inserted weak couplings	• Diagonal entries sampled uniformly from [−2.0, 2.0] • Ordinary first-band entries sampled uniformly from [−2.0, 2.0] • Every 8th coupling has random sign and magnitude in [2e-11, 1e-10] with the position shifted between samples • Conventional deflation uses $\tau = \varepsilon(\|d_1\| + \|d_2\|)$ • Initially $\tau \leq$ 4e-14 so even the smallest inserted weak link is at least 499.99999999999994 times the largest possible initial τ	(Demmel et al. [11])
Wilkinson tridiagonal	• For $n = 2m + 1$ • diagonal = m, m−1, ..., 1, 0, 1, ..., m−1, m • First off-diagonal entries equal 1	(Wilkinson [10])
Unequal variance-covariance Positive-definite with strongly unequal variances	• Begin with a dense correlation matrix R • Choose standard deviations s_i log-uniformly from [1.0, 10.0] • $S = \text{diag}(s_1, ..., s_n)$ • A = S R S • diagonal entries(A) range from 1.0 to 100.0	(Davies & Higham [5])

For each matrix family, 5 deterministic matrices are tested at n = 63, 255, 511, 1023.
For the Wilkinson family one matrix is tested at each size.

Common computational pathway

①

Preprocessing, common to both algorithms

For dense inputs, both algorithms begin with Householder tridiagonalization. Cost of tridiagonalization and corresponding eigenvector update is excluded from the comparison.

②

Convergence criteria, common to both algorithms

- A 1 x 1 block is complete
- A 2 x 2 block is diagonalized directly by a single Jacobi rotation

- A passive split is accepted if the subdiagonal coupling satisfies $|b| \leq \tau$, where $\tau = \varepsilon (|d_1| + |d_2|)$, and ε = 1e-14.
- An active split is accepted if both Jacobi rotation fill-ins satisfy $|sa| \leq \tau$ and $|sg| \leq \tau$.

③

Relative weights of arithmetic operations

Operation	Weight	Source if available
floating-point addition	1	(Fog [3])
floating-point subtraction	1	(Fog [3])
floating-point multiplication	1	(Fog [3])
floating-point division	4	(Fog [3])
$\sqrt{x}$	4	our estimate

④

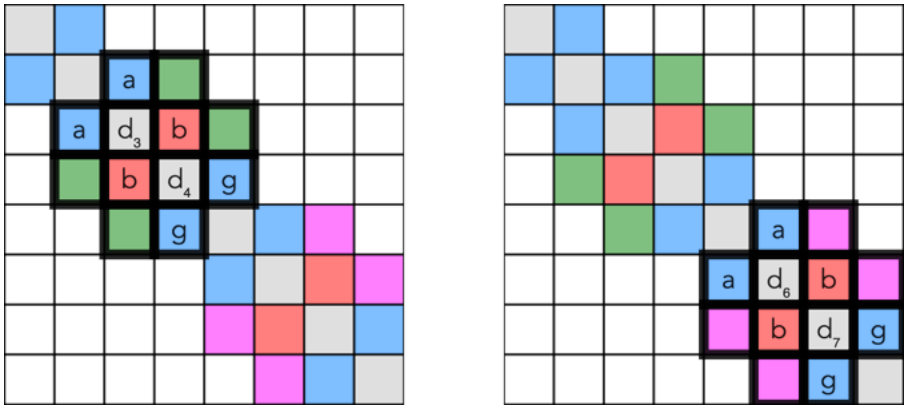
Shared arithmetic models

Quantity	Operation model	Weighted work	Source
hypot(x, y)	hardware-optimized	8	our estimate
Wilkinson shift calculation + initial use	4 additions+subtractions + 2 multiplications + 2 divisions + 1 hypot	4 + 2 + 2(4) + 1(8) = 22	our implementation and cost model
QR sweep m x m block including shift	6(m − 1) + 4 additions+subtractions + 8(m − 1) − 2 (for final-position multiplications) + 2 shift multiplications + 1(m − 1) + 2 divisions + 1(m − 1) + 1 hypot	26(m − 1) − 2 (for final-position multiplications) + 22 shift	(Reinsch [8]) same leading weighted term 26m in our sweep-position cost model
Eigenvector update in addition to QR sweep for m x m active block	(m − 1) eigenvector rotations × 6n operations	6n(m − 1)	our cost model
Conventional deflation test per cut	1 addition + 1 multiplication	2	our cost model
Terminal 2 x 2 diagonalization	5 + 2n additions+subtractions + 3 + 4n multiplications + 1 division + 2 hypot	5 + 2n + 3 + 4n + 1(4) + 2(8) = 28 + 6n	our cost model

Here n is the size of the original matrix and m is the size of the current active block.

How can we evaluate active splits in parallel?

More than one split within a block can be accepted before any rotation is performed,
provided that the data required by their tests do not intersect.



Legend

- highlighted on left image: entries affected by rotation at cut 3
- highlighted on right image: entries affected by rotation at cut 6
- gray: diagonal entries
- blue: first off-diagonal band
- red at cut 3: accepted active split
- red at cut 6: additional active split being tested

• green: fill-in around the first split, already found negligible
• magenta: fill-in that must be tested for the additional split

Here the candidate cuts are 3 and 6, so their highlighted neighborhoods are disjoint.
Cuts only two positions apart share one first-band coupling;
therefore simultaneous active splits must be at least three positions apart.

The candidate cuts can therefore be partitioned into three non-intersecting rows, where all candidates can be tested and applied in parallel.

