

Introduction

The bidiagonal QR route for computing SVD of an $m \times n$ ($m \geq n$) matrix A proceeds as follows: (Watkins [1])

- ① Reduce A to upper bidiagonal form by left and right Householder transformations.

The resulting $m \times n$ bidiagonal matrix has zero rows below the leading $n \times n$ bidiagonal block;

this square bidiagonal block B is retained for phase 2.

- ② Choose a shift and apply a sweep of alternating right and left Givens rotations within the active block,

chasing the temporary bulge from top to bottom and restoring bidiagonal form.

- ③ Scan the superdiagonal for conventional deflation opportunities between sweeps.

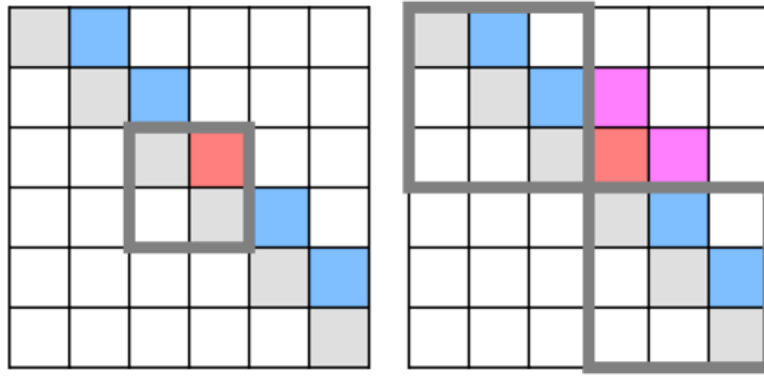
Whenever a superdiagonal entry becomes sufficiently small, set it to zero and split the block at that position.

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

If $m < n$, apply the algorithm to A^T instead, so the singular values remain unchanged, and the left and right singular vectors exchange roles.

Divide-and-conquer is an alternative method for computing the bidiagonal SVD.

Bidiagonal QR nevertheless remains implemented in LAPACK DBDSQR, which supports computation and accumulation of the left and right singular vectors. (LAPACK [2]) CAT targets this QR method; comparison with divide-and-conquer is outside the present study.



We investigate if the same terminal 2×2 diagonalization can be applied at a non-negligible coupling.

The transformation diagonalizes the 2×2 bidiagonal block (selected in the left image)

and annihilates the selected superdiagonal entry (red).

The transformation creates up to two new fill-in entries, shown in purple. If both new entries are below the accepted tolerance, they may be discarded and allow an early split.

We call this augmented bidiagonal QR strategy Chase And Trim (CAT).

In between bulge chases, conventional deflations are applied as usual.

After that, the remaining couplings are examined for active deflation feasibility. When a split is accepted, both rotations are applied and the child blocks continue recursively.

We compare the two algorithms for computing singular values and singular vectors.

① Standard bidiagonal QR with only conventional deflation.

② CAT in addition to ①.

- Examines the rotation associated with the larger outer coupling.
- Uses two inexpensive sufficient conditions to reject infeasible candidates.
- Candidates surviving both screens are passed to LAPACK DLASV2 (LAPACK [6])

The resulting left and right rotations are used to verify both cross-boundary fill-ins.
If both satisfy the tolerance, the rotations are applied and the block is split.



LBR product for active deflation

Let B be an $n \times n$ upper-bidiagonal matrix: $B =$

$$\begin{bmatrix} d_1 & e_1 & 0 & 0 & 0 & 0 \\ 0 & d_2 & e_2 & 0 & 0 & 0 \\ 0 & 0 & d_3 & e_3 & 0 & 0 \\ 0 & 0 & 0 & d_4 & e_4 & 0 \\ 0 & 0 & 0 & 0 & d_5 & e_5 \\ 0 & 0 & 0 & 0 & 0 & d_6 \end{bmatrix}$$

For the displayed example, $n = 6$, $k = 3$, the selected coupling is e_3 , between d_3 and d_4 .

Extract the selected 2×2 block: $K =$

$$\begin{bmatrix} d_3 & e_3 \\ 0 & d_4 \end{bmatrix}$$

Compute the local 2×2 SVD to obtain left and right rotations L_2 and R_2 .
These rotations diagonalize K:

$$L_2 K R_2 = \left[\begin{array}{c|c} \sigma_3 & 0 \\ \hline 0 & \sigma_4 \end{array} \right]$$

The diagonal values may be signed, depending on the rotation convention.
Embed L_2 and R_2 into $n \times n$ identity matrices L and R , acting only on the selected rows and columns.

We now examine their effect on the full bidiagonal matrix, including neighboring couplings.

First, apply the right rotation R :

$$B R = \left[\begin{array}{c|c|c|c|c|c} d_1 & e_1 & 0 & 0 & 0 & 0 \\ \hline 0 & d_2 & e_2 & 0 & 0 & 0 \\ \hline 0 & 0 & d_3 & e_3 & 0 & 0 \\ \hline 0 & 0 & 0 & d_4 & e_4 & 0 \\ \hline 0 & 0 & 0 & 0 & d_5 & e_5 \\ \hline 0 & 0 & 0 & 0 & 0 & d_6 \end{array} \right] \times \left[\begin{array}{c|c|c|c|c|c} 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & c^R & -s^R & 0 & 0 \\ \hline 0 & 0 & s^R & c^R & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right] =$$

$$\left[\begin{array}{c|c|c|c|c|c} d_1 & e_1 & 0 & 0 & 0 & 0 \\ \hline 0 & d_2 & c^R e_2 & -s^R e_2 & 0 & 0 \\ \hline 0 & 0 & c^R d_3 + s^R e_3 & c^R e_3 - s^R d_3 & 0 & 0 \\ \hline 0 & 0 & s^R d_4 & c^R d_4 & e_4 & 0 \\ \hline 0 & 0 & 0 & 0 & d_5 & e_5 \\ \hline 0 & 0 & 0 & 0 & 0 & d_6 \end{array} \right]$$

Next, apply the left rotation L:

$$L B R = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & c^L & s^L & 0 & 0 \\ 0 & 0 & -s^L & c^L & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} d_1 & e_1 & 0 & 0 & 0 & 0 \\ 0 & d_2 & c^R e_2 & -s^R e_2 & 0 & 0 \\ 0 & 0 & c^R d_3 + s^R e_3 & c^R e_3 - s^R d_3 & 0 & 0 \\ 0 & 0 & s^R d_4 & c^R d_4 & e_4 & 0 \\ 0 & 0 & 0 & 0 & d_5 & e_5 \\ 0 & 0 & 0 & 0 & 0 & d_6 \end{bmatrix} =$$

$$\begin{bmatrix} d_1 & e_1 & 0 & 0 & 0 & 0 \\ 0 & d_2 & c^R e_2 & -s^R e_2 & 0 & 0 \\ 0 & 0 & c^L (c^R d_3 + s^R e_3) + s^L s^R d_4 & c^L (c^R e_3 - s^R d_3) + s^L c^R d_4 & s^L e_4 & 0 \\ 0 & 0 & -s^L (c^R d_3 + s^R e_3) + c^L s^R d_4 & -s^L (c^R e_3 - s^R d_3) + c^L c^R d_4 & c^L e_4 & 0 \\ 0 & 0 & 0 & 0 & d_5 & e_5 \\ 0 & 0 & 0 & 0 & 0 & d_6 \end{bmatrix}$$

By construction, the local SVD rotations annihilate both off-diagonal entries of the selected block:

$$\begin{aligned} -s^L (c^R d_3 + s^R e_3) + c^L s^R d_4 &= 0 \\ c^L (c^R e_3 - s^R d_3) + s^L c^R d_4 &= 0. \end{aligned}$$

$$L B R = \begin{bmatrix} d_1 & e_1 & 0 & 0 & 0 & 0 \\ 0 & d_2 & c^R e_2 & -s^R e_2 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 & s^L e_4 & 0 \\ 0 & 0 & 0 & \sigma_4 & c^L e_4 & 0 \\ 0 & 0 & 0 & 0 & d_5 & e_5 \\ 0 & 0 & 0 & 0 & 0 & d_6 \end{bmatrix}$$

The selected coupling is eliminated, but two cross-boundary fill-ins remain:

- left fill = $-s^R e_2$
- right fill = $s^L e_4$.

The same fill-ins result whether we compute $L(B R)$ or $(L B)R$.

Early deflation is feasible when both fill-in magnitudes satisfy the local DBDSQR deflation tolerance τ (LAPACK [2]):

- $|-s^R e_2| \leq \tau$
- $|s^L e_4| \leq \tau$.



From fill-in to CAT inequalities

Consider the same upper-bidiagonal matrix: $B =$

$$\begin{bmatrix} d_1 & e_1 & 0 & 0 & 0 & 0 \\ 0 & d_2 & e_2 & 0 & 0 & 0 \\ 0 & 0 & d_3 & e_3 & 0 & 0 \\ 0 & 0 & 0 & d_4 & e_4 & 0 \\ 0 & 0 & 0 & 0 & d_5 & e_5 \\ 0 & 0 & 0 & 0 & 0 & d_6 \end{bmatrix}$$

Suppose we seek to rotate and eliminate the coupling e_3 , between d_3 and d_4 .

Denote

- $L = |e_2|$
- $R = |e_4|$

(L and R denote neighboring coupling magnitudes, not the rotation matrices of the preceding section)

- τ = directional deflation tolerance for the selected index (LAPACK [2]).

Define μ_k as the recurrence value for deflation at e_k .

If the chase proceeds left-to-right:

- $\mu_1 = |d_1|$
- $\mu_{k+1} = \frac{|d_{k+1}| \mu_k}{\mu_k + |e_k|}$

- (deflation threshold at e_k) or $\tau_k = \text{tol } \mu_k$

If the chase proceeds right-to-left inside a block with h columns:

- $\mu_h = |d_h|$
- $\mu_k = \frac{|d_k| \mu_{k+1}}{\mu_{k+1} + |e_k|}$

- (deflation threshold at e_k) or $\tau_k = \text{tol } \mu_{k+1}$

(note that $\mu_k \geq 0$ for all k and, for the cut at e_k , we denote τ_k by τ)

The two rotations create the fill-ins identified earlier:

- the right rotation creates a left fill-in of magnitude $L \mid s^R \mid$
- the left rotation creates a right fill-in of magnitude $R \mid s^L \mid$

For active deflation, both magnitudes must be no larger than τ .

