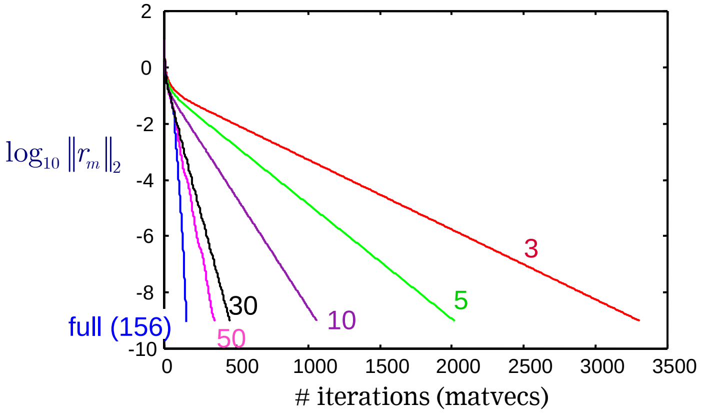
VirginiaTech
Invent the Future

Not-quite restarted GMRES

Example: Restarted GMRES

$$-\nabla \cdot (a \nabla u) = 0$$

$$u_s = 0, u_w = 1, u_n = 1, u_e = 0$$



Restarted GMRES

Restarted GMRES often leads to slow convergence or even stagnation

This poor convergence is caused by the loss of information when we restart the iteration from scratch – its hard to maintain orthogonality against vectors you've thrown away.

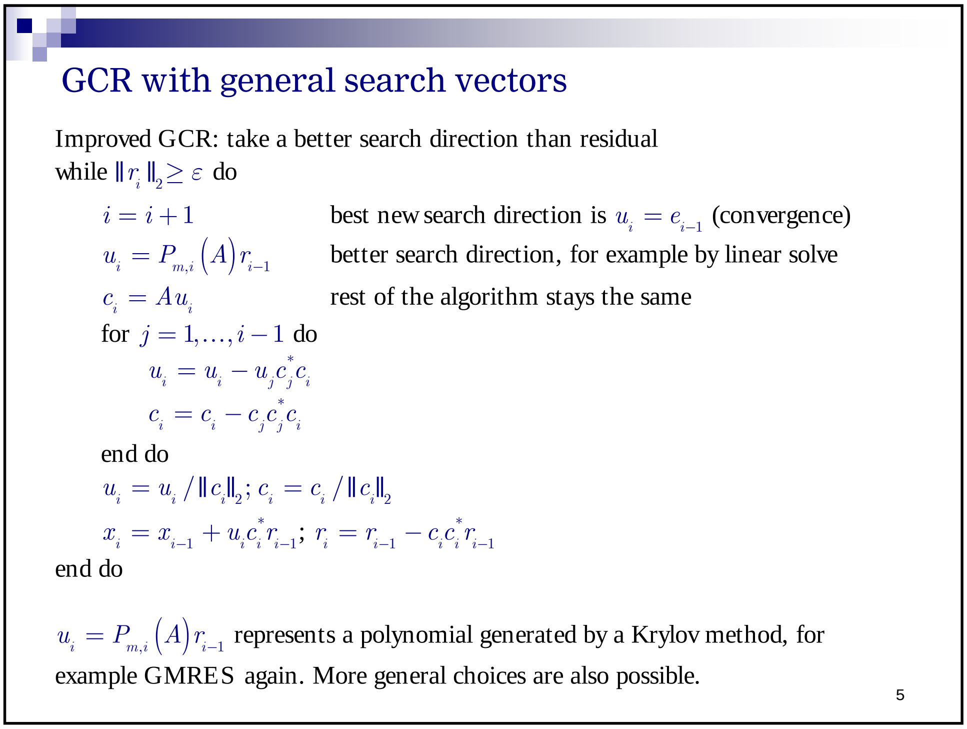
We discuss several methods that try to remedy this problem by keeping selected information from the Krylov space or otherwise including extra information:

- GMRES*/GMRESR (Vuik & van der Vorst)
- Flexible GMRES (Saad)
- GCRO (de Sturler & Fokkema, de Sturler)
- GMRESDR (Morgan)
- GCROT (de Sturler)