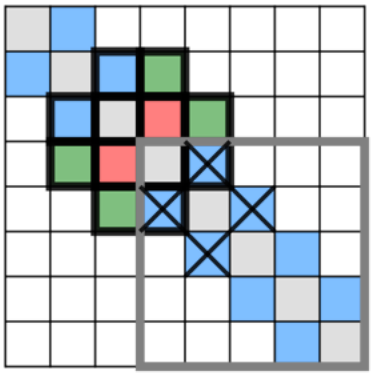
• row 1: i, i + 3, i + 6, ...
• row 2: i + 1, i + 4, i + 7, ...
• row 3: i + 2, i + 5, i + 8, ...

The three rows are processed successively.
After the rotations in one row are applied, the next row is tested using the updated matrix.

After the three rows have been processed, the new subblocks are rescanned before bulge chasing resumes.

① After an active split

• red: pivot zeroed by the rotation
• highlighted: other entries changed by the rotation
• wide outline: the child block
• green: the negligible fill-in from the rotation
• blue crossed positions: the candidate cuts to be re-tested



Why two new candidates?

For an interior candidate at entry j (crossed), the test uses itself, d_j, d_{j+1}, and its local couplings a and g.

- At the first cut, j = k + 1, the test uses the changed d_{k+1} and g from the completed rotation.
- At the second cut, j = k + 2, the same changed coupling g appears as that candidate's local a.
- For j ≥ k + 3, the test uses none of the changed quantities.

- By symmetry, only the last two cuts of a left child need to be retested.
- A child bounded by active splits on both sides requires retesting the first two and the last two cuts.

② After a passive split no rotation is performed

- no matrix entries inside either child are changed
- CAT testing was not performed after the passive split in the parent block
- therefore each child is passed back to the recursive block procedure with inspection scope = All

The recursive call therefore carries information on inspection scope:

- All: inspect every cut
- Left: after an active split, inspect the first two cuts
- Right: after an active split, inspect the last two cuts
- Left and right: after active splits on both sides, inspect the first two and the last

---

Arithmetic work of active split evaluation and execution

For an accepted active split, the implementation forms the Jacobi rotation as

- $\Delta = d_1 - d_2$

(For the usual accepted candidate, Δ has already been computed during testing and reused)

- $\sigma = 1$ if $\Delta \geq 0$, and -1 otherwise
- HalfDelta = $0.5\,\Delta$
- Radius = hypot(HalfDelta, b)
- Denominator = HalfDelta + σ Radius
- Norm = hypot(Denominator, b)
- $\text{InverseNorm} = \dfrac{1}{\text{Norm}}$
- c = |Denominator| InverseNorm
- s = σ b InverseNorm
- Midpoint = d_2 + HalfDelta
- $d_1' = \text{Midpoint} + \sigma\ \text{Radius}$
- $d_2' = \text{Midpoint} - \sigma\ \text{Radius}$

The rotation is applied to the local pivot block as shown below. The accepted fill-ins (−s a) and (s g) are written directly as zero.

$$T = \begin{bmatrix} \cdot & \bullet & 0 & 0 & 0 & 0 & 0 & 0 \\ \bullet & \cdot & a & 0 & 0 & 0 & 0 & 0 \\ 0 & a & d_1 & b & 0 & 0 & 0 & 0 \\ 0 & 0 & b & d_2 & g & 0 & 0 & 0 \\ 0 & 0 & 0 & g & \cdot & \bullet & 0 & 0 \\ 0 & 0 & 0 & 0 & \bullet & \cdot & \bullet & 0 \\ 0 & 0 & 0 & 0 & 0 & \bullet & \cdot & \bullet \\ 0 & 0 & 0 & 0 & 0 & 0 & \bullet & \cdot \end{bmatrix} \qquad G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & c & -s & 0 & 0 & 0 & 0 \\ 0 & 0 & s & c & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T' = G^{\intercal} T G = \begin{bmatrix} \cdot & \bullet & 0 & 0 & 0 & 0 & 0 & 0 \\ \bullet & \cdot & ca & -sa & 0 & 0 & 0 & 0 \\ 0 & ca & d_1' & 0 & sg & 0 & 0 & 0 \\ 0 & -sa & 0 & d_2' & cg & 0 & 0 & 0 \\ 0 & 0 & sg & cg & \cdot & \bullet & 0 & 0 \\ 0 & 0 & 0 & 0 & \bullet & \cdot & \bullet & 0 \\ 0 & 0 & 0 & 0 & 0 & \bullet & \cdot & \bullet \\ 0 & 0 & 0 & 0 & 0 & 0 & \bullet & \cdot \end{bmatrix}$$

Candidate testing

If $M \leq \tau$, the candidate is accepted without further screening arithmetic.

Otherwise $\Delta = d_1 - d_2$ is formed, at a cost of one subtraction.
The factors in $|b| M$ and $|\Delta| \tau$ are compared before either product is formed.
Only when those comparisons do not determine the result are the two products evaluated.

Testing path	Arithmetic work
$M \le \tau$	0
Decision before product evaluation	1
Two products required	3

Accepted split

The Jacobi construction costs 10 arithmetic operations when Δ was already formed by the screen,
and 11 operations when Δ must be formed during the rotation construction.
The base count includes the division and two hypot operations once.
Weighted work adds 3 for the division and 7 for each hypot.

The surviving first-band entry c a or c g adds one multiplication at an endpoint; an interior split has both entries and therefore adds two multiplications.

Every accepted active CAT split also applies the Jacobi rotation to the full $n \times n$ eigenvector matrix.
This costs $6n$ arithmetic operations.

