# (Q)OR Krylov Methods in Finite Precision: Inherent Characteristics.

Jens–Peter M. Zemke

Talk at the

Gesellschaft für Angewandte Mathematik und Mechanik

Workshop

Applied and Numerical Linear Algebra
with special emphasis on
Applications in Medicine and Biology

September 12–13, 2003
Technische Universität Braunschweig, Germany

## Outline.


OR and QOR (Krylov Subspace) Methods 3

Finite Precision Issues 8

Inherent Characteristics: 9
- "best iterates" 10
- "best errors" 12
- "basis vectors" 14
- "best residuals" 15

Examples: 16
- CG (Omin), BiCG (Biomin) 17
- Orthores and Ores 18

## OR and QOR (Krylov Subspace) Methods.

- o Lanczos based methods (short–term methods)
- o Arnoldi based methods (long–term methods)

- o eigensolvers: $Av = v\lambda$
- o linear system solvers: $Ax = b$
  - o (quasi-) orthogonal residual approaches: (Q)OR
  - o (quasi-) minimal residual approaches: (Q)MR

Extensions:

- o Lanczos based methods:
  - o look-ahead
  - o product-type (LTPMs)
  - o applied to normal equations (CGN)
- o Arnoldi based methods:
  - o restart (thin/thick, explicit/implicit)
  - o truncation (standard/optimal)

Krylov method: compute approximations from Krylov subspace

$$\mathcal{K}_m \equiv \mathcal{K}_m(A, q) \equiv \mathsf{span}\{q, Aq, \ldots, A^{m-1}q\}, \quad K_m = [q, Aq, \ldots, A^{m-1}q]$$

*All* Krylov methods based on Hessenberg decomposition

$$\begin{aligned} AQ_m &= Q_m C_m, \\ AQ_k &= Q_{k+1}\underline{C}_k = Q_k C_k + q_{k+1} c_{k+1,k} e_k^T \quad \forall\ k < m \end{aligned}$$

of Krylov subspace ($C_k$ Hessenberg).

Linear system context: $Ax = r_0$, $q \equiv q_1 \equiv r_0 / \|r_0\|$.

$k$th approximation $x_k \in \mathcal{K}_k$ for $x \equiv A^{-1}r_0$ expressed in terms of new basis:

$$x_k = Q_k z_k, \quad z_k \in \mathbb{K}^k.$$

Task: find *computable* expression for $z_k$.

Essentially two approaches: (Q)OR and (Q)MR.

Observe

$$
\begin{aligned}
-r_k = Ax_k - r_0 &= AQ_k z_k - Q_k e_1 \|r_0\| \\
&= Q_k(C_k z_k - e_1\|r_0\|) + q_{k+1} c_{k+1,k} z_{kk}.
\end{aligned}
$$

$\widehat{Q}_k^H$ defined by partition of pseudo-inverse:

$$
Q_{k+1}^{\dagger} Q_{k+1} = \begin{pmatrix} \widehat{Q}_k^H \\ \widehat{q}_{k+1}^H \end{pmatrix} \left[ Q_k, q_{k+1} \right] = I_{k+1}.
$$

Apply $\widehat{Q}_k^H$ to the left:

$$
-\widehat{Q}_k^H r_k = C_k z_k - e_1 \|r_0\|.
$$

When possible, set $z_k = C_k^{-1} e_1 \|r_0\|$:

$\Rightarrow$ annihilate projected residual.

Reason for notion (quasi)-orthogonal residual method, (Q)OR.

Galërkin approach. Project linear system onto smaller linear system:

$$\begin{aligned}(A, r_0) &\rightarrow (AQ_k, Q_k\|r_0\|e_1)\\ &\rightarrow (\widehat{Q}_k^H AQ_k, \widehat{Q}_k^H Q_k\|r_0\|e_1) = (C_k, \|r_0\|e_1),\end{aligned}$$

$\Rightarrow$ second matrix $\widehat{Q}_k$ (bi)orthogonal to $Q_k$.

(Q)OR related to small square matrix $C_k$.

Variety of (Q)OR methods:

OR ($\widehat{Q}_k = Q_k$, *Bubnov*-Galërkin): FOM, Orthores, CG-Ores, CG-Omin, CG-Odir, SymmLQ

QOR ($\widehat{Q}_k \neq Q_k$, *Petrov*-Galërkin): Biores, Biomin, Biodir, QOR

(Q)OR relies on iterated solution of small systems with system matrix $C_k$. Hessenberg structure of $C_k$ admits computation of decomposition along with computation of coefficients.