A more general GCR

```
GCR: $Ax = b$
Choose $x_0$ (e.g. $x_0 = 0$) and tolerance $\varepsilon$; set $r_0 = b - Ax_0$; $i = 0$
while $\| r_i \|_2 \geq \varepsilon$ do
    $i = i + 1$                                    $r_i$ adds search vector to $K_{i-1}(A, r_0)$
    $u_i = r_{i-1}$; $c_i = Au_i$                    $Ar_{i-1}$ extends $K_{i-1}(A, Ar_0)$
    for $j = 1, \ldots, i-1$ do                    (start QR decomposition)
        $u_i = u_i - u_j c_j^* c_i$                 Orthogonalize $c_i$ against previous $c_j$ and
        $c_i = c_i - c_j c_j^* c_i$                 update $u_i$ such that $Au_i = c_i$ maintained
    end do
    $u_i = u_i / \| c_i \|_2$; $c_i = c_i / \| c_i \|_2$    Normalize; (end QR decomposition)
    $x_i = x_{i-1} + u_i c_i^* r_{i-1}$              Project new $c_i$ out of residual and update
    $r_i = r_{i-1} - c_i c_i^* r_{i-1}$              solution accordingly; note $r_i \perp c_j$ for $j \leq i$
end do
```

What happens if $c_i \perp r_{i-1}$



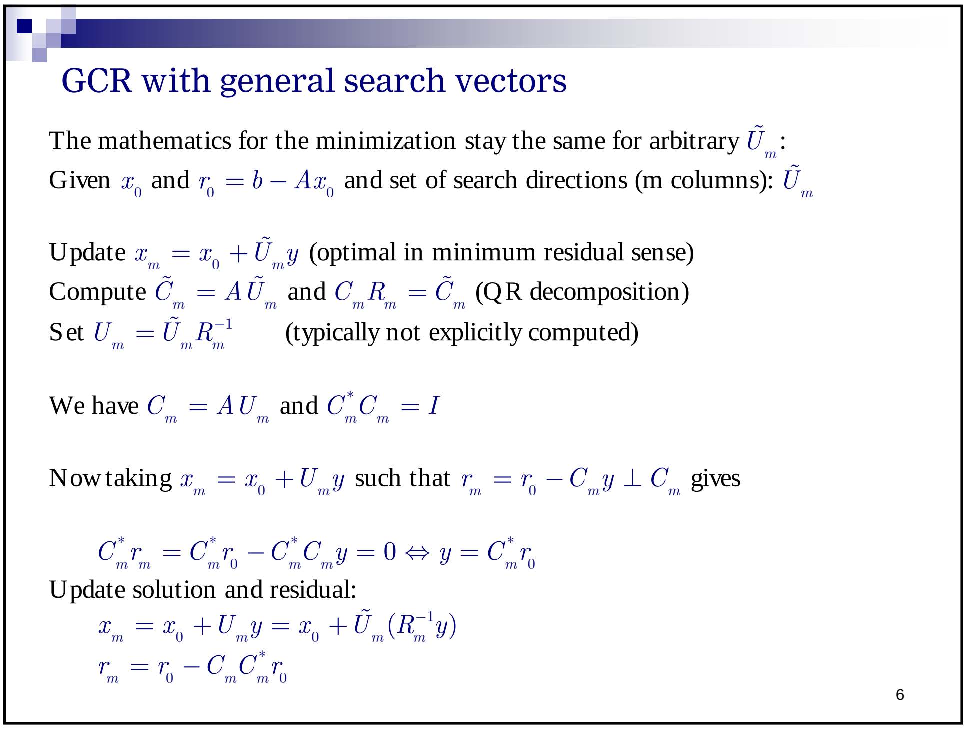
GCR with general search vectors

Improved GCR: take a better search direction than residual

```
while $\|r_i\|_2 \geq \varepsilon$ do
    $i = i + 1$                          best new search direction is $u_i = e_{i-1}$ (convergence)
    $u_i = P_{m,i}\left(A\right) r_{i-1}$    better search direction, for example by linear solve
    $c_i = A u_i$                        rest of the algorithm stays the same
    for $j = 1, \ldots, i-1$ do
        $u_i = u_i - u_j c_j^* c_i$
        $c_i = c_i - c_j c_j^* c_i$
    end do
    $u_i = u_i / \|c_i\|_2$; $c_i = c_i / \|c_i\|_2$
    $x_i = x_{i-1} + u_i c_i^* r_{i-1}$; $r_i = r_{i-1} - c_i c_i^* r_{i-1}$
end do
```

$u_i = P_{m,i}\left(A\right) r_{i-1}$ represents a polynomial generated by a Krylov method, for example GMRES again. More general choices are also possible.