Quantity	Weighted work breakdown	Total weighted work
Jacobi construction cached Δ	10 base arithmetic count + 1(3) + 2(7)	27
Jacobi construction no cached Δ	11 arithmetic operations + 1(3) + 2(7)	28
In-place update of A endpoint	1 surviving first-band multiplication	1
In-place update of A interior	2 surviving first-band multiplications	2
Eigenvector matrix update	2n additions or subtractions + 4n multiplications	6n
Endpoint cached Δ	27 + 1 + 6n	6n + 28
Interior cached Δ	27 + 2 + 6n	6n + 29
Endpoint no cached Δ	28 + 1 + 6n	6n + 29
Interior no cached Δ	28 + 2 + 6n	6n + 30

The normal successful screening path carries the already computed Δ into the Jacobi construction.
The $M \leq \tau$ path does not compute Δ during screening, so Δ is computed later during the Jacobi construction.

Parallel evaluation and execution of non-intersecting active splits were described in the preceding section.

Reported quantities

Ⓐ

For sweep positions and weighted arithmetic

- absolute reduction = baseline − candidate
- relative reduction = $\dfrac{\text{baseline} - \text{candidate}}{\text{baseline}}$

Ⓑ

For eigenvalue reference and error measurement

- Reference eigenvalues of the original generated matrix A are computed by LAPACK DSYEVD (LAPACK [7]).
- Errors are computed as the absolute difference from the LAPACK reference eigenvalue

and reported side by side for the candidate and the standard implementation

Measure	Source
Scaled eigenvalue error vs LAPACK	(DLMF [6])

①

Sweep positions

For an active block of size m, one complete QR sweep contains m − 1 sweep positions.
Total sweep positions are the sum of m − 1 over all active blocks and all sweeps.

②

Weighted arithmetic work W

Weighted arithmetic work is accumulated after each algorithmic event using the operation counts and weights shown in the tables for

- QR sweeps, including shift calculation and eigenvector updates
- conventional deflation tests
- terminal 2 x 2 rotations and eigenvector updates
- CAT candidate acceptance tests
- CAT rotations and eigenvector updates

Householder reduction, comparisons, and diagnostic-only checks are excluded.

③

Conventional and active deflations for the candidate method
(Counts are summed over all tested samples at each matrix size)

- Conventional deflations are ordinary QR deflations
- Active CAT deflations are one-rotation splits of active blocks
- Terminal 2 x 2 diagonalizations are excluded from both counts

$$\bullet\ \text{Active CAT fraction} = \frac{\text{active CAT deflations}}{\text{active CAT deflations} + \text{conventional deflations}}$$

④

$$\text{Mean} \ \frac{\text{smaller child block}}{\text{parent-block}} \ \text{ratio}$$

For a parent block of size $m > 2$, let n_l and n_r be the child sizes.

$$\text{smaller-child ratio} = \frac{\min\left(n_l, n_r\right)}{m}$$

$$\frac{1}{m} \le \text{smaller-child ratio} \le \frac{1}{2}$$

The mean ratio for each matrix size and family is reported for

- active splits, candidate
- passive splits, candidate
- passive splits, standard

⑤

Weighted work decomposition

The total weighted arithmetic work of the candidate method is decomposed into

- QR iteration
- CAT tests
- CAT rotations

Each component's percentage of the total is reported.

⑥

Scaled eigenvalue error vs LAPACK

$$\text{Scaled eigenvalue error} = \frac{|\hat{\lambda} - \lambda|}{\max(1, |\lambda|)} \quad \left(\text{DLMF [6]}\right).$$

For each family-size combination, we report the largest error among all eigenvalues of all tested matrices.

⑦

Eigenvector orthogonality error

$$\| V^T V - I \|_F$$

For each family-size combination, we report the largest error among all tested matrices.

⑧

Relative eigenpair residual

$$\frac{\| T V - V \Lambda \|_F}{\| T \|_F}$$

(computed on the tridiagonal matrix T).
For each family-size combination, we report the largest residual among all tested matrices.

Pseudocode

Input: real symmetric matrix A

Main procedure

T ← Householder tridiagonalization of A

V ← I

Apply the recursive block procedure below to the whole tridiagonal matrix, with inspection scope = All

---

①

Recursive block procedure receives
• an active block
• one of four scopes to inspect for deflation opportunities:
• All to inspect entire band
• Left to inspect 2 left band entries
• Right to inspect 2 right band entries
• Left and right to inspect 2 left and 2 right band entries

• If block size = 1
• return

• If block size = 2
• diagonalize by one Jacobi rotation
• apply the rotation to V
• return

• If block size > 2
• look for passive deflations within inspection scope
• If passive deflation opportunities occurred
• set every coupling below tolerance to zero
• partition the active block at those cuts into child blocks
• set inspection scope = All
• recurse child blocks at ①
• If passive deflation did not occur, look for active splits within inspection scope, in three non-intersecting rows
• Rows proceed in sequence
• Cuts within each row can be tested in parallel

• Rotations within each row can be applied in parallel

Row 1: cuts i, i + 3, i + 6, ...

• test all cuts in row 1

• accumulate every cut that satisfies the CAT acceptance test

• apply all accepted row-1 CAT rotations

• apply the corresponding rotations to V

Row 2: cuts i + 1, i + 4, i + 7, ...

• test every row 2 cut whose neighborhood does not intersect a CAT cut accepted in row 1

• accumulate every remaining cut that satisfies the acceptance test

• apply all accepted row 2 rotations

• apply the corresponding rotations to V

Row 3: cuts i + 2, i + 5, i + 8, ...