Since $L(BR) = (LB)R$, the left and right rotations may be constructed in either order.

Suppose we start with the right rotation, which is

- constructed from diagonalization of $K^T K$
- allows deflation only if $L \mid s^R \mid \leq \tau$

For the selected e , define quantities of local $K^T K = \left[\begin{array}{c|c} d_3^2 & d_3 e_3 \\ \hline d_3 e_3 & e_3^2 + d_4^2 \end{array} \right]$

- $A^R = d_3^2$
- $C^R = e_3^2 + d_4^2$
- $q^R = \text{abs}(d_3 e_3)$

(for a nontrivial rotation considered below, $q^R > 0$)

- $\Delta^R = \text{abs}(A^R - C^R)$

$\left(\text{if } \Delta^R = 0, \text{abs}(\theta^R) = \frac{\pi}{4} \text{ and feasibility is decided directly from } \frac{L}{\sqrt{2}} \leq \tau \right)$

Let θ^R be the angle of the smaller right rotation, so

- $s^R = \sin \theta^R$
- $c^R = \cos \theta^R$

the rotation angle is computed as

$$\text{abs}(\tan(2\theta^R)) = \frac{2 q^R}{\Delta^R}$$

From

$$\tan(2\theta^R) = \frac{2 \tan\theta^R}{1 - (\tan\theta^R)^2}$$

we obtain

$$\begin{aligned} \frac{\Delta^R}{2 q^R} &= \frac{1 - (\tan\theta^R)^2}{2 \text{abs}(\tan \theta^R)} \\ &= \frac{1}{2 \text{abs}(\tan\theta^R)} - \frac{\text{abs}(\tan\theta^R)}{2} \end{aligned}$$

$$\text{For the shorter rotation, } \text{abs}(\theta^R) \leq \frac{\pi}{4}$$

Therefore $\text{abs}(\tan \theta^R)$ increases with $\text{abs}(\theta^R)$, while the Jacobi ratio above decreases.

For feasibility testing only, define

$$\bullet s = \text{abs}(s^R)$$

Since s also increases with $\text{abs}(\theta^R)$, the Jacobi ratio decreases as s increases.

Work backward from the limit of feasible rotation:

$$s = \frac{\tau}{L}$$

Since $s = \text{abs}(\sin \theta^R)$,

$$\text{abs}(\tan \theta^R) = \frac{s}{\sqrt{1 - s^2}}, \text{ and therefore}$$

$$\frac{\Delta^R}{2 q^R} = \frac{1 - 2s^2}{2s \sqrt{1 - s^2}}$$

Substituting $s = \frac{\tau}{L}$ into the above equation gives

$$\Delta^R \tau \sqrt{L^2 - \tau^2} = q^R (L^2 - 2 \tau^2).$$

Because the Jacobi ratio decreases as s increases, the right rotation is feasible when

$$\Delta^R \tau \sqrt{L^2 - \tau^2} \geq q^R (L^2 - 2 \tau^2)$$



Rejection conditions

The right fill is too large for deflation when

$$\Delta^R \tau \sqrt{L^2 - \tau^2} < q^R (L^2 - 2 \tau^2)$$

From this exact rejection condition, we derive two early conditions,
both for $L \geq 2\tau$

①

- $\sqrt{L^2 - \tau^2} < L$

- If $L \geq 2\tau$, $L^2 - 2\tau^2 \geq L(L - \tau)$

Substitute the above two bounds into

$$\Delta^R \tau \sqrt{L^2 - \tau^2} < q^R (L^2 - 2\tau^2)$$

results in

$$\Delta^R \tau L < q^R L (L - \tau)$$

and, since $L > 0$,

$$\Delta^R \tau < q^R (L - \tau).$$

②

A simpler rejection condition follows from the first.

If $L \geq 2\tau$, then $L - \tau \geq \tau$.

Therefore, if $\Delta^R \leq q^R$,

$$\Delta^R \tau \leq q^R \tau \leq q^R (L - \tau).$$

Thus we can also reject when

$$L \geq 2\tau \text{ and } \Delta^R \leq q^R.$$

The simpler condition is derived from the stronger one.

Although derived second, it is tested first because it requires only one addition.

The stronger condition requires a subtraction and two multiplications.

It is evaluated only when $\Delta^R > q^R$.

The stronger condition

$$\Delta^R \tau < q^R (L - \tau)$$

needs no separate $L \geq 2\tau$ comparison.

It is evaluated only when $\Delta^R > q^R$, so

$$q^R \tau \leq \Delta^R \tau < q^R (L - \tau).$$

Dividing by q^R gives $\tau < L - \tau$, hence $L > 2\tau$

($q^R > 0$ was established previously)



Evaluation order

Candidate cuts are examined from the destination of the QR chase backward.

This prioritizes the region where deflations are more likely to occur.

The first feasible cut is accepted. DBDSQR then resumes block processing, checking for further deflations before performing another QR sweep.

For each candidate cut, Rejection CAT proceeds as follows:

- Evaluate the rotation associated with the larger outer coupling.
 - Apply the inexpensive rejection condition first.
- Apply the stronger rejection condition only if the first test fails to reject.
- Pass candidates surviving both tests directly to LAPACK DLASV2.

The two rejection conditions, in order, are:

①

Reject if

$$L \geq 2\tau \text{ and } \Delta^R \leq q^R$$

②

Reject if

$$\Delta^R \tau < q^R (L - \tau)$$

③

For surviving candidates, compute both rotations

using LAPACK DLASV2 (LAPACK [6])

④

Verify the two cross-boundary fill-ins against τ .

Reject the candidate immediately if either fill-in exceeds τ .

⑤

If both fill-ins satisfy τ , apply both rotations,
update the singular vectors, and split the active block.



Rejection: pseudocode

Input: active upper-bidiagonal block $B[l \dots r]$, tolerance τ_k at each cut k , and QR chase direction idir .

After conventional deflation, examine the surviving couplings in the direction opposite to the QR chase:

- If $\text{idir} = 1$ (left-to-right chase), scan $k = r - 1, \dots, l$.
- If $\text{idir} = 2$ (right-to-left chase), scan $k = l, \dots, r - 1$.

At each candidate cut, $b = e_k$. Apply the following tests in scan order.

1. Obtain the local quantities

$$d_1 = B[k, k]$$

$$d_2 = B[k+1, k+1]$$

$$b = B[k, k+1]$$

$$L = |e_{k-1}| \text{ if } k > l, \text{ otherwise } 0$$

$$R = |e_{k+1}| \text{ if } k + 1 < r, \text{ otherwise } 0$$

$$\tau = \tau_k, M = \max(L, R)$$

2. If $M \leq \tau$, go to step 6.

3. Choose the less permissive rotation and construct its Gram matrix quantities.

If $L \geq R$, use $K^T K$:

$$\begin{aligned}
A &= d_1^2 \\
C &= b^2 + d_2^2 \\
q &= d_1 b \\
\text{Otherwise, use } K K^T: \\
A &= d_1^2 + b^2 \\
C &= d_2^2 \\
q &= b d_2 \\
\Delta &= A - C
\end{aligned}$$

4. First rejection screen:

If $|\Delta| \leq |q|$:

If $M \geq 2\tau$, reject this cut and continue to the next k .

Otherwise, proceed directly to step 6.

5. Second rejection screen:

If $|\Delta| > |q|$, compute:

$$M - \tau$$

$$|\Delta| \tau$$

$$|q| (M - \tau)$$

If $|\Delta| \tau < |q| (M - \tau)$, reject this cut and continue to the next k

6. Compute the local 2×2 SVD using LAPACK DLASV2 (LAPACK [6])

Obtain right rotation (c^R, s^R), left rotation (c^L, s^L), and signed diagonal values

Use the same ordering convention as the production algorithm.

7. Verify the cross-boundary fills.

If $k > l$ and $|s^R e_{k-1}| > \tau$, reject this cut and continue

If $k + 1 < r$ and $|s^L e_{k+1}| > \tau$, reject this cut and continue

8. If the cut is accepted

Scale existing outer couplings by c^R and c^L where present

Replace the two local diagonal entries with the returned signed values

$$\text{Set } e_k = 0$$

Accumulate the right rotation in V^T and the left rotation in U
 Return to bidiagonal QR to process the resulting blocks
 for conventional deflation
 if no conventional deflation occurs, active deflation
 if no active deflation occurs, new bulge chase

9. If every candidate is rejected, perform the ordinary bidiagonal QR sweep.



Why do the two rejection screens miss so little?

The first two tests reject a candidate before the exact fill condition is evaluated.

In the graph below, the green curve shows the exact rejection boundary, while the blue curve shows the inexpensive rejection boundary from the second screen.

For the less permissive rotation, define

- $M = \max(L, R)$

- $r = \frac{\Delta}{q}$

- $x = \frac{M}{\tau}$

Recall that $M \geq 2\tau$, so $x \geq 2$ is part of our rejection setting.

Substituting r and $M = \tau x$ into the exact rejection condition

$$\Delta \tau \sqrt{M^2 - \tau^2} < q (M^2 - 2\tau^2)$$

results in

$$r \sqrt{x^2 - 1} < x^2 - 2,$$

or, equivalently,

$$r < \frac{x^2 - 2}{\sqrt{x^2 - 1}}$$

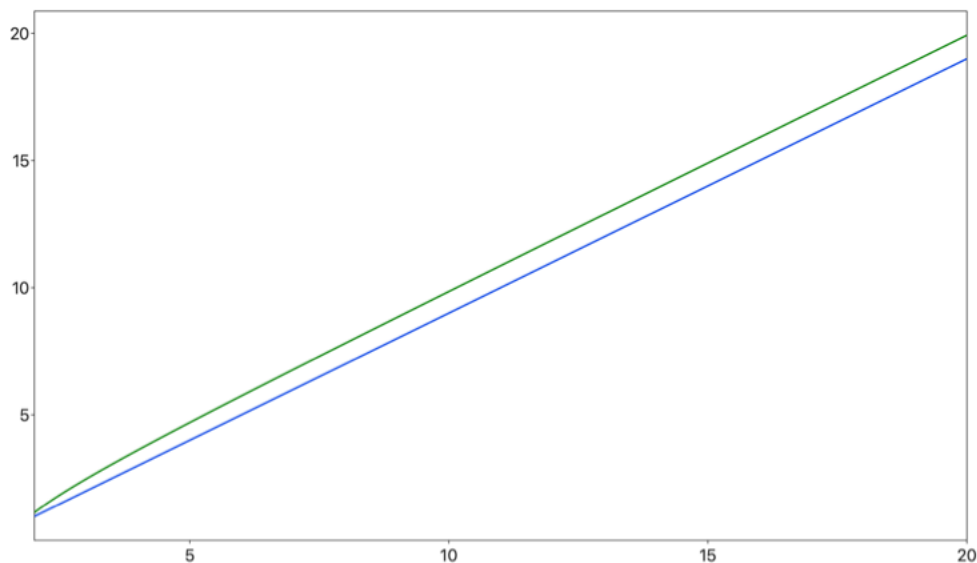
Substituting r and $M = \tau x$ into the second rejection screen

$$\Delta \tau < q (M - \tau)$$

results in

$$r < x - 1.$$

Thus, for $x \geq 2$, the blue boundary lies below the green boundary. The region between them shows the small gap that corresponds to possible candidates rejected by the exact test but not by the inexpensive screen.