**Example:** LR decomposition

$$C_k = B_k R_k, \quad c_{k+1,k} = b_{k+1,k} r_{kk}.$$

$B_k$ bidiagonal since $C_k$ Hessenberg. Matrix of *direction vectors* $P_m \equiv Q_m R_m^{-1}$. Split Hessenberg decomposition:

$$\begin{aligned} AP_m &= Q_m B_m, \\ AP_k &= Q_k B_k + q_{k+1} b_{k+1,k} e_k^T \quad \forall\ k < m. \end{aligned}$$

Split Hessenberg decomposition and $Q_k = P_k R_k$ define two *coupled* recurrences.

Methods based on LR (LDLT, LDMT): CG-Omin, Biomin (BiCG)
Methods based on QR (LQ): FOM, SymmLQ, QOR

## Finite Precision Issues.

Finite precision analog of Hessenberg decomposition:

$$AQ_k - Q_k C_k = M_k - F_k, \qquad M_k = q_{k+1} c_{k+1,k} e_k^T.$$

Error term $F_k$ depends on method and implementation.

Orthores, Ores and Biores:

$$|f_k| \leq \gamma_n |A||q_k| + \gamma_g |Q_{k+1}||c_k|, \qquad g = \text{nnz}(c_k).$$

Omin, Biomin:

$$-F_k = AF_k^{(P)} + F_k^{(R)} L_k^{-1} C_k^{(0)}.$$

$(P)$ errors from direction vector recurrence,
$(R)$ errors from residual recurrence,
$L_k$ certain bidiagonal matrix, $C_k^{(0)}$ scaled Hessenberg.

Some methods: a priori bound, often: only a posteriori bounds possible.

**Inherent Characteristics.**

Inherent: We suppose no errors made in solution of small systems, i.e., we think of "best solution" possible with information at hand.
Characteristics: Behaviour common to *all* (Q)OR methods, indicators of potential drawbacks and traps.

**Assumption:** $A$, $C_k$ diagonalisable.

**Notation:** Eigendecompositions defined as

$$\begin{aligned} AV &= V\Lambda, & \hat{V}^H &\equiv V^{-1}, & \check{V}^T &\equiv V^{-1} \\ C_k S_k &= S_k \Theta_k, & \hat{S}_k^H &\equiv S_k^{-1}, & \check{S}_k^T &\equiv S_k^{-1} \end{aligned}$$

Characteristic polynomials: $\chi_{C_k}(z) \equiv \det(zI - C_k)$.

**Toolkit:** Lagrange, Hessenberg eigenvalue-eigenvector-relations ($i \leq j$)

$$\check{s}_{il} s_{jl} = \frac{\chi_{C_{i-1}}(\theta_l) \left( \prod_{\ell=i}^{j-1} c_{\ell+1,\ell} \right) \chi_{C_{j+1:k}}(\theta_l)}{\chi'_{C_k}(\theta_l)}$$

## Inherent Characteristics: "best iterates".

Diagonalised (slightly re-written) form of $AQ_k - Q_k C_k = M_k - F_k$:

$$\widehat{v}_i^H q_{k+1} = \widehat{v}_i^H Q_k \left[ \frac{\lambda_i - \theta_j}{c_{k+1,k} s_{kj}} \right] s_j + \widehat{v}_i^H F_k \left[ \frac{1}{c_{k+1,k} s_{kj}} \right] s_j.$$

"Best solution" coefficients in terms of Ritz values and Ritz vectors:

$$\begin{aligned} \frac{z_k}{\|r_0\|} &= C_k^{-1} e_1 = S_k \Theta_k^{-1} \check{S}_k^T e_1 \\ &= \sum_{j=1}^{k} \frac{\check{s}_{1j}}{\theta_j} s_j = \sum_{j=1}^{k} \left( \frac{c_{k+1,k} \check{s}_{1j} s_{kj}}{(\lambda_i - \theta_j)\,\theta_j} \right) \left( \frac{\lambda_i - \theta_j}{c_{k+1,k} s_{kj}} \right) s_j \end{aligned}$$