• test every row 3 cut whose neighborhood does not intersect a CAT cut accepted in row 1 or row 2

• accumulate every remaining cut that satisfies the acceptance test

• apply all accepted row 3 rotations

• apply the corresponding rotations to V

• If one or more CAT rotations are performed

• partition the active block at the CAT cuts

• assign each resulting child its required inspection scope as

• Left for a new left boundary

• Right for a new right boundary

• Left and right for new boundaries on both sides

• send each child to ①

• If no active or passive deflations

• perform one shifted QR sweep through the active block

• apply the sweep rotations to V

• set inspection scope = All

• return to ①

General dense symmetric

Sweep positions

n	baseline	candidate	relative reduction	absolute reduction
63	4,510.6	4,399	2.47%	111.6
255	68,404.2	66,580.4	2.67%	1,823.8
511	264,366.6	255,646.2	3.3%	8,720.4
1023	1,023,981.2	989,346.2	3.38%	34,635

Weighted arithmetic work W

n	baseline	candidate	relative reduction	absolute reduction
63	1,838,245.6	1,804,699.4	1.82%	33,546.2
255	106,650,081.6	103,997,731.8	2.49%	2,652,349.8
511	818,233,848.8	791,983,268.8	3.21%	26,250,580
1023	6,314,961,740.8	6,104,272,057.2	3.34%	210,689,683.6

Conventional and active deflations for the candidate method
(Counts are summed over all tested samples at each matrix size)

n	conventional	active CAT	$\frac{\text{active CAT}}{\text{all deflations}}$
63	277	28	9.18%
255	1,135	130	10.28%
511	2,286	259	10.18%
1023	4,564	537	10.53%

$$\text{Mean } \frac{\text{smaller child block}}{\text{parent-block}} \text{ ratio}$$

n	active candidate	passive candidate	passive standard
63	6.37%	5.18%	5.29%
255	2.61%	1.74%	1.83%
511	1.07%	1.04%	1.04%
1023	0.86%	0.59%	0.63%

Weighted work decomposition

n	component	W
63	QR iteration	99.33%
	CAT tests	0.55%
	CAT rotations	0.13%
255	QR iteration	99.81%
	CAT tests	0.15%
	CAT rotations	0.04%
511	QR iteration	99.9%
	CAT tests	0.08%
	CAT rotations	0.02%
1023	QR iteration	99.95%
	CAT tests	0.04%
	CAT rotations	0.01%

Scaled eigenvalue error vs LAPACK

n	baseline	candidate
63	8.27e-15	8.27e-15
255	2.22e-14	2.48e-14
511	1.18e-14	1.27e-14
1023	6.54e-14	6.54e-14

Eigenvector orthogonality error

n	baseline	candidate
63	1.79e-14	1.89e-14
255	7.45e-14	6.95e-14
511	1.43e-13	1.38e-13
1023	2.80e-13	2.69e-13

Relative eigenpair residual

n	baseline	candidate
63	3.24e-15	7.49e-15
255	5.91e-15	7.25e-15
511	7.40e-15	8.32e-15
1023	9.44e-15	9.99e-15

Dense correlation

Sweep positions

n	baseline	candidate	relative reduction	absolute reduction
63	4,257	4,150.8	2.49%	106.2
255	65,809.6	63,055.4	4.19%	2,754.2
511	251,149	241,519.2	3.83%	9,629.8
1023	962,795	928,868.4	3.52%	33,926.6

Weighted arithmetic work W

n	baseline	candidate	relative reduction	absolute reduction
63	1,735,136	1,702,913.8	1.86%	32,222.2
255	102,607,508.8	98,496,722	4.01%	4,110,786.8
511	777,338,560	748,224,783.8	3.75%	29,113,776.2
1023	5,937,684,024.4	5,731,188,821.2	3.48%	206,495,203.2

Conventional and active deflations for the candidate method
(Counts are summed over all tested samples at each matrix size)

n	conventional	active CAT	$\frac{\text{active CAT}}{\text{all deflations}}$
63	273	32	10.49%
255	1,105	159	12.58%
511	2,228	315	12.39%
1023	4,480	621	12.17%

$$\text{Mean } \frac{\text{smaller child block}}{\text{parent-block}} \text{ ratio}$$

n	active candidate	passive candidate	passive standard
63	7.96%	4.98%	5.29%
255	2.71%	1.73%	1.84%
511	1.14%	1.07%	1.08%
1023	0.68%	0.57%	0.63%

Weighted work decomposition

n	component	W
63	QR iteration	99.34%
	CAT tests	0.51%
	CAT rotations	0.15%
255	QR iteration	99.81%
	CAT tests	0.14%
	CAT rotations	0.05%
511	QR iteration	99.9%
	CAT tests	0.07%
	CAT rotations	0.03%
1023	QR iteration	99.95%
	CAT tests	0.04%
	CAT rotations	0.01%

Scaled eigenvalue error vs LAPACK

n	baseline	candidate
63	2.33e-15	2.69e-15
255	3.66e-15	4.62e-15
511	5.00e-15	5.75e-15
1023	7.88e-15	6.77e-15

Eigenvector orthogonality error

n	baseline	candidate
63	1.83e-14	2.00e-14
255	6.93e-14	6.81e-14
511	1.36e-13	1.33e-13
1023	2.67e-13	2.58e-13