To examine the gap for large x , rewrite the exact rejection boundary as

$$\frac{x^2 - 2}{\sqrt{x^2 - 1}} = x \left(1 - \frac{2}{x^2} \right) \left(1 - \frac{1}{x^2} \right)^{-0.5}$$

Define

$$u = \frac{1}{x^2}$$

$$v = 1 - u$$

$$f(u) = \left(1 - \frac{1}{x^2} \right)^{-0.5} = (1 - u)^{-0.5} = v^{-0.5}$$

$$\text{Since } \frac{dv}{du} = -1$$

$$f'(u) = \left(\frac{1}{2} \right) v^{-1.5}$$

$$f''(u) = \left(\frac{3}{4} \right) v^{-2.5}$$

At $u = 0$, $v = 1$, so

$$f(0) = 1, f'(0) = \frac{1}{2}, f''(0) = \frac{3}{4}$$

The Taylor expansion around $u = 0$ is

$$f(u) = f(0) + f'(0)u + \frac{f''(0)u^2}{2!} + \dots$$

$$= 1 + \frac{u}{2} + \frac{3u^2}{8} + \frac{5u^3}{16} + \dots$$

Multiplying by $(1 - 2u)$, substituting $u = \frac{1}{x^2}$, and multiplying by x gives

$$\frac{x^2 - 2}{\sqrt{x^2 - 1}} = x - \frac{3}{2x} - \frac{5}{8x^3} - \dots$$

Subtracting the inexpensive rejection boundary $x - 1$ gives

$$G(x) = 1 - \frac{3}{2x} - \frac{5}{8x^3} - \dots$$

The Taylor series converges here because $u = \frac{1}{x^2} < 1$.

Thus, the absolute gap approaches 1 from below, while the gap relative to x approaches 0.

Its leading relative term is $\frac{1}{x} = \frac{\tau}{M}.$



Singular-vector accumulation

Accumulation of the Householder reflectors from phase 1 is excluded from this comparison.

The phase-2 singular-vector matrices $\mathbf{U}$ and $\mathbf{V}^T$ are initialized to identity $\mathbf{I}_n$.

During phase 2 one left and one right Givens rotation is applied by:

- Each QR chase position
- Terminal 2×2 diagonalization
 - An accepted CAT split

Conventional deflation requires no additional singular-vector rotation.

For an active bidiagonal block of order h , the affected panels are:

- $n \times h$ columns of U
- $h \times n$ rows of V^T

Example: $n = 9$, $h = 5$, active indices 3...7.

Filled dots identify the active panel; blue marks the two vectors updated by one rotation.

U: two columns updated

[illegible]

V^T : two rows updated

○	○	○	○	○	○	○	○	○
○	○	○	○	○	○	○	○	○
●	●	●	●	●	●	●	●	●
●	●	●	●	●	●	●	●	●
●	●	●	●	●	●	●	●	●
●	●	●	●	●	●	●	●	●
○	○	○	○	○	○	○	○	○
○	○	○	○	○	○	○	○	○

Each rotation updates two vectors of length n .
For each entry pair, four multiplications and two additions or subtractions are required.

- One singular-vector rotation: $6n$
- One paired left and right update: $2(6n) = 12n$
- One QR sweep on an active block of order h : $12n(h - 1)$
 - One accepted CAT split: $12n$
 - One terminal 2×2 diagonalization: $12n$

These updates are performed by the modified LAPACK DBDSQR production function.



Spectral-norm bound for accepted splits

The same LAPACK DLASV2 rotations diagonalize a terminal 2×2 block or an interior one.

An interior rotation differs only in the two cross-boundary fill-ins that CAT discards.

Let $B' = L B R$ be the matrix after the rotations and $\tilde{B}$ the matrix after trimming. Then $B' = \tilde{B} + E$.

The discarded perturbation in the four affected coordinates is:

$$E = \begin{bmatrix} 0 & 0 & -s^R e_2 & 0 \\ 0 & 0 & 0 & s^L e_4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$E^T E = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -s^R e_2 & 0 & 0 & 0 \\ 0 & s^L e_4 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} 0 & 0 & -s^R e_2 & 0 \\ 0 & 0 & 0 & s^L e_4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & (s^R e_2)^2 & 0 \\ 0 & 0 & 0 & (s^L e_4)^2 \end{bmatrix}$$

The two fill-ins occupy different rows and columns, so $E^T E$ is diagonal.

$$\lambda_{\max}(E^T E) = \max \left((s^R e_2)^2, (s^L e_4)^2 \right).$$

$$\|E\|_2 = \sqrt{\lambda_{\max}(E^T E)} = \max(|s^R e_2|, |s^L e_4|).$$

CAT accepts only if $|s^R e_2| \leq \tau$ and $|s^L e_4| \leq \tau$.

Therefore $\|E\|_2 \leq \tau$.

Conventional deflation discards only e_3 : $E_p = \begin{bmatrix} 0 & e_3 \\ 0 & 0 \end{bmatrix}$

$$E_p^T E_p = \begin{bmatrix} 0 & 0 \\ 0 & e_3^2 \end{bmatrix}$$

$$\|E_p\|_2 = \sqrt{\lambda_{\max}(E_p^T E_p)} = |e_3| \leq \tau.$$

Thus conventional and active splits enforce the same spectral-norm bound on each discarded perturbation.



Computational methods

Both methods begin with Householder bidiagonalization of the original rectangular matrix A .

The identical upper-bidiagonal matrix B is supplied to both phase-2 algorithms.

Both use the same instrumented production LAPACK DBDSQR implementation for:

- Shift selection and bidiagonal QR bulge chasing
 - Conventional deflation
- Terminal 2×2 diagonalization using LAPACK DLASV2
 - Singular-vector accumulation

CAT adds only:

- Screening for additional deflation opportunities between QR sweeps
- Applying the same LAPACK DLASV2 local diagonalization to selected nonterminal 2×2 blocks
- Verifying and discarding negligible fill-ins, updating singular vectors, and splitting the active block

Computational effort is evaluated for phase 2 only.

The common Householder bidiagonalization is excluded from the comparison.

The weighted-work model uses structure-aware arithmetic counts for standard LAPACK bidiagonal QR iteration.

The local 2×2 diagonalization and singular-vector rotation costs are modeled from LAPACK DLASV2 and BLAS DROT.

Shift computation is charged using the branch-dependent work recorded by the instrumented DBDSQR routine.

We report QR sweep positions and weighted arithmetic work, not execution time.



Matrix families used for testing

Matrix family	Construction for the experiment Matrix dimensions: $2n \times n$	Adjacent σ relation if applicable	Source if applicable
Power law, $\alpha = 0.5$ Specified singular values with power-law decay	• Set $\sigma_i = i^{-0.5}$, $i = 1, \dots, n$ • Construct a $2n \times n$ diagonal matrix Σ containing the specified singular values • Construct U and V as products of 4 $\max(m, n)$ plane rotations, with angles in $[0.2, 1.35]$ radians • Construct $A = U \Sigma V^T$	$\sigma_{k+1} = \sigma_k \left(\frac{k}{k+1} \right)^{0.5}$	(LAPACK [8])
$\alpha = 1.0$	• Set $\sigma_i = i^{-1.0}$, $i = 1, \dots, n$ • Otherwise as above	$\sigma_{k+1} = \sigma_k \left(\frac{k}{k+1} \right)^{1.0}$	(LAPACK [8])
$\alpha = 2.0$	• Set $\sigma_i = i^{-2.0}$, $i = 1, \dots, n$ • Otherwise as above	$\sigma_{k+1} = \sigma_k \left(\frac{k}{k+1} \right)^{2.0}$	(LAPACK [8])
$\alpha = 3.0$	• Set $\sigma_i = i^{-3.0}$, $i = 1, \dots, n$ • Otherwise as above	$\sigma_{k+1} = \sigma_k \left(\frac{k}{k+1} \right)^{3.0}$	(LAPACK [8])
IID noise Independent uniformly distributed entries	• Sample each entry independently from $[-1.0, 1.0]$		
Uniform $[1.0, 10.0]$ Uniformly sampled singular values	• Sample singular values from $[1.0, 10.0]$ • Construct a $2n \times n$ diagonal matrix Σ containing the specified singular values • Construct U and V as products of 4 $\max(m, n)$ plane rotations, with angles in $[0.2, 1.35]$ radians • Construct $A = U \Sigma V^T$		(LAPACK [8])
$[1.0, 100.0]$	• Sample singular values from $[1.0, 100.0]$ • Otherwise as above		(LAPACK [8])
$[1.0, 1000.0]$	• Sample singular values from $[1.0, 1000.0]$ • Otherwise as above		(LAPACK [8])
	• Sample singular values from $[1.0, 10000.0]$		

[1.0, 10000.0]	• Otherwise as above		(LAPACK [8])
Linear Linearly spaced singular values	• Set $\sigma_i = 0.001 + (i-1)h$, $i = 1, \dots, n$ • $h = \frac{1000.0 - 0.001}{n-1}$ • Construct a $2n \times n$ diagonal matrix Σ containing the specified singular values • Construct U and V as products of $4 \max(m, n)$ plane rotations, with angles in [0.2, 1.35] radians • Construct $A = U \Sigma V^T$	$\sigma_{k+1} = \sigma_k + h$	(LAPACK [8])
Geometric Geometrically spaced singular values	• Set $\sigma_i = 0.001 \times r^{(i-1)}$, $i = 1, \dots, n$ • $r = \left(10^6\right)^{\frac{1}{n-1}}$ • Construct a $2n \times n$ diagonal matrix Σ containing the specified singular values • Construct U and V as products of $4 \max(m, n)$ plane rotations, with angles in [0.2, 1.35] radians • Construct $A = U \Sigma V^T$	$\sigma_{k+1} = r \sigma_k$	(LAPACK [8])
Tightly clustered [1.0, 1.001] Tightly clustered singular values	• Sample singular values from [1.0, 1.001] • Construct a $2n \times n$ diagonal matrix Σ containing the specified singular values • Construct U and V as products of $4 \max(m, n)$ plane rotations, with angles in [0.2, 1.35] radians • Construct $A = U \Sigma V^T$		(LAPACK [8])

For each matrix family, 3 deterministic matrices are tested at $n = 64, 128, 256, 512$.
The random seeds are 1, 2, 3.