5



GCR with general search vectors

The mathematics for the minimization stay the same for arbitrary $\tilde{U}_m$:
Given x_0 and $r_0 = b - Ax_0$ and set of search directions (m columns): $\tilde{U}_m$

Update $x_m = x_0 + \tilde{U}_m y$ (optimal in minimum residual sense)
Compute $\tilde{C}_m = A\tilde{U}_m$ and $C_m R_m = \tilde{C}_m$ (QR decomposition)
Set $U_m = \tilde{U}_m R_m^{-1}$ (typically not explicitly computed)

We have $C_m = AU_m$ and $C_m^* C_m = I$

Now taking $x_m = x_0 + U_m y$ such that $r_m = r_0 - C_m y \perp C_m$ gives

$$C_m^* r_m = C_m^* r_0 - C_m^* C_m y = 0 \Leftrightarrow y = C_m^* r_0$$

Update solution and residual:

$$x_m = x_0 + U_m y = x_0 + \tilde{U}_m (R_m^{-1} y)$$
$$r_m = r_0 - C_m C_m^* r_0$$

6

GCR with general search vectors: GMRESR

So take GCR and improve by better approximation to error than residual

Typically take a Krylov method to compute an approximation to e_{i-1}.
Any method is possible and we can vary these methods from one step to next
Hybrid or nested method is referred to as GMRES*

Mostly GMRES itself is taken, referred to as GMRESR (recursive)
Most popular, because it is the most robust, especially if we only take modest number of iterations (cannot make residual worse in those steps)

Still some potential problems:

- Inner method is restarted GMRES; can still stagnate for hard problem.
- If inner GMRES converges poorly, the outer optimization is not so useful
- We solve two separate optimizations rather than one global one: not optimal
    - Optimality $r_k \perp C_k$ ignores in inner GMRES, so typically update that destroys orthogonality. Restored by outer correction, but often direction computed/evaluated that are removed in outer ,minimization.

GCRO (GCR with inner orthogonalization)

Alternative: keep inner search directions orthogonal to C_k

Augmented/Adapted Arnoldi: $r_k = r_0 - C_k C_k^* r_0, \quad x_k = x_0 + U_k C_k^* r_0$

Set $v_1 = r_k \big/ \|r_k\|$, and iterate $h_{i+1,i} v_{i+1} = A v_i - C_k C_k^* A v_i - V_i V_i^* A v_i$

Recurrence: $A V_m = C_k C_k^* A V_m + V_{m+1} \underline{H}_m$ or

$$A V_m - C_k C_k^* A V_m = \left(I - C_k C_k^*\right) A V_m = V_{m+1} \underline{H}_m$$

Optimal update over range$\left([U_k \ V_m]\right)$: $x_{k+1} = x_0 + U_k z + V_m y$

$$r_{k+1} = r_0 - A U_k z - A V_m y = C_k C_k^* r_0 + v_1 \|r_k\| - C_k z - C_k C_k^* A V_m y - V_{m+1} \underline{H}_m y$$

$$r_{k+1} = C_k \left(C_k^* r_0 - C_k^* A V_m y - z\right) + V_{m+1} \left(\eta_1 \|r_k\| - \underline{H}_m y\right)$$

Choose y and z to minimize $\|r_k\|_2$.

Note that $C_k \perp V_{m+1}$ and that optimal z always makes C_k part zero.