Relative eigenpair residual

n	baseline	candidate
63	6.00e-15	6.10e-15
255	4.28e-15	5.11e-15
511	5.14e-15	5.55e-15
1023	5.16e-15	6.00e-15

Dense clustered spectrum

Sweep positions

n	baseline	candidate	relative reduction	absolute reduction
63	3,072.4	2,928.6	4.68%	143.8
255	44,922	43,646.6	2.84%	1,275.4
511	173,848.8	170,206.6	2.1%	3,642.2
1023	677,716.2	664,926.2	1.89%	12,790

Weighted arithmetic work W

n	baseline	candidate	relative reduction	absolute reduction
63	1,254,155.6	1,206,135	3.83%	48,020.6
255	70,048,989.6	68,214,883	2.62%	1,834,106.6
511	538,100,652.8	527,415,054.2	1.99%	10,685,598.6
1023	4,179,612,662.4	4,103,130,476.6	1.83%	76,482,185.8

Conventional and active deflations for the candidate method
(Counts are summed over all tested samples at each matrix size)

n	conventional	active CAT	active CAT / all deflations
63	246	44	15.17%
255	1,075	171	13.72%
511	2,205	313	12.43%
1023	4,365	681	13.5%

Mean $\frac{\text{smaller child block}}{\text{parent-block}}$ ratio

n	active candidate	passive candidate	passive standard
63	16.16%	9.57%	10.4%
255	6.83%	4.38%	4.75%
511	4.33%	2.76%	2.96%
1023	2.98%	1.61%	1.77%

Weighted work decomposition

n	component	W
63	QR iteration	99.11%
	CAT tests	0.59%
	CAT rotations	0.3%
255	QR iteration	99.77%
	CAT tests	0.15%
	CAT rotations	0.08%
511	QR iteration	99.89%
	CAT tests	0.08%
	CAT rotations	0.04%
1023	QR iteration	99.94%
	CAT tests	0.04%
	CAT rotations	0.02%

Scaled eigenvalue error vs LAPACK

n	baseline	candidate
63	3.41e-15	3.70e-15
255	8.44e-15	8.29e-15
511	1.00e-14	1.03e-14
1023	2.88e-14	2.88e-14

Eigenvector orthogonality error

n	baseline	candidate
63	1.65e-14	1.80e-14
255	5.94e-14	5.95e-14
511	1.15e-13	1.16e-13
1023	2.26e-13	2.26e-13

Relative eigenpair residual

n	baseline	candidate
63	5.01e-15	8.01e-15
255	5.21e-15	7.35e-15
511	6.56e-15	8.19e-15
1023	8.08e-15	9.02e-15

Dense graded spectrum

Sweep positions

n	baseline	candidate	relative reduction	absolute reduction
63	2,882	2,643	8.29%	239
255	46,479.8	41,732	10.21%	4,747.8
511	185,770.6	165,225.4	11.06%	20,545.2
1023	743,769.4	655,959.8	11.81%	87,809.6

Weighted arithmetic work W

n	baseline	candidate	relative reduction	absolute reduction
63	1,175,950	1,090,503	7.27%	85,447
255	72,488,154.4	65,261,210	9.97%	7,226,944.4
511	575,050,574.4	512,151,173.4	10.94%	62,899,401
1023	4,587,159,126.4	4,048,405,271.2	11.74%	538,753,855.2

Conventional and active deflations for the candidate method
(Counts are summed over all tested samples at each matrix size)

n	conventional	active CAT	$\frac{\text{active CAT}}{\text{all deflations}}$
63	235	69	22.7%
255	977	274	21.9%
511	1,936	579	23.02%
1023	3,775	1,271	25.19%

$$\text{Mean} \quad \frac{\text{smaller child block}}{\text{parent-block}} \quad \text{ratio}$$

n	active candidate	passive candidate	passive standard
63	7.17%	4.87%	5.29%
255	3.12%	1.95%	1.83%
511	2.22%	1.27%	1.04%
1023	1.29%	0.82%	0.59%

Weighted work decomposition

n	component	W
63	QR iteration	98.85%
	CAT tests	0.64%
	CAT rotations	0.51%
255	QR iteration	99.72%
	CAT tests	0.15%
	CAT rotations	0.13%
511	QR iteration	99.86%
	CAT tests	0.07%
	CAT rotations	0.07%
1023	QR iteration	99.93%
	CAT tests	0.03%
	CAT rotations	0.04%

Scaled eigenvalue error vs LAPACK

n	baseline	candidate
63	6.02e-14	6.03e-14
255	5.73e-14	5.74e-14
511	5.90e-14	5.90e-14
1023	7.29e-14	7.24e-14

Eigenvector orthogonality error

n	baseline	candidate
63	1.80e-14	1.62e-14
255	6.11e-14	5.68e-14
511	1.22e-13	1.12e-13
1023	2.50e-13	2.21e-13

Relative eigenpair residual

n	baseline	candidate
63	1.00e-15	1.78e-14
255	1.86e-15	1.68e-14
511	2.78e-15	1.49e-14
1023	4.12e-15	1.51e-14

General symmetric tridiagonal

Sweep positions

n	baseline	candidate	relative reduction	absolute reduction
63	4,717	4,677.2	0.84%	39.8
255	73,982.6	72,014.6	2.66%	1,968
511	297,682.4	289,968.8	2.59%	7,713.6
1023	1,120,001.6	1,098,933.6	1.88%	21,068