Reported quantities

①

Computational effort

Quantity	Formula	Reported
Sweep positions	$\sum (h - 1)$, over all active blocks and QR sweeps	Average across samples
Weighted arithmetic work W		Average across samples
CAT test work		Average across samples
Accepted CAT work		Average across samples

②

Conventional and active deflations

Conventional deflations are ordinary bidiagonal QR splits.
 Active CAT deflations are the splits accepted by the production library.
 Terminal 2×2 diagonalizations are excluded from both counts.

$$\text{Active CAT fraction} = \frac{\text{active CAT}}{(\text{active CAT} + \text{conventional})}.$$

For a parent block of size m and children of sizes n_l and n_r ,

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

$$\text{The ratio lies between } \frac{1}{m} \text{ and } \frac{1}{2}$$

We report the mean smaller-child ratio separately for

- active CAT splits
- conventional splits in CAT
- conventional splits in standard QR.

Counts are summed across samples; mean balances are pooled across individual splits.

③

Accuracy

Quantity	Formula	Reported
Max absolute σ error	$\max_i \sigma_i - \sigma_i^* $	Maximum across samples
Max scaled σ error (floor = 1)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{\max(1, \sigma_i^*)} \right)$	Maximum across samples
Max relative σ error (for $\sigma_i^* \neq 0$)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{ \sigma_i^* } \right)$	Maximum across samples
Left singular-vector orthogonality	$\ U^T U - I\ _F$	Maximum across samples
Right singular-vector orthogonality	$\ V^T V - I\ _F$	Maximum across samples
Relative SVD residual	$\frac{\ B - U \Sigma V^T\ _F}{\ B\ _F}$	Maximum across samples

The reference singular values σ_i^* are computed by LAPACK DGESVD from the original rectangular matrix A.

④

CAT rejection screens

Quantity	Reported
Total CAT candidate inspections	Sum across samples
First-test rejections	Sum across samples
Second-test rejections	Sum across samples

⑤

Production verification

Quantity	Reported
Missed CAT opportunities	Sum across samples
Sent to production	Sum across samples
Rejected by production	Sum across samples



Weighted-work accounting

①

Relative weights of arithmetic operations

Operation	Weight
Floating-point addition	1
Floating-point subtraction	1
Floating-point multiplication	1
Floating-point division	4
Square root	4
hypot(x, y)	8

②

Arithmetic models

Quantity	Operation model	Weighted work	Source
Singular-vector rotation	One rotation applied to U or V	$6n$	LAPACK [7]
Conventional convergence scan	Initial test + up to $(h - 1)$ coupling tests (stops at first deflation)	1 + $7(h - 1)$ (full scan only)	LAPACK [2]
Shift computation	Shift selection + optional DLAS2 computation + optional negligibility test (cost recorded by production)	7–53 (branch-dependent)	LAPACK [2] LAPACK [3]
Zero-shift bidiagonal QR sweep, bidiagonal updates	Two Givens + simplified updates per chase position + endpoint correction + shift selection & computation	$36(h - 1)$ + 2 + shift work	LAPACK [2]
Shifted bidiagonal QR sweep, bidiagonal updates	Two Givens + full updates per chase position + net boundary correction $(7 - 2)$ + shift selection & computation	$48(h - 1)$ + 5 + shift work	LAPACK [2]
Singular-vector updates during QR sweep	One U rotation + one V rotation per chase position	$2(6)n(h - 1)$	LAPACK [5]
CAT rejection, test 1	Three squares + one diagonal sum + one off-diagonal product + one diagonal difference + tolerance doubling	3 + 1 + 1 + 1 + 1 = 7	This work
CAT rejection, test 2	Three squares + one diagonal sum + one off-diagonal product + one diagonal difference + margin subtraction + two products	3 + 1 + 1 + 1 + 1 + 2 = 9	This work
Production local 2×2 SVD	DLASV2 computes candidate rotations (no singular-vector updates)	76	LAPACK [6]
Terminal 2×2 diagonalization	DLASV2 computation + rotations applied to U and V	76 + $2(6)n$	LAPACK [6] LAPACK [7]
Accepted CAT, endpoint	DLASV2 computation + one outer-fill verification + one surviving-coupling update + rotations applied to U and V	76 + 1 + 1 + $2(6)n$	LAPACK [6] LAPACK [7]
Accepted CAT, interior	DLASV2 computation + two outer-fill verifications + one surviving-coupling update	76 + 2 + 1	LAPACK [6] LAPACK [7]

+ two surviving-coupling updates + rotations applied to U and V	+ 2 + $2(6)n$	LAPACK [7]
--	--------------------	------------

Here n is the full bidiagonal order and h is the current active-block size.
The shift cost is supplied by the instrumented DBDSQR algorithm for each QR sweep.

CAT work is charged as follows:

- Rejection-screen tests
- Local DLASV2 computation
 - Fill-in verification
- Surviving-coupling updates
- Singular-vector rotations

Comparisons, assignments, sign changes, and diagnostic-only computations are excluded.

The cost of bidiagonalization is also excluded.



Power-law singular values $\alpha = 0.5$

Each matrix size is tested with 3 independently generated samples.

Sweep positions
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	4,242.0	3,894.3	8.20%	347.7
128	17,054.3	15,312.7	10.21%	1,741.7
256	65,767.7	60,153.7	8.54%	5,614.0
512	262,386.0	243,579.3	7.17%	18,806.7

Weighted arithmetic work W
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	3,497,748.3	3,267,018.7	6.60%	230,729.7
128	27,146,569.7	24,592,512.3	9.41%	2,554,057.3
256	205,683,217.0	188,989,868.7	8.12%	16,693,348.3
512	1,626,621,179.0	1,513,387,256.3	6.96%	113,233,922.7

Conventional and active deflations
Counts summed across samples
Mean balances pooled across all splits

n	Passive standard	Passive CAT	Active CAT	Active fraction	Active CAT	Passive CAT	Passive standard
					<u>smaller child</u> parent	<u>smaller child</u> parent	<u>smaller child</u> parent
64	181	126	58	31.52%	0.0552	0.0515	0.0475
128	371	259	116	30.93%	0.0293	0.0333	0.0289
256	756	515	241	31.88%	0.0200	0.0223	0.0178
512	1494	1041	462	30.74%	0.0178	0.0180	0.0114

CAT-specific work
Average across samples
Percentages of total CAT algorithm work

n	CAT test work	% of CAT work	Accepted CAT work	% of CAT work
64	34,404.3	1.05%	16,356.0	0.50%
128	134,031.7	0.55%	62,409.3	0.25%
256	520,377.0	0.28%	253,054.0	0.13%
512	2,052,159.7	0.14%	958,206.0	0.06%

CAT rejection distribution
Counts summed across samples
Percentages of total candidate inspections

n	Total candidates	Test 1 rejections	% of candidates	Test 2 rejections	% of candidates
64	12307	3775	30.67%	8474	68.86%
128	48695	18080	37.13%	30499	62.63%
256	191569	81495	42.54%	109833	57.33%
512	766971	373130	48.65%	393379	51.29%

CAT screening verification
Counts summed across samples

n	Missed opportunities	Sent to production	Rejected by production
64	0	58	0
128	0	116	0
256	0	241	0
512	0	462	0

Numerical accuracy and singular-vector quality
Maximum across all matrix sizes and samples

Quantity	Formula	baseline	candidate
Max absolute σ error	$\max_i \sigma_i - \sigma_i^* $	3.109e-15	2.054e-15
Max scaled σ error (floor = 1)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{\max(1, \sigma_i^*)} \right)$	3.109e-15	2.054e-15
Max relative σ error (for $\sigma_i^* \neq 0$)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{ \sigma_i^* } \right)$	1.243e-14	1.366e-14
Left singular-vector orthogonality	$\ U^T U - I\ _F$	1.134e-13	1.120e-13
Right singular-vector orthogonality	$\ V^T V - I\ _F$	1.080e-13	1.051e-13
Relative SVD residual	$\frac{\ B - U \Sigma V^T\ _F}{\ B\ _F}$	8.862e-15	8.636e-15



Power-law singular values $\alpha = 1.0$

Each matrix size is tested with 3 independently generated samples.