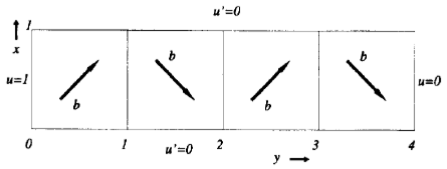
So, only need to minimize $\left(\eta_1 \|r_k\| - \underline{H}_m y\right)$ as in standard GMRES; then pick z

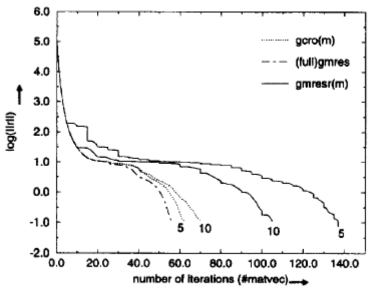
GCRO vs GMRESR

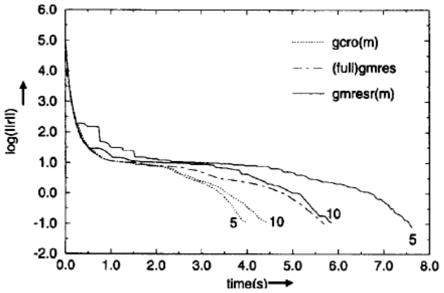
$$-(u_{xx}+u_{yy})+bu_x+cu_y=0$$

on $[0,1]\times[0,4]$, where

$$b(x,y)=\begin{cases}200 & \text{for } 0\leqslant y\leqslant 1,\\ -200 & \text{for } 1<y\leqslant 2,\\ 200 & \text{for } 2<y\leqslant 3,\\ -200 & \text{for } 3<y\leqslant 4,\end{cases}$$







FGMRES

Flexible GMRES (FGMRES): a GMRES with general search vectors
The general search vectors can be determined by an iterative solver, but usually presented as resulting from a variable preconditioner (really the same, of course).

Given x_0, $r_0 = b - Ax_0$
Set $v_1 = r_0 / \|r_0\|_2$ and iterate $h_{i+1,i}v_{i+1} = AK_i^{-1}v_i - V_iV_i^*AK_i^{-1}v_i$
This gives $AK_i^{-1}v_i = V_{i+1}h_i$ (h_i is column vector of Hessenberg matrix)
Recurrence : $A\left[K_1^{-1}v_1\; K_2^{-1}v_2\; K_3^{-1}v_3 \ldots K_m^{-1}v_m\right] = V_{m+1}\underline{H}_m \Leftrightarrow AZ_m = V_{m+1}\underline{H}_m$

Update $x_m = x_0 + Z_m y$ such that $\|r_m\|_2 = \|r_0 - AZ_m y\|_2$ minimal

We proceed as for GMRES:
$\|r_0 - AZ_m y\|_2 = \|V_{m+1}\eta_1 \|r_0\|_2 - V_{m+1}\underline{H}_m y\|_2 = \|\eta_1 \|r_0\|_2 - \underline{H}_m y\|_2$
And the final norm can be minimized by solving a small least squares problem just as for GMRES.

How to restart – What to keep

- GMRES optimal in number of iterations, but very expensive in time and memory (unless convergence very rapid).
- Limit resources: restart after m steps with best solution as initial guess or orthogonalize only against latest m vectors.
- Unfortunately, these strategies can slow down convergence drastically and even prevent convergence.
- Solution: keep some carefully selected subspace.
- Various possible choices (and ways to use them).
  - Approx. invariant subspace. Based on assumptions on normality and spectrum. Not for general problems, but often works.
  - Optimality requires orthogonal projection. After making a correction in some subspace the importance of that subspace is determined by its contribution in maintaining orthogonality.

Convergence GMRES (motivational)

Consider $Ax = b$, and relate convergence to polynomials.

$$x_m = x_0 + z_m \text{ where } z_m \in \operatorname{span}\{r_0, Ar_0, A^2 r_0, \ldots, A^{m-1} r_0\}$$
$$r_m = r_0 - Az_m \in \operatorname{span}\{r_0, Ar_0, \ldots, A^m r_0\}$$