Weighted arithmetic work W

n	baseline	candidate	relative reduction	absolute reduction
63	1,922,163.2	1,918,114.6	0.21%	4,048.6
255	115,343,073.2	112,474,497.2	2.49%	2,868,576
511	921,324,614.8	898,224,370	2.51%	23,100,244.8
1023	6,907,001,366.8	6,780,068,100.8	1.84%	126,933,266

Conventional and active deflations for the candidate method
(Counts are summed over all tested samples at each matrix size)

n	conventional	active CAT	active CAT / all deflations
63	281	24	7.87%
255	1,113	139	11.1%
511	2,280	241	9.56%
1023	4,530	512	10.15%

Mean $\frac{\text{smaller child block}}{\text{parent-block}}$ ratio

n	active candidate	passive candidate	passive standard
63	8.4%	5.03%	5.26%
255	4.2%	2.69%	2.73%
511	3.53%	1.92%	1.95%
1023	3.68%	1.82%	1.94%

Weighted work decomposition

n	component	W
63	QR iteration	99.36%
	CAT tests	0.54%
	CAT rotations	0.1%
255	QR iteration	99.82%
	CAT tests	0.14%
	CAT rotations	0.04%
511	QR iteration	99.91%
	CAT tests	0.07%
	CAT rotations	0.02%
1023	QR iteration	99.96%
	CAT tests	0.04%
	CAT rotations	0.01%

Scaled eigenvalue error vs LAPACK

n	baseline	candidate
63	5.66e-15	5.44e-15
255	1.10e-14	1.14e-14
511	1.63e-14	2.03e-14
1023	2.80e-14	2.66e-14

Eigenvector orthogonality error

n	baseline	candidate
63	1.89e-14	1.99e-14
255	8.12e-14	7.71e-14
511	1.52e-13	1.51e-13
1023	2.98e-13	2.89e-13

Relative eigenpair residual

n	baseline	candidate
63	6.43e-15	7.44e-15
255	6.73e-15	8.81e-15
511	8.44e-15	9.68e-15
1023	9.88e-15	1.10e-14

Weak-link symmetric tridiagonal

Sweep positions

n	baseline	candidate	relative reduction	absolute reduction
63	4,192.8	3,928.8	6.3%	264
255	66,363.4	59,745.6	9.97%	6,617.8
511	242,933.2	228,201.8	6.06%	14,731.4
1023	698,696	541,684.6	22.47%	157,011.4

Weighted arithmetic work W

n	baseline	candidate	relative reduction	absolute reduction
63	1,709,102.8	1,614,817.4	5.52%	94,285.4
255	103,473,004.4	93,354,318.2	9.78%	10,118,686.2
511	751,908,724	707,093,826.6	5.96%	44,814,897.4
1023	4,308,953,758	3,342,789,977.2	22.42%	966,163,780.8

Conventional and active deflations for the candidate method
(Counts are summed over all tested samples at each matrix size)

n	conventional	active CAT	$\frac{\text{active CAT}}{\text{all deflations}}$
63	262	43	14.1%
255	1,087	166	13.25%
511	2,161	355	14.11%
1023	4,398	635	12.62%

$$\text{Mean } \frac{\text{smaller child block}}{\text{parent-block}} \text{ ratio}$$

n	active candidate	passive candidate	passive standard
63	5.98%	5.18%	5.29%
255	3.35%	2.68%	2.48%
511	2.87%	2.25%	2.27%
1023	4.28%	2.78%	2.45%

Weighted work decomposition

n	component	W
63	QR iteration	99.16%
	CAT tests	0.62%
	CAT rotations	0.22%
255	QR iteration	99.78%
	CAT tests	0.16%
	CAT rotations	0.06%
511	QR iteration	99.89%
	CAT tests	0.08%
	CAT rotations	0.03%
1023	QR iteration	99.94%
	CAT tests	0.04%
	CAT rotations	0.02%

Scaled eigenvalue error vs LAPACK

n	baseline	candidate
63	7.12e-15	8.33e-15
255	2.13e-14	1.37e-14
511	2.52e-14	2.75e-14
1023	3.65e-14	3.28e-14

Eigenvector orthogonality error

n	baseline	candidate
63	1.82e-14	1.95e-14
255	7.32e-14	7.27e-14
511	1.38e-13	1.42e-13
1023	2.78e-13	2.48e-13

Relative eigenpair residual

n	baseline	candidate
63	3.96e-15	7.68e-15
255	5.79e-15	7.44e-15
511	6.92e-15	7.68e-15
1023	7.30e-15	8.62e-15

Wilkinson tridiagonal

Sweep positions

n	baseline	candidate	relative reduction	absolute reduction
63	4,040	3,866	4.31%	174
255	52,275	44,050	15.73%	8,225
511	176,855	154,835	12.45%	22,020
1023	611,034	579,244	5.2%	31,790

Weighted arithmetic work W

n	baseline	candidate	relative reduction	absolute reduction
63	1,646,716	1,585,580	3.71%	61,136
255	81,522,070	68,980,592	15.38%	12,541,478
511	547,470,154	480,480,975	12.24%	66,989,179
1023	3,768,723,550	3,575,512,388	5.13%	193,211,162

Conventional and active deflations for the candidate method

(Counts are summed over all tested samples at each matrix size)

n	conventional	active CAT	$\frac{\text{active CAT}}{\text{all deflations}}$
63	57	4	6.56%
255	113	136	54.62%
511	187	320	63.12%
1023	711	306	30.09%

$$\text{Mean } \frac{\text{smaller child block}}{\text{parent-block}} \text{ ratio}$$

n	active candidate	passive candidate	passive standard
63	4.41%	5.72%	5.64%
255	1.43%	2.65%	1.86%
511	0.43%	2.11%	1.05%
1023	0.33%	0.7%	0.59%