Sweep positions
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	3,743.7	3,354.3	10.40%	389.3
128	15,099.3	13,320.0	11.78%	1,779.3
256	58,113.3	52,513.7	9.64%	5,599.7
512	230,607.7	208,883.0	9.42%	21,724.7

Weighted arithmetic work W
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	3,087,033.3	2,820,982.3	8.62%	266,051.0
128	24,036,566.7	21,425,639.0	10.86%	2,610,927.7
256	181,757,276.0	165,115,423.3	9.16%	16,641,852.7
512	1,429,653,862.3	1,298,244,499.3	9.19%	131,409,363.0

Conventional and active deflations
Counts summed across samples
Mean balances pooled across all splits

n	Passive standard	Passive CAT	Active CAT	Active fraction	Active CAT	Passive CAT	Passive standard
					<u>smaller child</u>	<u>smaller child</u>	<u>smaller child</u>
					parent	parent	parent
64	183	124	62	33.33%	0.0410	0.0580	0.0492
128	372	242	133	35.47%	0.0307	0.0342	0.0296
256	740	475	275	36.67%	0.0191	0.0267	0.0178
512	1473	978	510	34.27%	0.0235	0.0230	0.0134

CAT-specific work
Average across samples
Percentages of total CAT algorithm work

n	CAT test work	% of CAT work	Accepted CAT work	% of CAT work
64	31,542.7	1.12%	17,484.0	0.62%
128	124,255.3	0.58%	71,556.0	0.33%
256	479,908.7	0.29%	288,758.0	0.17%
512	1,833,199.3	0.14%	1,057,767.3	0.08%

CAT rejection distribution
Counts summed across samples
Percentages of total candidate inspections

n	Total candidates	Test 1 rejections	% of candidates	Test 2 rejections	% of candidates
64	11150	2861	25.66%	8227	73.78%
128	44454	13660	30.73%	30661	68.97%
256	174604	65855	37.72%	108474	62.13%
512	679882	309670	45.55%	369702	54.38%

CAT screening verification
Counts summed across samples

n	Missed opportunities	Sent to production	Rejected by production
64	0	62	0
128	0	133	0
256	0	275	0
512	0	510	0

Numerical accuracy and singular-vector quality
Maximum across all matrix sizes and samples

Quantity	Formula	baseline	candidate
Max absolute σ error	$\max_i \sigma_i - \sigma_i^* $	1.998e-15	1.998e-15
Max scaled σ error (floor = 1)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{\max(1, \sigma_i^*)} \right)$	1.998e-15	1.998e-15
Max relative σ error (for $\sigma_i^* \neq 0$)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{ \sigma_i^* } \right)$	1.343e-14	1.498e-14
Left singular-vector orthogonality	$\ U^T U - I\ _F$	1.132e-13	1.063e-13
Right singular-vector orthogonality	$\ V^T V - I\ _F$	1.027e-13	9.887e-14
Relative SVD residual	$\frac{\ B - U \Sigma V^T\ _F}{\ B\ _F}$	9.526e-15	6.275e-15



Power-law singular values $\alpha = 2.0$

Each matrix size is tested with 3 independently generated samples.

Sweep positions
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	3,184.7	2,703.7	15.10%	481.0
128	13,283.0	11,100.3	16.43%	2,182.7
256	50,815.7	42,427.3	16.51%	8,388.3
512	198,626.3	167,002.0	15.92%	31,624.3

Weighted arithmetic work W
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	2,626,224.7	2,284,435.0	13.01%	341,789.7
128	21,129,398.7	17,888,285.0	15.34%	3,241,113.7
256	158,914,738.7	133,547,963.7	15.96%	25,366,775.0
512	1,231,347,969.7	1,038,419,998.0	15.67%	192,927,971.7

Conventional and active deflations
Counts summed across samples
Mean balances pooled across all splits

n	Passive standard	Passive CAT	Active CAT	Active fraction	Active CAT	Passive CAT	Passive standard
					<u>smaller child</u> parent	<u>smaller child</u> parent	<u>smaller child</u> parent
64	184	117	67	36.41%	0.0462	0.0621	0.0504
128	364	217	155	41.67%	0.0360	0.0409	0.0312
256	717	433	306	41.41%	0.0326	0.0342	0.0210
512	1445	925	540	36.86%	0.0393	0.0329	0.0201

CAT-specific work
Average across samples
Percentages of total CAT algorithm work

n	CAT test work	% of CAT work	Accepted CAT work	% of CAT work
64	28,212.7	1.23%	18,895.3	0.83%
128	113,734.7	0.64%	83,394.0	0.47%
256	419,427.3	0.31%	321,315.3	0.24%
512	1,567,808.7	0.15%	1,120,003.3	0.11%

CAT rejection distribution
Counts summed across samples
Percentages of total candidate inspections

n	Total candidates	Test 1 rejections	% of candidates	Test 2 rejections	% of candidates
64	9834	1934	19.67%	7833	79.65%
128	40410	11243	27.82%	29012	71.79%
256	151557	52904	34.91%	98346	64.89%
512	576118	240818	41.80%	334760	58.11%

CAT screening verification
Counts summed across samples

n	Missed opportunities	Sent to production	Rejected by production
64	0	67	0
128	0	155	0
256	0	307	1
512	0	540	0

Numerical accuracy and singular-vector quality
Maximum across all matrix sizes and samples

Quantity	Formula	baseline	candidate
Max absolute σ error	$\max_i \sigma_i - \sigma_i^* $	1.554e-15	1.554e-15
Max scaled σ error (floor = 1)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{\max(1, \sigma_i^*)} \right)$	1.554e-15	1.554e-15
Max relative σ error (for $\sigma_i^* \neq 0$)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{ \sigma_i^* } \right)$	4.373e-12	4.373e-12
Left singular-vector orthogonality	$\ U^T U - I\ _F$	9.763e-14	9.131e-14
Right singular-vector orthogonality	$\ V^T V - I\ _F$	9.180e-14	8.885e-14
Relative SVD residual	$\frac{\ B - U \Sigma V^T\ _F}{\ B\ _F}$	7.853e-15	1.028e-14



Power-law singular values $\alpha = 3.0$

Each matrix size is tested with 3 independently generated samples.

Sweep positions
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	3,391.0	2,562.3	24.44%	828.7
128	12,757.0	9,730.7	23.72%	3,026.3
256	49,181.0	37,685.0	23.37%	11,496.0
512	193,896.7	146,793.3	24.29%	47,103.3

Weighted arithmetic work W
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	2,782,416.7	2,165,969.3	22.16%	616,447.3
128	20,270,308.0	15,700,605.7	22.54%	4,569,702.3
256	153,712,579.0	118,629,318.7	22.82%	35,083,260.3
512	1,201,747,800.3	913,034,408.0	24.02%	288,713,392.3

Conventional and active deflations
Counts summed across samples
Mean balances pooled across all splits

n	Passive standard	Passive CAT	Active CAT	Active fraction	Active CAT	Passive CAT	Passive standard
					<u>smaller child</u>	<u>smaller child</u>	<u>smaller child</u>
					parent	parent	parent
64	179	101	79	43.89%	0.0543	0.0666	0.0517
128	360	200	163	44.90%	0.0447	0.0451	0.0328
256	706	433	292	40.28%	0.0463	0.0439	0.0278
512	1393	845	594	41.28%	0.0500	0.0449	0.0282

CAT-specific work
Average across samples
Percentages of total CAT algorithm work

n	CAT test work	% of CAT work	Accepted CAT work	% of CAT work
64	28,690.7	1.32%	22,282.0	1.03%
128	107,217.7	0.68%	87,704.0	0.56%
256	396,811.3	0.33%	306,624.7	0.26%
512	1,444,296.0	0.16%	1,232,016.7	0.13%

CAT rejection distribution
Counts summed across samples
Percentages of total candidate inspections

n	Total candidates	Test 1 rejections	% of candidates	Test 2 rejections	% of candidates
64	10002	1973	19.73%	7950	79.48%
128	37905	9746	25.71%	27996	73.86%
256	142752	47167	33.04%	95293	66.75%
512	527850	208881	39.57%	318375	60.32%

CAT screening verification
Counts summed across samples

n	Missed opportunities	Sent to production	Rejected by production
64	0	79	0
128	0	163	0
256	0	292	0
512	0	594	0

Numerical accuracy and singular-vector quality
Maximum across all matrix sizes and samples

Quantity	Formula	baseline	candidate
Max absolute σ error	$\max_i \sigma_i - \sigma_i^* $	1.110e-15	1.110e-15
Max scaled σ error (floor = 1)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{\max(1, \sigma_i^*)} \right)$	1.110e-15	1.110e-15
Max relative σ error (for $\sigma_i^* \neq 0$)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{ \sigma_i^* } \right)$	2.749e-09	2.749e-09
Left singular-vector orthogonality	$\ U^T U - I\ _F$	9.414e-14	8.556e-14
Right singular-vector orthogonality	$\ V^T V - I\ _F$	8.909e-14	8.636e-14
Relative SVD residual	$\frac{\ B - U \Sigma V^T\ _F}{\ B\ _F}$	6.011e-15	1.041e-14



IID rectangular noise

Each matrix size is tested with 3 independently generated samples.

Sweep positions
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	4,827.0	4,521.0	6.34%	306.0
128	19,547.0	18,862.3	3.50%	684.7
256	74,429.0	70,713.3	4.99%	3,715.7
512	294,825.0	281,380.0	4.56%	13,445.0

Weighted arithmetic work W
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	3,980,431.7	3,780,048.3	5.03%	200,383.3
128	31,114,895.7	30,225,105.7	2.86%	889,790.0
256	232,774,827.7	221,953,473.0	4.65%	10,821,354.7
512	1,827,696,606.7	1,747,477,652.3	4.39%	80,218,954.3

Conventional and active deflations
Counts summed across samples
Mean balances pooled across all splits

n	Passive standard	Passive CAT	Active CAT	Active fraction	Active CAT	Passive CAT	Passive standard
					<u>smaller child</u> parent	<u>smaller child</u> parent	<u>smaller child</u> parent
64	182	148	36	19.57%	0.0558	0.0517	0.0506
128	378	318	59	15.65%	0.0271	0.0321	0.0312
256	752	597	161	21.24%	0.0176	0.0196	0.0181
512	1518	1207	312	20.54%	0.0111	0.0125	0.0105

CAT-specific work
Average across samples
Percentages of total CAT algorithm work

n	CAT test work	% of CAT work	Accepted CAT work	% of CAT work
64	38,234.7	1.01%	10,152.7	0.27%
128	157,845.3	0.52%	31,742.0	0.11%
256	581,383.7	0.26%	169,051.3	0.08%
512	2,282,261.7	0.13%	647,092.7	0.04%

CAT rejection distribution
Counts summed across samples
Percentages of total candidate inspections

n	Total candidates	Test 1 rejections	% of candidates	Test 2 rejections	% of candidates
64	13600	3848	28.29%	9716	71.44%
128	56672	18256	32.21%	38357	67.68%
256	212363	83558	39.35%	128644	60.58%
512	844927	378779	44.83%	465836	55.13%

CAT screening verification
Counts summed across samples

n	Missed opportunities	Sent to production	Rejected by production
64	0	36	0
128	0	59	0
256	0	161	0
512	0	312	0

Numerical accuracy and singular-vector quality
Maximum across all matrix sizes and samples

Quantity	Formula	baseline	candidate
Max absolute σ error	$\max_i \sigma_i - \sigma_i^* $	2.593e-13	2.842e-13
Max scaled σ error (floor = 1)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{\max(1, \sigma_i^*)} \right)$	1.268e-14	9.451e-15
Max relative σ error (for $\sigma_i^* \neq 0$)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{ \sigma_i^* } \right)$	1.268e-14	9.451e-15
Left singular-vector orthogonality	$\ U^T U - I\ _F$	1.124e-13	1.122e-13
Right singular-vector orthogonality	$\ V^T V - I\ _F$	1.170e-13	1.128e-13
Relative SVD residual	$\frac{\ B - U \Sigma V^T\ _F}{\ B\ _F}$	7.119e-15	7.255e-15



Uniform singular values [1.0, 10.0]

Each matrix size is tested with 3 independently generated samples.