Thus "best iterate" defined by

$$\frac{x_k}{\|r_0\|} = Q_k \frac{z_k}{\|r_0\|} = \sum_{j=1}^{k} \left( \frac{c_{k+1,k} \check{s}_{1j} s_{kj}}{(\lambda_i - \theta_j)\,\theta_j} \right) \left[ Q_k \left( \frac{\lambda_i - \theta_j}{c_{k+1,k} s_{kj}} \right) s_j \right].$$

**Theorem:** Expression for (left) eigenpart of "best iterate":

$$\frac{\hat{v}_i^H x_k}{\|r_0\|} = \left[\sum_{j=1}^{k} \frac{c_{k+1,k}\check{s}_{1j}s_{kj}}{(\lambda_i - \theta_j)\,\theta_j}\right] \hat{v}_i^H q_{k+1} - \hat{v}_i^H F_k \left[\sum_{j=1}^{k} \left(\frac{\check{s}_{1j}}{(\lambda_i - \theta_j)\,\theta_j}\right) s_j\right].$$

Interpretation based on Hessenberg EER:

$$\sum_{j=1}^{k} \frac{c_{k+1,k}\check{s}_{1j}s_{kj}}{(\lambda_i - \theta_j)\theta_j} = \prod_{l=1}^{k} c_{l+1,l} \left(\sum_{j=1}^{k} \frac{1}{\chi'_{C_k}(\theta_j)(\lambda_i - \theta_j)} \cdot \frac{1}{\theta_j}\right)$$

Lagrange interpolation of $x^{-1}$ at knots $\theta_j$, $j \in \underline{k}$.

$$\sum_{j=1}^{k} \frac{\check{s}_{1j}s_{\ell j}}{(\lambda_i - \theta_j)\,\theta_j} = \prod_{l=1}^{\ell} c_{l+1,l} \left(\sum_{j=1}^{k} \frac{1}{\chi'_{C_k}(\theta_j)(\lambda_i - \theta_j)} \cdot \frac{\chi_{C_{\ell+1:k}}(\theta_j)}{\theta_j}\right)$$

Lagrange interpolation of $x^{-1}$ weighted by $\chi_{C_{\ell+1:k}}(\theta_j)$.

**Inherent Characteristics: "best errors".**

Similar expression for error $x - x_k$:

$$\begin{aligned}\hat{v}_i^H (x - x_k) &= \hat{v}_i^H A^{-1} r_0 - \|r_0\| \hat{v}_i^H Q_k z_k \\ &= \|r_0\| \hat{v}_i^H Q_k \left(\lambda_i^{-1} I_k - C_k^{-1}\right) e_1\end{aligned}$$

Representation in terms of Ritz vectors $s_j$:

$$\begin{aligned}\left(\lambda_i^{-1} I_k - C_k^{-1}\right) e_1 &= \left(S \lambda_i^{-1} \check{S}^T - S \Theta_k^{-1} \check{S}^T\right) e_1 \\ &= \sum_{j=1}^{k} \left(\frac{\check{s}_{1j}}{\lambda_i} - \frac{\check{s}_{1j}}{\theta_j}\right) s_j.\end{aligned}$$

Thus, we obtain

$$\begin{aligned}\frac{\hat{v}_i^H (x - x_k)}{\|r_0\|} &= \hat{v}_i^H Q_k \left( \sum_{j=1}^{k} \left(\frac{\check{s}_{1j}}{\lambda_i} - \frac{\check{s}_{1j}}{\theta_j}\right) s_j \right) \\ &= \hat{v}_i^H Q_k \left( \sum_{j=1}^{k} \left(\frac{c_{k+1,k} \check{s}_{1j} s_{kj}}{(0 - \theta_j)\lambda_i}\right) \left(\frac{\lambda_i - \theta_j}{c_{k+1,k} s_{kj}}\right) s_j \right).\end{aligned}$$

**Theorem:** Expression for (left) eigenpart of "best error":

$$\frac{\widehat{v}_i^H(x-x_k)}{\|r_0\|} = \left[\sum_{j=1}^{k} \frac{c_{k+1,k}\check{s}_{1j}s_{kj}}{0-\theta_j}\right] \frac{\widehat{v}_i^H q_{k+1}}{\lambda_i} - \frac{\widehat{v}_i^H F_k}{\lambda_i} \left[\sum_{j=1}^{k} \left(\frac{\check{s}_{1j}}{0-\theta_j}\right) s_j\right].$$