Weighted work decomposition

n	component	W
63	QR iteration	99.38%
	CAT tests	0.52%
	CAT rotations	0.1%
255	QR iteration	99.55%
	CAT tests	0.15%
	CAT rotations	0.31%
511	QR iteration	99.72%
	CAT tests	0.07%
	CAT rotations	0.21%
1023	QR iteration	99.91%
	CAT tests	0.04%
	CAT rotations	0.05%

Scaled eigenvalue error vs LAPACK

n	baseline	candidate
63	2.70e-14	3.79e-14
255	5.31e-14	6.36e-14
511	6.85e-14	4.74e-14
1023	2.20e-13	2.50e-13

Eigenvector orthogonality error

n	baseline	candidate
63	1.69e-14	1.70e-14
255	6.68e-14	6.10e-14
511	1.26e-13	1.15e-13
1023	2.43e-13	2.16e-13

Relative eigenpair residual

n	baseline	candidate
63	4.01e-15	5.28e-15
255	7.19e-15	1.08e-14
511	1.01e-14	9.40e-15
1023	9.83e-15	8.08e-15

Unequal variance-covariance

Sweep positions

n	baseline	candidate	relative reduction	absolute reduction
63	3,688	3,547.2	3.82%	140.8
255	57,325	54,884.8	4.26%	2,440.2
511	223,559.8	213,831.4	4.35%	9,728.4
1023	870,186	828,439.6	4.8%	41,746.4

Weighted arithmetic work W

n	baseline	candidate	relative reduction	absolute reduction
63	1,503,668.8	1,456,730.6	3.12%	46,938.2
255	89,386,816	85,726,245	4.1%	3,660,571
511	691,975,983.2	662,413,146.4	4.27%	29,562,836.8
1023	5,366,659,071.6	5,111,292,730.2	4.76%	255,366,341.4

Conventional and active deflations for the candidate method
(Counts are summed over all tested samples at each matrix size)

n	conventional	active CAT	$\frac{\text{active CAT}}{\text{all deflations}}$
63	265	39	12.83%
255	1,103	155	12.32%
511	2,174	363	14.31%
1023	4,319	764	15.03%

$$\text{Mean} \ \frac{\text{smaller child block}}{\text{parent-block}} \ \text{ratio}$$

n	active candidate	passive candidate	passive standard
63	4.59%	5.55%	5.3%
255	3.01%	2.01%	1.96%
511	1.8%	1.16%	1.16%
1023	1.14%	0.71%	0.65%

Weighted work decomposition

n	component	W
63	QR iteration	99.27%
	CAT tests	0.52%
	CAT rotations	0.22%
255	QR iteration	99.83%
	CAT tests	0.11%
	CAT rotations	0.06%
511	QR iteration	99.91%
	CAT tests	0.05%
	CAT rotations	0.03%
1023	QR iteration	99.96%
	CAT tests	0.02%
	CAT rotations	0.02%

Scaled eigenvalue error vs LAPACK

n	baseline	candidate
63	2.98e-14	2.98e-14
255	3.86e-14	3.20e-14
511	6.15e-14	3.16e-14
1023	7.47e-14	1.04e-13

Eigenvector orthogonality error

n	baseline	candidate
63	1.94e-14	1.74e-14
255	6.59e-14	6.47e-14
511	1.29e-13	1.27e-13
1023	2.56e-13	2.47e-13

Relative eigenpair residual

n	baseline	candidate
63	2.18e-15	5.35e-15
255	4.80e-15	8.50e-15
511	5.62e-15	7.61e-15
1023	6.40e-15	8.19e-15

Conclusions

Chase And Trim is a method for creating valid matrix splits earlier than conventional QR deflation, which

- reduces the amount of subsequent QR iteration on the larger child block
- allows the smaller child block to be processed independently of the larger one

①
Structural effect

Relative reduction of number of QR sweep positions,
defined as the sum of $m - 1$ over all QR sweeps on all active blocks of size m,
ranged

- from 0.8% for General symmetric tridiagonal, n = 63,
- to 22.5% for Weak-link symmetric tridiagonal, n = 1023.
- The graded-spectrum family also showed a consistently large average reduction of 10.3%.

The comparatively large reductions for the weak-link and graded-spectrum families
are consistent with the local rotation acceptance criterion by favoring smaller rotations

- Weak links directly provide small candidate couplings $|b|$
- Graded spectra tend to produce larger local diagonal separations $|\Delta|$ as QR progresses
- Small neighboring couplings $|a|$ and $|g|$ strengthen the effect further.

②
Weighted arithmetic work

Relative reduction of weighted arithmetic work ranged

- from 0.2% for General symmetric tridiagonal, n = 63,
- to 22.4% for Weak-link symmetric tridiagonal, n = 1023.

Why weighted work tracks sweep positions

For an active block of size m, the charged QR sweep work is
$W(\text{sweep}) = (26 + 6n)(m - 1) + 20.$

Therefore, if $P = \Sigma(m - 1)$ over all QR sweeps and S is the number of sweeps,
$W(\text{QR sweeps}) = (26 + 6n)P + 20S.$

The dominant QR work is therefore proportional to the sweep-position count P.
CAT reduces P through earlier splitting, while testing and accepted rotations consume part of that saving,
so the weighted-work reduction closely tracks, but does not exactly equal, the sweep-position reduction.