Assume $A = U \Lambda U^{-1}$ (diagonalizable), then residual at step m

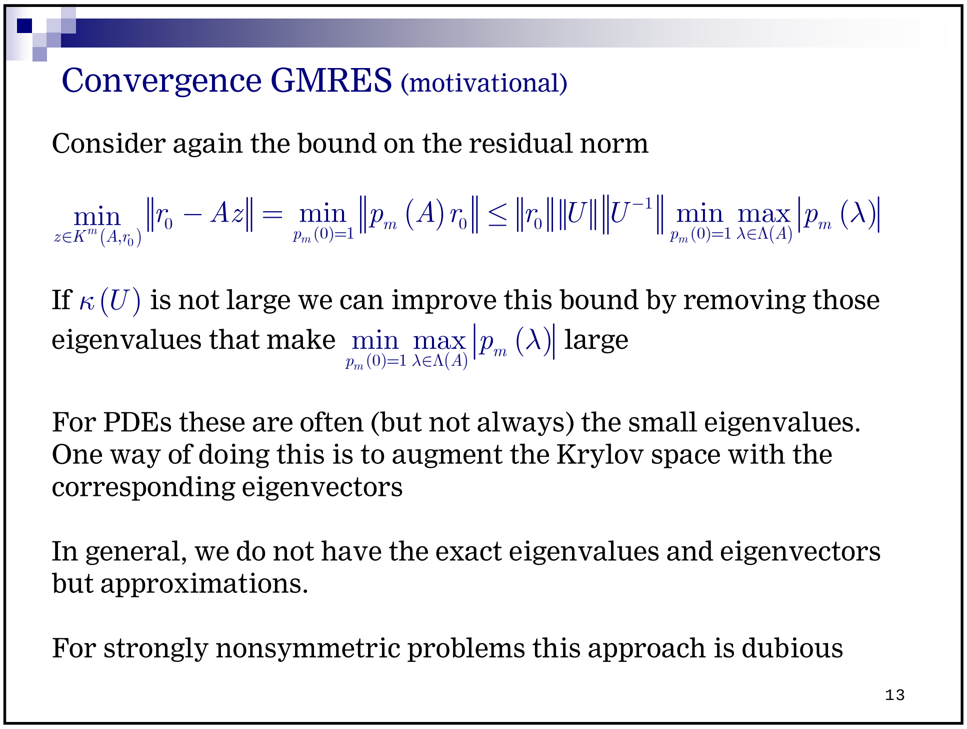
$$\min_{z \in K^m(A, r_0)} \|r_0 - Az\| = \min_{p_m(0)=1} \|p_m(A) r_0\| \le \|r_0\| \|U\| \|U^{-1}\| \min_{p_m(0)=1} \max_{\lambda \in \Lambda(A)} |p_m(\lambda)|$$

For normal matrix this bound is sharp. For highly nonnormal matrix this bound may not be useful.

$\kappa(U)$ small: convergence determined by minimal polynomial

Clustered eigenvalues yield fast convergence: preconditioning

Eigenvalues surrounding origin yields very poor convergence.



Convergence GMRES (motivational)