Sweep positions
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	4,663.3	4,418.7	5.25%	244.7
128	17,829.3	16,729.3	6.17%	1,100.0
256	69,393.0	65,970.7	4.93%	3,422.3
512	265,427.3	246,697.7	7.06%	18,729.7

Weighted arithmetic work W
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	3,846,446.0	3,693,420.7	3.98%	153,025.3
128	28,385,409.3	26,826,775.3	5.49%	1,558,634.0
256	217,042,631.7	207,078,435.0	4.59%	9,964,196.7
512	1,645,545,702.0	1,532,305,775.3	6.88%	113,239,926.7

Conventional and active deflations
Counts summed across samples
Mean balances pooled across all splits

n	Passive standard	Passive CAT	Active CAT	Active fraction	Active CAT	Passive CAT	Passive standard
					<u>smaller child</u>	<u>smaller child</u>	<u>smaller child</u>
					parent	parent	parent
64	180	151	30	16.57%	0.0810	0.0587	0.0597
128	367	296	77	20.64%	0.0451	0.0360	0.0362
256	738	596	151	20.21%	0.0260	0.0203	0.0202
512	1472	1152	345	23.05%	0.0220	0.0169	0.0157

CAT-specific work
Average across samples
Percentages of total CAT algorithm work

n	CAT test work	% of CAT work	Accepted CAT work	% of CAT work
64	37,205.7	1.01%	8,462.0	0.23%
128	138,912.0	0.52%	41,428.7	0.15%
256	540,127.0	0.26%	158,557.3	0.08%
512	1,976,443.0	0.13%	715,542.7	0.05%

CAT rejection distribution
Counts summed across samples
Percentages of total candidate inspections

n	Total candidates	Test 1 rejections	% of candidates	Test 2 rejections	% of candidates
64	13297	4028	30.29%	9239	69.48%
128	50284	17910	35.62%	32297	64.23%
256	198089	81210	41.00%	116728	58.93%
512	740532	367759	49.66%	372427	50.29%

CAT screening verification
Counts summed across samples

n	Missed opportunities	Sent to production	Rejected by production
64	0	30	0
128	0	77	0
256	0	151	0
512	0	346	1

Numerical accuracy and singular-vector quality
Maximum across all matrix sizes and samples

Quantity	Formula	baseline	candidate
Max absolute σ error	$\max_i \sigma_i - \sigma_i^* $	6.573e-14	5.684e-14
Max scaled σ error (floor = 1)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{\max(1, \sigma_i^*)} \right)$	6.792e-15	6.273e-15
Max relative σ error (for $\sigma_i^* \neq 0$)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{ \sigma_i^* } \right)$	6.792e-15	6.273e-15
Left singular-vector orthogonality	$\ U^T U - I\ _F$	1.121e-13	1.089e-13
Right singular-vector orthogonality	$\ V^T V - I\ _F$	1.096e-13	1.057e-13
Relative SVD residual	$\frac{\ B - U \Sigma V^T\ _F}{\ B\ _F}$	7.002e-15	6.966e-15



Uniform singular values [1.0, 100.0]

Each matrix size is tested with 3 independently generated samples.

Sweep positions
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	4,258.7	4,069.0	4.45%	189.7
128	16,511.3	15,791.7	4.36%	719.7
256	66,414.3	62,428.7	6.00%	3,985.7
512	251,588.0	239,057.7	4.98%	12,530.3

Weighted arithmetic work W
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	3,512,684.0	3,402,628.3	3.13%	110,055.7
128	26,287,330.3	25,318,835.7	3.68%	968,494.7
256	207,719,522.0	195,957,926.0	5.66%	11,761,596.0
512	1,559,736,737.0	1,484,825,532.0	4.80%	74,911,205.0

Conventional and active deflations
Counts summed across samples
Mean balances pooled across all splits

n	Passive standard	Passive CAT	Active CAT	Active fraction	Active CAT	Passive CAT	Passive standard
					<u>smaller child</u> parent	<u>smaller child</u> parent	<u>smaller child</u> parent
64	182	151	33	17.93%	0.0661	0.0581	0.0584
128	369	304	67	18.06%	0.0509	0.0394	0.0381
256	747	599	147	19.71%	0.0319	0.0239	0.0225
512	1480	1166	332	22.16%	0.0269	0.0217	0.0187

CAT-specific work
Average across samples
Percentages of total CAT algorithm work

n	CAT test work	% of CAT work	Accepted CAT work	% of CAT work
64	34,050.7	1.00%	9,306.7	0.27%
128	129,954.0	0.51%	36,050.0	0.14%
256	503,066.0	0.26%	154,358.7	0.08%
512	1,895,512.3	0.13%	688,584.7	0.05%

CAT rejection distribution
Counts summed across samples
Percentages of total candidate inspections

n	Total candidates	Test 1 rejections	% of candidates	Test 2 rejections	% of candidates
64	12248	4040	32.98%	8175	66.75%
128	47466	18666	39.32%	28733	60.53%
256	187490	89106	47.53%	98237	52.40%
512	717658	386231	53.82%	331094	46.14%

CAT screening verification
Counts summed across samples

n	Missed opportunities	Sent to production	Rejected by production
64	0	33	0
128	0	67	0
256	0	147	0
512	0	333	1

Numerical accuracy and singular-vector quality
Maximum across all matrix sizes and samples

Quantity	Formula	baseline	candidate
Max absolute σ error	$\max_i \sigma_i - \sigma_i^* $	6.537e-13	6.111e-13
Max scaled σ error (floor = 1)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{\max(1, \sigma_i^*)} \right)$	2.298e-14	1.787e-14
Max relative σ error (for $\sigma_i^* \neq 0$)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{ \sigma_i^* } \right)$	2.298e-14	1.787e-14
Left singular-vector orthogonality	$\ U^T U - I\ _F$	1.086e-13	1.061e-13
Right singular-vector orthogonality	$\ V^T V - I\ _F$	1.071e-13	1.031e-13
Relative SVD residual	$\frac{\ B - U \Sigma V^T\ _F}{\ B\ _F}$	6.833e-15	6.967e-15



Uniform singular values [1.0, 1000.0]

Each matrix size is tested with 3 independently generated samples.

Sweep positions
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	4,105.7	3,899.3	5.03%	206.3
128	16,466.0	15,570.3	5.44%	895.7
256	63,274.3	60,069.7	5.06%	3,204.7
512	252,865.0	236,295.0	6.55%	16,570.0

Weighted arithmetic work W
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	3,386,653.7	3,259,967.7	3.74%	126,686.0
128	26,214,594.3	24,961,021.3	4.78%	1,253,573.0
256	197,894,543.3	188,526,443.7	4.73%	9,368,099.7
512	1,567,638,568.0	1,467,609,586.7	6.38%	100,028,981.3

Conventional and active deflations
 Counts summed across samples
 Mean balances pooled across all splits

n	Passive standard	Passive CAT	Active CAT	Active fraction	Active CAT	Passive CAT	Passive standard
					<u>smaller child</u>	<u>smaller child</u>	<u>smaller child</u>
					parent	parent	parent
64	182	154	30	16.30%	0.0712	0.0564	0.0575
128	370	309	64	17.16%	0.0505	0.0350	0.0364
256	752	620	128	17.11%	0.0312	0.0217	0.0213
512	1490	1188	314	20.91%	0.0283	0.0211	0.0189

CAT-specific work
 Average across samples
 Percentages of total CAT algorithm work

n	CAT test work	% of CAT work	Accepted CAT work	% of CAT work
64	32,545.7	1.00%	8,460.7	0.26%
128	127,801.0	0.51%	34,434.7	0.14%
256	476,333.7	0.25%	134,407.3	0.07%
512	1,866,831.0	0.13%	651,250.7	0.04%

CAT rejection distribution
 Counts summed across samples
 Percentages of total candidate inspections

n	Total candidates	Test 1 rejections	% of candidates	Test 2 rejections	% of candidates
64	11735	3989	33.99%	7716	65.75%
128	46793	18867	40.32%	27862	59.54%
256	180377	97196	53.88%	83053	46.04%
512	709362	391921	55.25%	317126	44.71%

CAT screening verification
Counts summed across samples

n	Missed opportunities	Sent to production	Rejected by production
64	0	30	0
128	0	64	0
256	0	128	0
512	0	315	1

Numerical accuracy and singular-vector quality
Maximum across all matrix sizes and samples

Quantity	Formula	baseline	candidate
Max absolute σ error	$\max_i \sigma_i - \sigma_i^* $	5.571e-12	6.139e-12
Max scaled σ error (floor = 1)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{\max(1, \sigma_i^*)} \right)$	1.279e-13	1.352e-13
Max relative σ error (for $\sigma_i^* \neq 0$)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{ \sigma_i^* } \right)$	1.279e-13	1.352e-13
Left singular-vector orthogonality	$\ U^T U - I\ _F$	1.048e-13	1.012e-13
Right singular-vector orthogonality	$\ V^T V - I\ _F$	1.098e-13	1.036e-13
Relative SVD residual	$\frac{\ B - U \Sigma V^T\ _F}{\ B\ _F}$	6.866e-15	6.639e-15



Uniform singular values [1.0, 10000.0]

Each matrix size is tested with 3 independently generated samples.