③
Candidate algorithm overhead

The additional computation consists of:
• testing candidate cuts, which was reduced algebraically to a small number of arithmetic operations
• constructing and applying successful one-rotation splits

Across the tested cases, the largest combined weighted work CAT cost was 1.2%.

④
Numerical accuracy

A successful CAT rotation creates only two disjoint symmetric fill-in pairs.
The algorithm requires both fill magnitudes to be no larger than the same local tolerance used for deflation.
Because the two discarded symmetric pairs involve disjoint index pairs, the spectral norm of the discarded matrix
is the larger fill magnitude, and is therefore also no larger than that local

tolerance.

ⓐ Largest scaled eigenvalue error versus LAPACK
- standard: 2.20e-13
- candidate: 2.50e-13

ⓑ Largest eigenvector orthogonality error $\| V^T V - I \|_F$
- standard: 2.98e-13
- candidate: 2.89e-13

ⓒ Largest relative eigenpair residual $\dfrac{\| T V - V \Lambda \|_F}{\| T \|_F}$
- standard: 1.01e-14
- candidate: 1.78e-14

Limitations

- The reported weighted-work reductions apply only to the eigenvector-accumulating branch,
where the reduction in QR sweep work exceeded the CAT overhead in every tested family-size average.

- No arithmetic-span claim.
The CAT screening was organized to expose parallelism,
but we did not report span because a transparent accounting would require additional assumptions about concurrency and scheduling.

- CAT was not tested in conjunction with aggressive early deflation (AED) (Braman, Byers & Mathias [13]).

Production QR eigensolvers invoke AED selectively rather than at every sweep, so CAT and AED address deflation at different steps and are not inherently contradictory.

• The conclusions are limited to the tested matrix families, matrix sizes, and tolerance.

Other applications

The potential extension is bidiagonal QR iteration for the singular value decomposition,
where the corresponding local two-sided transformations have similarly sparse cross-cut fill.
This suggests testing whether the same principle can produce earlier SVD deflation.

References

① D. S. Watkins
Fundamentals of Matrix Computations, 3rd edition, Section 5.6: Francis's Algorithm
John Wiley & Sons, 2010.

② LAPACK
DSTEQR: Symmetric Tridiagonal Eigensolver Using Implicit QL or QR Iteration
LAPACK 3.12.1 documentation, Netlib.
https://www.netlib.org/lapack/explore-html/d3/d16/group__steqr.html

③ Agner Fog
Instruction Tables: Lists of Instruction Latencies, Throughputs and Micro-operation

Breakdowns for Intel, AMD and VIA CPUs
https://www.agner.org/optimize/instruction_tables.pdf

④ LAPACK/ScaLAPACK testing literature
Test matrices for symmetric eigenvalue problems include prescribed spectra
with arithmetic, geometric, and clustered eigenvalue distributions,
combined with orthogonal similarity transformations.
https://www.netlib.org/lapack/lawnspdf/lawn18.pdf

⑤ Philip I. Davies and Nicholas J. Higham
Numerically Stable Generation of Correlation Matrices and Their Factors
BIT Numerical Mathematics, 40(4), 640–651, 2000.
https://eprints.maths.manchester.ac.uk/323/

⑥ NIST Digital Library of Mathematical Functions
Section 3.1(v): Error Measures
Defines the mollified error $\varepsilon_m = |x^* - x|/\max(|x|, 1)$.
https://dlmf.nist.gov/3.1#v

⑦ LAPACK
DSYEVD: Symmetric Eigensolver
Computes all eigenvalues and, optionally, eigenvectors of a real symmetric matrix.
https://www.netlib.org/lapack/explore-html/d8/d30/
group__heevd_ga25b71a69f9921df0a6050aa5883f54f4.html

⑧ Christian H. Reinsch
A Stable, Rational QR Algorithm for the Computation of the Eigenvalues of an
Hermitian, Tridiagonal Matrix
Mathematics of Computation, 25(115), 591–597, 1971.
For a natural QR iteration on a symmetric tridiagonal matrix, gives
4n additions, 10n multiplications, 2n divisions, and n square roots per iteration.
DOI: 10.1090/S0025-5718-1971-0295555-4

⑨ Gene H. Golub and Charles F. Van Loan

Matrix Computations, 4th edition
Johns Hopkins University Press, 2013.
Standard reference for induced matrix norms, including the spectral norm $\|A\|_2$.

⑩ J. H. Wilkinson
The Algebraic Eigenvalue Problem
Clarendon Press, Oxford, 1965.

⑪ James W. Demmel, Osni A. Marques, Beresford N. Parlett, and Christof Vömel
A Testing Infrastructure for LAPACK's Symmetric Eigensolvers
LAPACK Working Note 182, April 2007.
Defines general glued symmetric tridiagonal matrices as tridiagonal blocks connected by interface glue factors.
https://www.netlib.org/lapack/lawnspdf/lawn182.pdf

⑫ Françoise Tisseur and Jack Dongarra
Parallelizing the Divide and Conquer Algorithm for the Symmetric Tridiagonal Eigenvalue Problem on Distributed Memory Architectures
LAPACK Working Note 132, 1998.
Describes rotation-based deflation of the diagonal-plus-rank-one problem when two poles are close.
https\://www\.netlib.org/lapack/lawnspdf/lawn132.pdf

⑬ Karen Braman, Ralph Byers, and Roy Mathias
The Multishift QR Algorithm. Part II: Aggressive Early Deflation
SIAM Journal on Matrix Analysis and Applications, 23(4), 948–973, 2002.
DOI: 10.1137/S0895479801384585