Reformulation again based on Hessenberg EER:

$$\sum_{j=1}^{k} \frac{c_{k+1,k}\check{s}_{1j}s_{kj}}{0-\theta_j} = \sum_{j=1}^{k} \frac{\prod_{p=1}^{k} c_{p+1,p}}{\chi'_{C_k}(\theta_j)(0-\theta_j)} = \frac{\prod_{p=1}^{k} c_{p+1,p}}{\chi_{C_k}(0)}.$$

$$\sum_{j=1}^{k} \frac{\check{s}_{1j}s_{lj}}{0-\theta_j} = \sum_{j=1}^{k} \frac{\left(\prod_{p=1}^{l-1} c_{p+1,p}\right)\chi_{C_{l+1:k}}(\theta_j)}{\chi'_{C_k}(\theta_j)(0-\theta_j)} = \frac{\left(\prod_{p=1}^{l-1} c_{p+1,p}\right)\chi_{C_{l+1:k}}(0)}{\chi_{C_k}(0)}.$$

## Inherent Characteristics: "basis vectors".

Influence of errors $f_l$ on the computed "basis" vectors given by

$$\frac{\hat{v}_i^H q_{k+1}}{\lambda_i} = \frac{\chi_{C_k}(0)}{\prod_{p=1}^k c_{p+1,p}} \left( \frac{\hat{v}_i^H (x - x_k)}{\|r_0\|} \right) + \sum_{l=1}^k \left[ \frac{\chi_{C_{l+1:k}}(0)}{\prod_{p=l+1}^k c_{p+1,p}} \left( \frac{\hat{v}_i^H f_l}{\lambda_i c_{l+1,l}} \right) \right].$$

Mixture of Krylov methods:

$$A^{-1} q_{k+1} = \frac{\chi_{C_k}(0)}{\prod_{p=1}^k c_{p+1,p}} \left( \frac{x - x_k}{\|r_0\|} \right) + \sum_{l=1}^k \left[ \frac{\chi_{C_{l+1:k}}(0)}{\prod_{p=l+1}^k c_{p+1,p}} \left( \frac{A^{-1} f_l}{c_{l+1,l}} \right) \right].$$

In terms of residual vectors:

$$q_{k+1} = \frac{\chi_{C_k}(0)}{\prod_{p=1}^k c_{p+1,p}} \left( \frac{r_k}{\|r_0\|} \right) + \sum_{l=1}^k \left[ \frac{\chi_{C_{l+1:k}}(0)}{\prod_{p=l+1}^k c_{p+1,p}} \left( \frac{f_l}{c_{l+1,l}} \right) \right].$$

## Inherent Characteristics: "best residuals".

The $k$th *true error* $x - x_k$ is composed of two terms, namely

$$\frac{x - x_k}{\|r_0\|} = \frac{\prod_{p=1}^{k} c_{p+1,p}}{\chi_{C_k}(0)} A^{-1} q_{k+1} - \sum_{l=1}^{k} \left[ \frac{\left(\prod_{p=1}^{l-1} c_{p+1,p}\right) \chi_{C_{l+1:k}}(0)}{\chi_{C_k}(0)} A^{-1} f_l \right].$$

This implies that the *true residual* $r_k$ can be expressed by

$$\frac{r_k}{\|r_0\|} = \frac{\prod_{p=1}^{k} c_{p+1,p}}{\chi_{C_k}(0)} q_{k+1} - \sum_{l=1}^{k} \left[ \frac{\left(\prod_{p=1}^{l-1} c_{p+1,p}\right) \chi_{C_{l+1:k}}(0)}{\chi_{C_k}(0)} f_l \right].$$

$\Rightarrow$ stopping criteria based on characteristics of $C_k$ and size of $f_l$.

**Idea:** model error behaviour by (one) Krylov method.

**Examples.**

Orthores and Orthomin variants are based on scaled Hessenberg decompositions:

$$e^T \underline{C}_m = 0, \qquad e = (1, \ldots, 1)^T.$$

(Orthores directly, Orthomin based on split Hessenberg decomposition.)

This implies:

$$\chi_{C_k}(0) = \prod_{\ell=1}^{k} c_{\ell+1,\ell} \quad \forall\ k \leq m.$$

(follows by induction; similar to Hyman's computation of determinant)

Very pleasant feature:

holds true in finite precision: Orthomin variants.

holds approximately true: Orthores variants.