Sweep positions
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	4,171.3	3,973.7	4.74%	197.7
128	15,932.3	15,305.7	3.93%	626.7
256	61,168.3	58,755.0	3.95%	2,413.3
512	241,092.7	228,322.3	5.30%	12,770.3

Weighted arithmetic work W
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	3,436,976.3	3,319,256.7	3.43%	117,719.7
128	25,363,125.7	24,532,280.3	3.28%	830,845.3
256	191,301,137.3	184,379,599.3	3.62%	6,921,538.0
512	1,494,624,237.0	1,418,002,723.7	5.13%	76,621,513.3

Conventional and active deflations
Counts summed across samples
Mean balances pooled across all splits

n	Passive standard	Passive CAT	Active CAT	Active fraction	Active CAT	Passive CAT	Passive standard
					<u>smaller child</u> parent	<u>smaller child</u> parent	<u>smaller child</u> parent
64	185	153	32	17.30%	0.0677	0.0570	0.0582
128	369	311	61	16.40%	0.0542	0.0349	0.0360
256	753	631	119	15.87%	0.0313	0.0191	0.0192
512	1503	1210	295	19.60%	0.0263	0.0186	0.0166

CAT-specific work
Average across samples
Percentages of total CAT algorithm work

n	CAT test work	% of CAT work	Accepted CAT work	% of CAT work
64	32,960.7	0.99%	9,024.7	0.27%
128	124,531.0	0.51%	32,821.3	0.13%
256	459,664.7	0.25%	124,956.7	0.07%
512	1,754,367.7	0.12%	611,842.7	0.04%

CAT rejection distribution
Counts summed across samples
Percentages of total candidate inspections

n	Total candidates	Test 1 rejections	% of candidates	Test 2 rejections	% of candidates
64	11960	4379	36.61%	7549	63.12%
128	45997	20190	43.89%	25746	55.97%
256	176408	104339	59.15%	71950	40.79%
512	685377	452645	66.04%	232437	33.91%

CAT screening verification
Counts summed across samples

n	Missed opportunities	Sent to production	Rejected by production
64	0	32	0
128	0	61	0
256	0	119	0
512	0	295	0

Numerical accuracy and singular-vector quality
Maximum across all matrix sizes and samples

Quantity	Formula	baseline	candidate
Max absolute σ error	$\max_i \sigma_i - \sigma_i^* $	6.003e-11	6.185e-11
Max scaled σ error (floor = 1)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{\max(1, \sigma_i^*)} \right)$	8.692e-13	8.692e-13
Max relative σ error (for $\sigma_i^* \neq 0$)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{ \sigma_i^* } \right)$	8.692e-13	8.692e-13
Left singular-vector orthogonality	$\ U^T U - I\ _F$	1.078e-13	1.033e-13
Right singular-vector orthogonality	$\ V^T V - I\ _F$	1.050e-13	1.020e-13
Relative SVD residual	$\frac{\ B - U \Sigma V^T\ _F}{\ B\ _F}$	6.809e-15	6.924e-15



Linearly spaced singular values [0.001, 1000.0]

Each matrix size is tested with 3 independently generated samples.

Sweep positions
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	3,969.0	3,778.0	4.81%	191.0
128	15,428.3	14,694.3	4.76%	734.0
256	59,781.7	56,624.7	5.28%	3,157.0
512	230,908.7	217,123.3	5.97%	13,785.3

Weighted arithmetic work W
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	3,271,980.7	3,157,270.0	3.51%	114,710.7
128	24,557,107.3	23,555,530.7	4.08%	1,001,576.7
256	186,956,789.3	177,770,717.0	4.91%	9,186,072.3
512	1,431,436,163.7	1,348,787,317.3	5.77%	82,648,846.3

Conventional and active deflations
Counts summed across samples
Mean balances pooled across all splits

n	Passive standard	Passive CAT	Active CAT	Active fraction	Active CAT	Passive CAT	Passive standard
					<u>smaller child</u>	<u>smaller child</u>	<u>smaller child</u>
					parent	parent	parent
64	179	150	36	19.35%	0.0774	0.0463	0.0525
128	373	293	80	21.45%	0.0402	0.0282	0.0313
256	756	553	203	26.85%	0.0192	0.0179	0.0183
512	1521	1056	470	30.80%	0.0091	0.0111	0.0104

CAT-specific work
Average across samples
Percentages of total CAT algorithm work

n	CAT test work	% of CAT work	Accepted CAT work	% of CAT work
64	30,383.3	0.96%	10,152.0	0.32%
128	114,205.7	0.48%	43,040.0	0.18%
256	426,519.0	0.24%	213,150.0	0.12%
512	1,595,036.0	0.12%	974,780.0	0.07%

CAT rejection distribution
Counts summed across samples
Percentages of total candidate inspections

n	Total candidates	Test 1 rejections	% of candidates	Test 2 rejections	% of candidates
64	11370	5590	49.16%	5744	50.52%
128	44163	27425	62.10%	16658	37.72%
256	170077	125568	73.83%	44306	26.05%
512	651840	540726	82.95%	110644	16.97%

CAT screening verification
Counts summed across samples

n	Missed opportunities	Sent to production	Rejected by production
64	0	36	0
128	0	80	0
256	0	203	0
512	0	470	0

Numerical accuracy and singular-vector quality
Maximum across all matrix sizes and samples

Quantity	Formula	baseline	candidate
Max absolute σ error	$\max_i \sigma_i - \sigma_i^* $	6.253e-12	6.480e-12
Max scaled σ error (floor = 1)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{\max(1, \sigma_i^*)} \right)$	9.862e-14	9.862e-14
Max relative σ error (for $\sigma_i^* \neq 0$)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{ \sigma_i^* } \right)$	9.862e-11	9.862e-11
Left singular-vector orthogonality	$\ U^T U - I\ _F$	9.981e-14	9.823e-14
Right singular-vector orthogonality	$\ V^T V - I\ _F$	1.027e-13	9.866e-14
Relative SVD residual	$\frac{\ B - U \Sigma V^T\ _F}{\ B\ _F}$	6.800e-15	6.905e-15



Geometric singular values [0.001, 1000.0]

Each matrix size is tested with 3 independently generated samples.

Sweep positions
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	4,411.7	2,410.7	45.36%	2,001.0
128	17,299.7	9,773.0	43.51%	7,526.7
256	68,952.3	39,396.3	42.86%	29,556.0
512	269,391.0	158,981.3	40.98%	110,409.7

Weighted arithmetic work W
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	3,594,005.3	2,022,057.7	43.74%	1,571,947.7
128	27,367,216.0	15,667,146.3	42.75%	11,700,069.7
256	214,952,792.3	123,582,998.3	42.51%	91,369,794.0
512	1,667,313,606.3	986,974,196.0	40.80%	680,339,410.3

Conventional and active deflations
Counts summed across samples
Mean balances pooled across all splits

n	Passive standard	Passive CAT	Active CAT	Active fraction	Active CAT	Passive CAT	Passive standard
					<u>smaller child</u> parent	<u>smaller child</u> parent	<u>smaller child</u> parent
64	143	59	94	61.44%	0.2257	0.2151	0.2206
128	301	145	172	54.26%	0.1943	0.1844	0.1580
256	590	331	326	49.62%	0.1816	0.1715	0.1327
512	1246	682	670	49.56%	0.1448	0.1502	0.1017

CAT-specific work
Average across samples
Percentages of total CAT algorithm work

n	CAT test work	% of CAT work	Accepted CAT work	% of CAT work
64	22,956.0	1.14%	26,530.0	1.31%
128	88,672.0	0.57%	92,576.7	0.59%
256	353,115.7	0.29%	342,370.0	0.28%
512	1,400,788.3	0.14%	1,389,698.7	0.14%

CAT rejection distribution
Counts summed across samples
Percentages of total candidate inspections

n	Total candidates	Test 1 rejections	% of candidates	Test 2 rejections	% of candidates
64	7784	594	7.63%	7096	91.16%
128	30222	2991	9.90%	27059	89.53%
256	121223	15830	13.06%	105067	86.67%
512	484993	81286	16.76%	403037	83.10%

CAT screening verification
Counts summed across samples

n	Missed opportunities	Sent to production	Rejected by production
64	0	94	0
128	0	172	0
256	0	326	0
512	0	670	0

Numerical accuracy and singular-vector quality
Maximum across all matrix sizes and samples

Quantity	Formula	baseline	candidate
Max absolute σ error	$\max_i \sigma_i - \sigma_i^* $	5.798e-12	5.230e-12
Max scaled σ error (floor = 1)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{\max(1, \sigma_i^*)} \right)$	5.224e-14	5.225e-14
Max relative σ error (for $\sigma_i^* \neq 0$)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{ \sigma_i^* } \right)$	1.887e-11	1.887e-11
Left singular-vector orthogonality	$\ U^T U - I\ _F$	1.074e-13	8.791e-14
Right singular-vector orthogonality	$\ V^T V - I\ _F$	1.081e-13	8.834e-14
Relative SVD residual	$\frac{\ B - U \Sigma V^T\ _F}{\ B\ _F}$	6.359e-15	7.356e-15



Tightly clustered singular values [1.0, 1.001]

Each matrix size is tested with 3 independently generated samples.

Sweep positions
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	4,654.0	4,425.7	4.91%	228.3
128	18,127.0	16,611.0	8.36%	1,516.0
256	68,350.0	65,058.0	4.82%	3,292.0
512	255,797.0	243,646.3	4.75%	12,150.7

Weighted arithmetic work W
Average across samples

n	baseline	candidate	relative reduction	absolute reduction
64	3,839,245.0	3,701,721.3	3.58%	137,523.7
128	28,864,566.3	26,640,592.7	7.70%	2,223,973.7
256	213,783,115.7	204,221,943.3	4.47%	9,561,172.3
512	1,585,835,391.3	1,513,378,412.0	4.57%	72,456,979.3

Conventional and active deflations
Counts summed across samples
Mean balances pooled across all splits

n	Passive standard	Passive CAT	Active CAT	Active fraction	Active CAT	Passive CAT	Passive standard
					<u>smaller child</u>	<u>smaller child</u>	<u>smaller child</u>
					parent	parent	parent
64	181	144	37	20.44%	0.0500	0.0543	0.0569
128	360	289	84	22.52%	0.0482	0.0351	0.0378
256	738	590	160	21.33%	0.0284	0.0206	0.0205
512	1482	1145	363	24.07%	0.0159	0.0111	0.0130

CAT-specific work
Average across samples
Percentages of total CAT algorithm work

n	CAT test work	% of CAT work	Accepted CAT work	% of CAT work
64	36,872.3	1.00%	10,434.7	0.28%
128	135,115.7	0.51%	45,194.0	0.17%
256	527,072.7	0.26%	168,006.7	0.08%
512	1,953,736.7	0.13%	752,870.7	0.05%

CAT rejection distribution
Counts summed across samples
Percentages of total candidate inspections

n	Total candidates	Test 1 rejections	% of candidates	Test 2 rejections	% of candidates
64	13317	4618	34.68%	8662	65.04%
128	49943	22070	44.19%	27789	55.64%
256	195358	88502	45.30%	106696	54.62%
512	731333	360431	49.28%	370538	50.67%

CAT screening verification
Counts summed across samples

n	Missed opportunities	Sent to production	Rejected by production
64	0	37	0
128	0	84	0
256	0	160	0
512	0	364	1

Numerical accuracy and singular-vector quality
Maximum across all matrix sizes and samples

Quantity	Formula	baseline	candidate
Max absolute σ error	$\max_i \sigma_i - \sigma_i^* $	1.177e-14	1.055e-14
Max scaled σ error (floor = 1)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{\max(1, \sigma_i^*)} \right)$	1.176e-14	1.055e-14
Max relative σ error (for $\sigma_i^* \neq 0$)	$\max_i \left(\frac{ \sigma_i - \sigma_i^* }{ \sigma_i^* } \right)$	1.176e-14	1.055e-14
Left singular-vector orthogonality	$\ U^T U - I\ _F$	1.303e-13	1.261e-13
Right singular-vector orthogonality	$\ V^T V - I\ _F$	1.263e-13	1.227e-13
Relative SVD residual	$\frac{\ B - U \Sigma V^T\ _F}{\ B\ _F}$	6.656e-15	6.743e-15



Spectral separation and CAT deflation

The observed CAT benefit varies substantially across spectral families and is most pronounced in the geometric group.