## Examples: CG (Omin), BiCG (Biomin).

Formula for true residual:

$$\frac{r_k}{\|r_0\|} = \frac{\prod_{p=1}^{k} c_{p+1,p}}{\chi_{C_k}(0)} q_{k+1} - \sum_{l=1}^{k} \left[ \frac{\left(\prod_{p=1}^{l-1} c_{p+1,p}\right) \chi_{C_{l+1:k}}(0)}{\chi_{C_k}(0)} f_l \right].$$

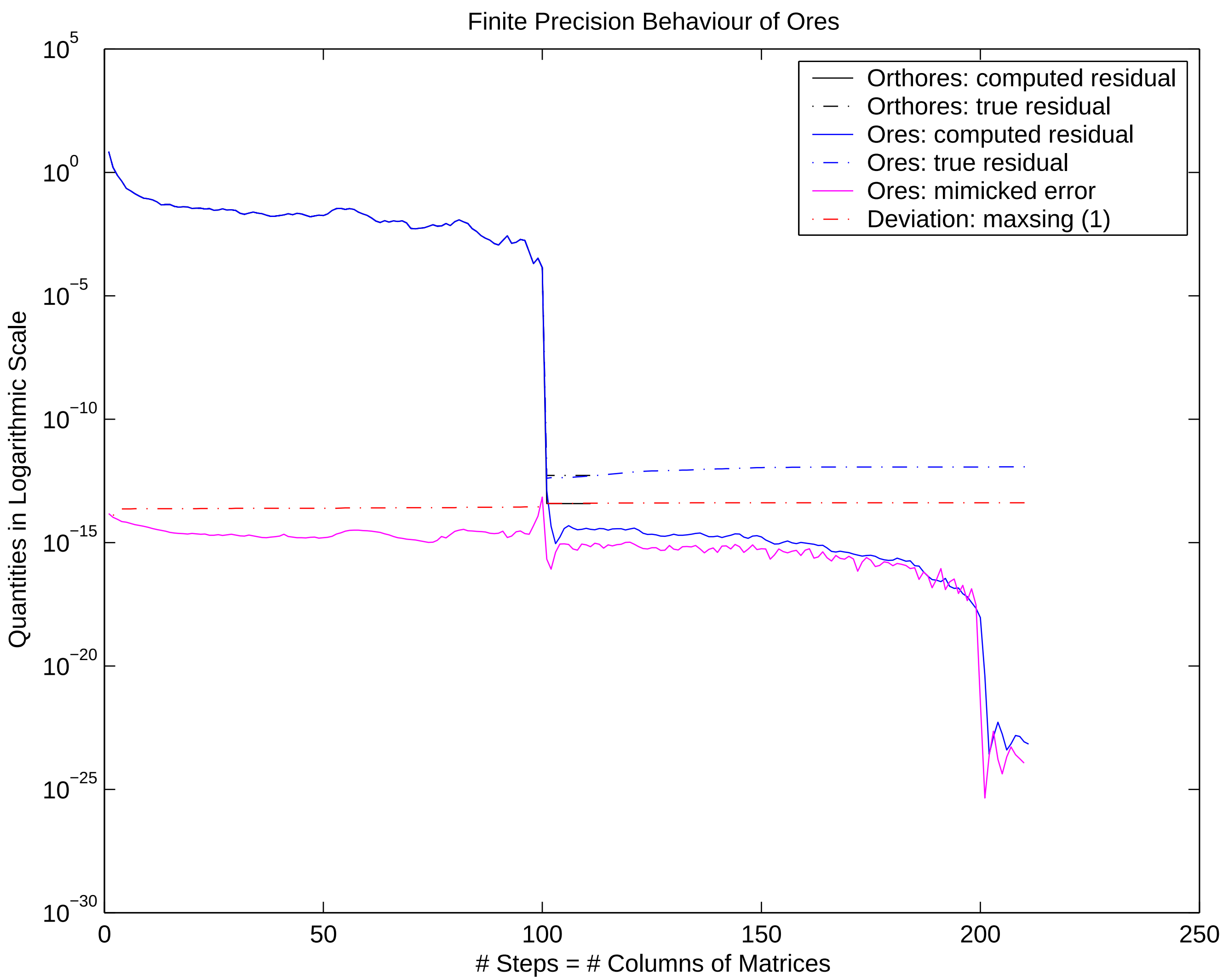
Due to underlying nice feature:

$$\begin{aligned} q_{k+1} - \frac{r_k}{\|r_0\|} &= \sum_{l=1}^{k} \frac{\chi_{C_{1:l-1}}(0)\chi_{C_{l+1:k}}(0)}{\chi_{C_k}(0)} f_l \\ &= \sum_{l=1}^{k} \left(-C_k^{-1}\right)_{ll} f_l. \end{aligned}$$

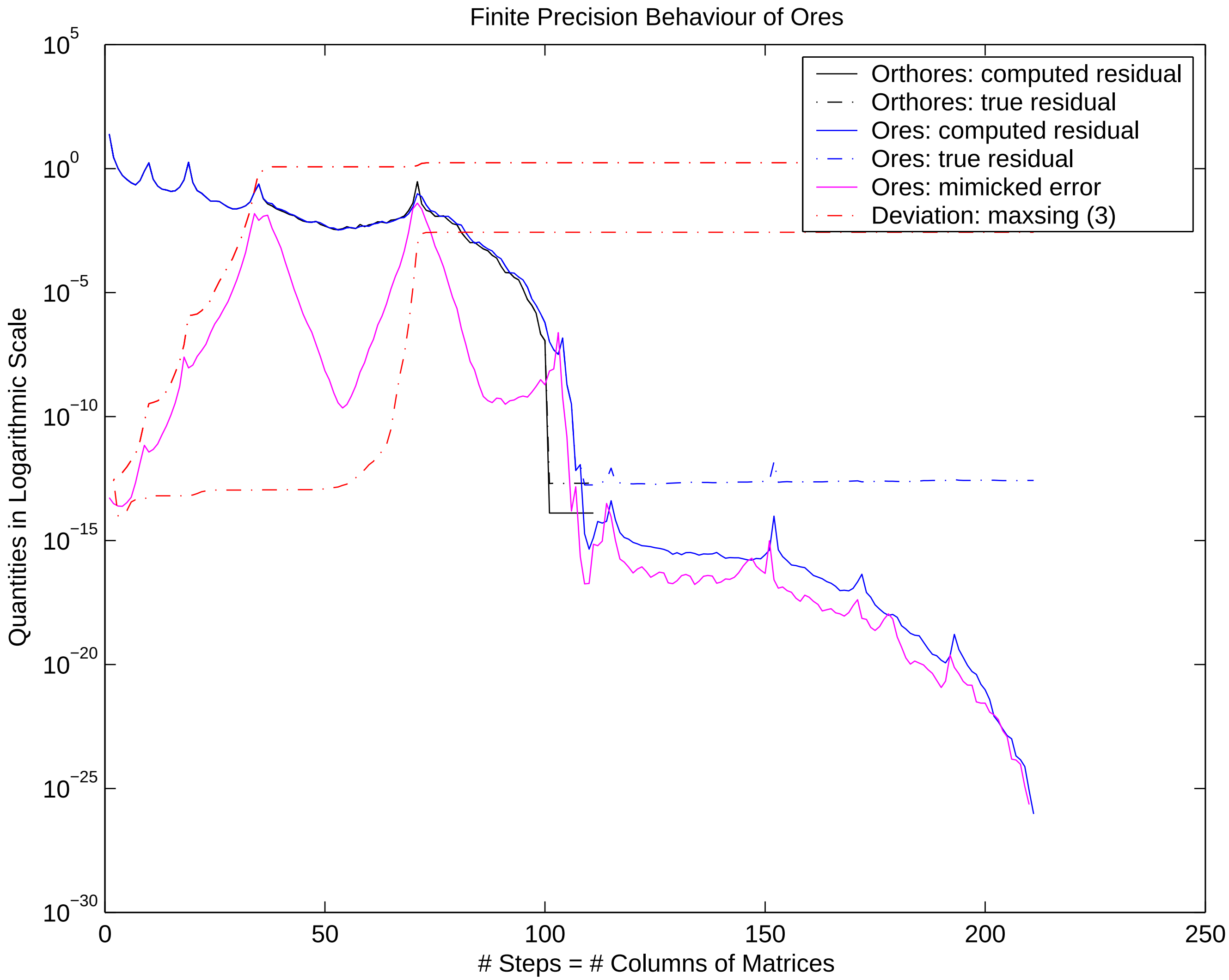
Computable bounds (twisted factorisation?):

$$\left\|\frac{r_k}{\|r_0\|} - q_{k+1}\right\| = \left\|\sum_{l=1}^{k} \left(C_k^{-1}\right)_{ll} f_l\right\| \le \operatorname{tr}\left|C_k^{-1}\right| \cdot \max_l \|f_l\|.$$