We examine how singular-value separation enters the local deflation condition.

We consider $|e_2| \geq |e_4|$, the commonly observed ordering, where

the right rotation then gives the more restrictive fill-in condition.

At the selected coupling, the local bidiagonal block is $K = \left[\begin{array}{c|c} d_3 & e_3 \\ \hline 0 & d_4 \end{array} \right]$

The right rotation diagonalizes local $K^T K = \left[\begin{array}{c|c} d_3^2 & d_3 e_3 \\ \hline d_3 e_3 & e_3^2 + d_4^2 \end{array} \right]$

The left rotation diagonalizes local $K K^T = \left[\begin{array}{c|c} d_3^2 + e_3^2 & e_3 d_4 \\ \hline e_3 d_4 & d_4^2 \end{array} \right]$

The corresponding diagonal differences and off-diagonal magnitudes are:

$$\Delta^R = |d_3^2 - d_4^2 - e_3^2|$$

$$q^R = |d_3 e_3|$$

$$\Delta^L = |d_3^2 + e_3^2 - d_4^2|$$

$$q^L = |e_3 d_4|$$

Recall the exact right-rotation feasibility condition:

$$\Delta^R \tau \sqrt{L^2 - \tau^2} \geq q^R (L^2 - 2\tau^2)$$

Substituting Δ^R , q^R , and $L = |e_2|$ gives:

$$|d_3^2 - d_4^2 - e_3^2| \tau \sqrt{e_2^2 - \tau^2} \geq |d_3 e_3| (e_2^2 - 2\tau^2).$$

For $|e_2| \gg \tau$, this condition is approximately:

$$\frac{|d_3^2 - d_4^2 - e_3^2|}{|d_3 e_3|} \gtrsim \frac{|e_2|}{\tau}$$

$$\text{or} \\ \tau |d_3^2 - d_4^2 - e_3^2| \geq |d_3 e_3 e_2|$$

Entries near convergence:

- e_3 gets smaller and e_3^2 becomes negligible
- diagonal magnitudes approach singular values
- diagonal entries are arranged approximately in spectral order

Let $M = \max(\sigma_3, \sigma_4)$ and $m = \min(\sigma_3, \sigma_4)$.

The corresponding approximate condition for the right rotation becomes

$$|e_2 e_3| \leq \tau \frac{M^2 - m^2}{M}$$

The left rotation must also satisfy its similar condition.

①

Matrices with linear σ spacing

For consecutive values, $m = M - h$, where $h > 0$.

$$M^2 - m^2 = 2 M h - h^2.$$

Therefore,

$$|e_2 e_3| \leq \tau \frac{2 M h - h^2}{M}$$

$$= \tau \left(2 h - \frac{h^2}{M} \right).$$

The linear factor $2h - \frac{h^2}{M}$ is approximately constant when $M \gg h$.

②

Matrices with geometric σ spacing

Let $\rho = \frac{1}{r}$, where $r > 1$ is the ratio used in the tables.

For consecutive values, $m = \rho M$, where $0 < \rho < 1$.

$$M^2 - m^2 = (1 - \rho^2) M^2.$$

Therefore,

$$\begin{aligned} |e_2 e_3| &\leq \tau \frac{(1 - \rho^2) M^2}{M} \\ &= \tau (1 - \rho^2) M. \end{aligned}$$

The geometric factor $(1 - \rho^2) M$ scales with the larger singular value and may facilitate interior deflations, consistent with the experiments.

The experiment also shows that geometric spectra produce more interior splits under both conventional QR deflation and active CAT deflation.



Conclusions

①

Algorithm

The Chase and Trim (CAT) method provides additional deflation opportunities for SVD,

augmenting conventional bidiagonal QR with local 2×2 SVD diagonalizations.

Conventional deflation opportunities are evaluated and applied first.

The remaining superdiagonal entries are then examined for active deflation.

When a feasible candidate is found, the selected coupling is eliminated,
the local diagonal entries and singular vectors are updated,
and the active block is split into two smaller blocks.

Deflation opportunities are favored by small superdiagonal entries and
large differences between neighboring diagonal entries.

These conditions become more common as QR iteration approaches
convergence.

We show that the spectral norm of the perturbation introduced by active
deflation does not exceed that permitted by conventional passive deflation.

We present efficient rejection tests that eliminate unsuitable candidates at low
cost.

Candidates that pass these tests are submitted to the production function,
which makes the final feasibility decision and, if successful,
applies the transformation to all three matrices without explicitly forming $K^T K$.

②

Reduction in computational work

After inclusion of candidate testing and accepted rotations,
CAT reduces weighted arithmetic work, ranging
from 4.4% for IID rectangular noise
to 41.0% for Geometric singular values $[0.001, 1000.0]$.
(pooled across samples and matrix sizes)

③

Dependence on spectral structure

The effectiveness of CAT depends strongly on the singular-value distribution.

The largest work reductions occur for Geometric singular values [0.001, 1000.0],
and the smallest for IID rectangular noise.

This difference may be explained in part by split balance.

The mean smaller-child-to-parent ratio for active CAT splits was
0.017 for the latter family,
compared with 0.167 for the former.
(pooled across samples and matrix sizes)

④

Dependence on matrix size

For a given matrix family, the relative reduction in weighted arithmetic work
varies comparatively little across the tested matrix sizes.

CAT reduces work by 9.8% at $n = 64$
and 11.0% at $n = 512$.

(pooled across samples and matrix families)

Thus the singular-value distribution appears to be a stronger predictor
of relative work reduction than matrix size over the tested range.

⑤

Candidate screening

Two inexpensive rejection conditions assess the feasibility of one of the two
rotations.

They eliminate approximately 99.91% of candidates before production
verification.

They are designed as sufficient conditions for rejection, and
no valid CAT opportunities were missed by the screens in the experiments.
Of the 9,546 candidates passing the screens, only 5 were rejected by production.

Of these, 2 failed the screened rotation,
while 3 failed only the unexamined rotation.

The narrow gap between exact and approximate rejection conditions
explains their close agreement for the screened rotation when $M \geq 2\tau$.

Because these few production rejections contribute negligible overhead,
we leave verification of the second rotation to production
rather than introduce another screening test.
Screening and accepted CAT rotations together account for approximately 0.23%
of total weighted work.

⑥

Numerical accuracy

CAT preserves the accuracy of the computed singular values and vectors.
The reported singular-value errors, left and right orthogonality errors,
and reconstruction residuals remain comparable to conventional QR.

These results establish 'Chase and Trim' as a useful addition to conventional
bidiagonal QR for the tested matrix families.



References

① D. S. Watkins

Fundamentals of Matrix Computations, 3rd edition
John Wiley & Sons, 2010.

② LAPACK

DBDSQR: Singular Value Decomposition of a Real Bidiagonal Matrix Using the
Implicit QR Algorithm

LAPACK documentation and source code, Netlib.
https://www.netlib.org/lapack/lapack_routine/dbdsqr.f

③ LAPACK

DLAS2: Singular Values of a Real 2×2 Triangular Matrix

Used to compute the bidiagonal QR shift.

LAPACK documentation and source code, Netlib.

https://www.netlib.org/lapack/lapack_routine/dlas2.f

④ LAPACK

DLARTG: Generation of a Real Plane Rotation

Used to construct the Givens rotations in bidiagonal QR iteration.

LAPACK documentation and source code, Netlib.

https://www.netlib.org/lapack/lapack_routine/dlartg.f90

⑤ LAPACK

DLASR: Application of a Sequence of Plane Rotations to a Real Matrix

Used to accumulate singular-vector transformations during QR sweeps.

LAPACK documentation and source code, Netlib.

https://www.netlib.org/lapack/lapack_routine/dlasr.f

⑥ LAPACK

DLASV2: Singular Value Decomposition of a Real 2×2 Upper-Triangular Matrix

Computes the two singular values and the left and right rotations that diagonalize the matrix.

LAPACK documentation and source code, Netlib.

https://www.netlib.org/lapack/lapack_routine/dlasv2.f

⑦ LAPACK/BLAS

DROT: Application of a Plane Rotation to Two Real Vectors

Used for singular-vector updates after terminal 2×2 diagonalization.

Reference BLAS source code, Netlib.

https://www.netlib.org/lapack/lapack_routine/drot.f

⑧ LAPACK

DLATMS: Generation of Random Matrices with Specified Singular Values or

Eigenvalues

LAPACK testing matrix generator, Netlib.

https://www.netlib.org/lapack/explore-html/da/d00/dlatms_8f_source.html

⑨ P. I. Davies and N. J. Higham

Numerically Stable Generation of Correlation Matrices and Their Factors

BIT Numerical Mathematics, 40(4), 640–651, 2000.

<https://doi.org/10.1023/A:1022384216930>

⑩ J. W. Demmel, O. A. Marques, B. N. Parlett, and C. Vömel

A Testing Infrastructure for LAPACK's Symmetric Eigensolvers

LAPACK Working Note 182, 2007.

<https://www.netlib.org/lapack/lawnspdf/lawn182.pdf>