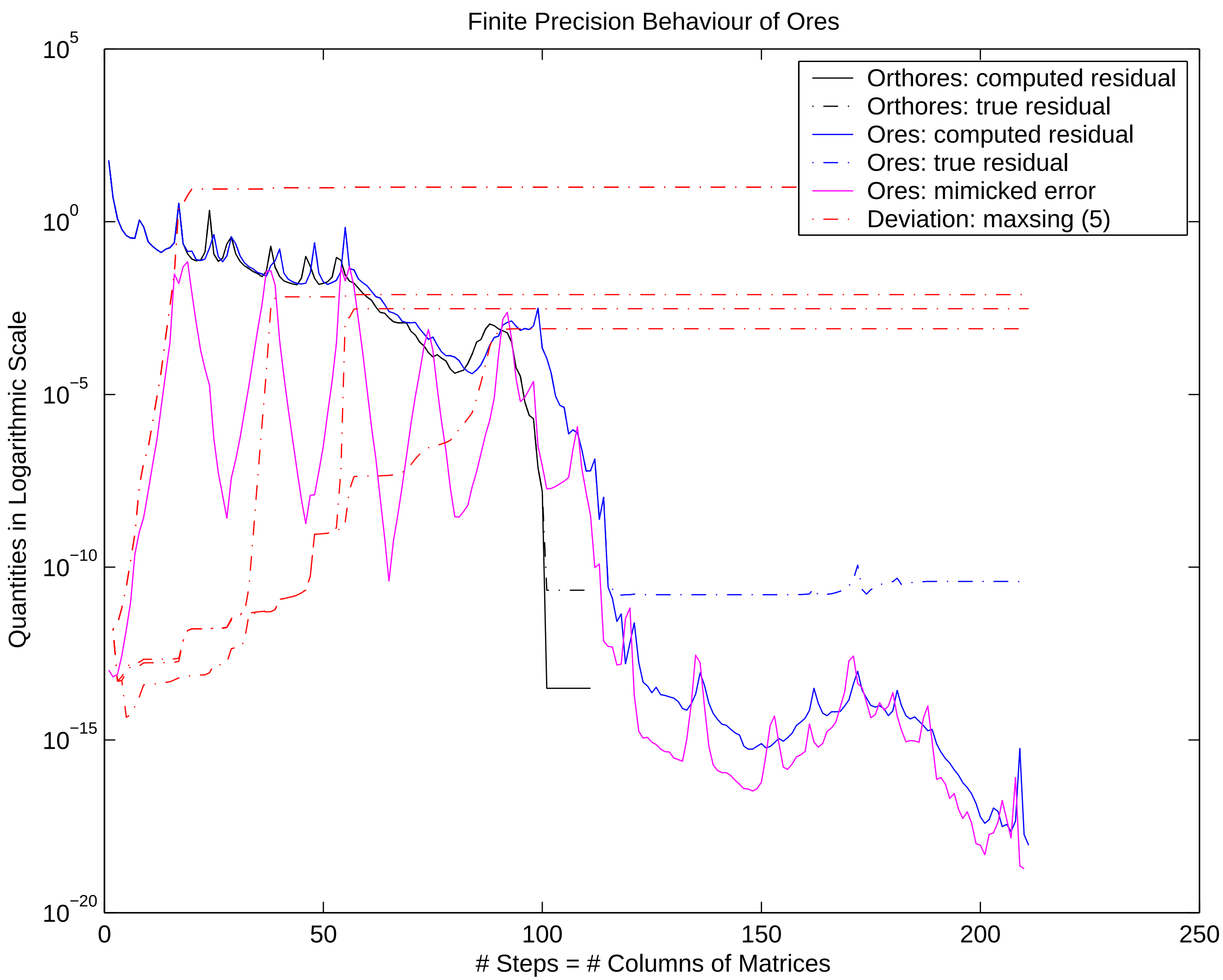
## Examples: Orthores and Ores. $T + \sigma A$, $\sigma = 0.0$



# Examples: Orthores and Ores. $T + \sigma A$, $\sigma = 0.1$



# Examples: Orthores and Ores. $T + \sigma A$, $\sigma = 0.2$



# Examples: Orthores and Ores. $T + \sigma A$, $\sigma = 1.0$