Consider again the bound on the residual norm

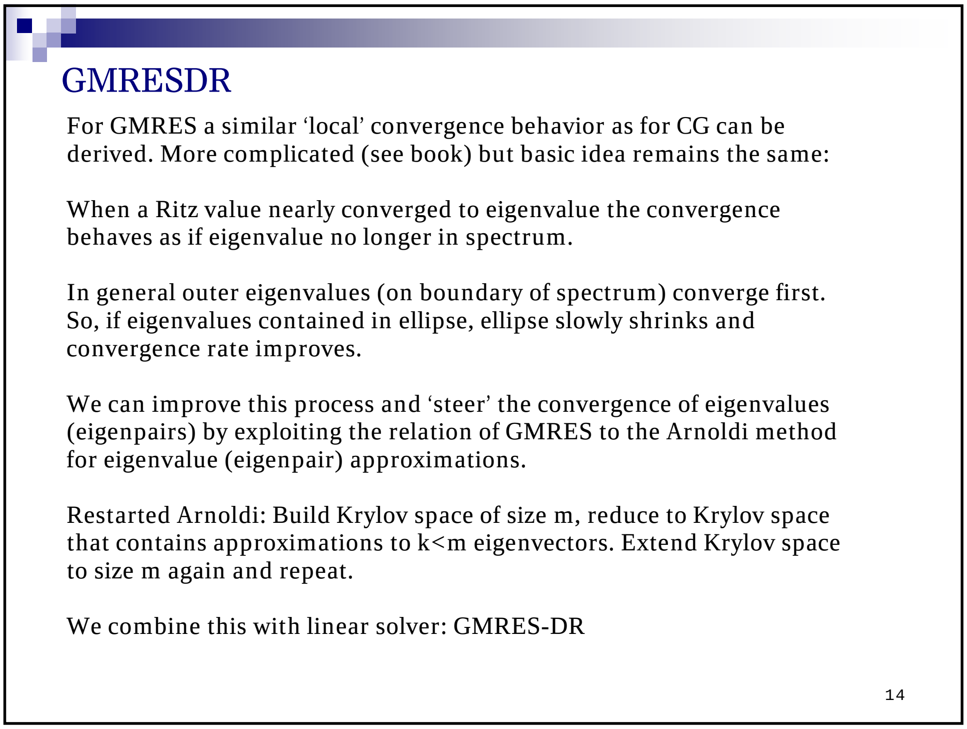
$$\min_{z\in K^m(A,r_0)}\left\|r_0-Az\right\|=\min_{p_m(0)=1}\left\|p_m(A)\,r_0\right\|\le\left\|r_0\right\|\left\|U\right\|\left\|U^{-1}\right\|\min_{p_m(0)=1}\max_{\lambda\in\Lambda(A)}\left|p_m(\lambda)\right|$$

If $\kappa(U)$ is not large we can improve this bound by removing those eigenvalues that make $\min_{p_m(0)=1}\max_{\lambda\in\Lambda(A)}\left|p_m(\lambda)\right|$ large

For PDEs these are often (but not always) the small eigenvalues. One way of doing this is to augment the Krylov space with the corresponding eigenvectors

In general, we do not have the exact eigenvalues and eigenvectors but approximations.

For strongly nonsymmetric problems this approach is dubious



GMRESDR

For GMRES a similar 'local' convergence behavior as for CG can be derived. More complicated (see book) but basic idea remains the same:

When a Ritz value nearly converged to eigenvalue the convergence behaves as if eigenvalue no longer in spectrum.

In general outer eigenvalues (on boundary of spectrum) converge first. So, if eigenvalues contained in ellipse, ellipse slowly shrinks and convergence rate improves.

We can improve this process and 'steer' the convergence of eigenvalues (eigenpairs) by exploiting the relation of GMRES to the Arnoldi method for eigenvalue (eigenpair) approximations.

Restarted Arnoldi: Build Krylov space of size m, reduce to Krylov space that contains approximations to k<m eigenvectors. Extend Krylov space to size m again and repeat.

We combine this with linear solver: GMRES-DR look in the next subsection at some needed details about some of the deflated Krylov methods. See [12] for more on deflated GMRES methods, including discussion of a related approach by De Sturler [10].

2.3. Keeping the subspace a Krylov subspace. In this subsection, we look at how an augmented subspace can still be a Krylov subspace. Notationally, we let m be the maximum dimension of the subspace and k be the number of approximate eigenvectors retained at a restart. Also, we let (θ_i, y_i) be a Ritz pair. Harmonic Ritz pairs [21, 31, 39, 27] are denoted as $(\tilde{\theta}_i, \tilde{y}_i)$. Let v_i be an Arnoldi vector from the Arnoldi recurrence [33]:

$$\begin{aligned} AV_m &= V_m H_m + h_{m+1,m} v_{m+1} e_m^T \\ &= V_{m+1} \bar{H}_m. \end{aligned} \tag{2.1}$$

Note that H_m denotes an m by m matrix, while $\bar{H}_m$ is $m+1$ by m. The ith coordinate vector is e_i. We refer to each pass through the Arnoldi iteration between restarts as one "cycle."

It was shown in [40] that when the Arnoldi method for eigenvalues is implicitly restarted with unwanted Ritz values as the shifts, the new initial vector is a combination of the desired Ritz vectors. And as given in [24] (see also [42]), the first k vectors of the new subspace are all combinations of the desired Ritz vectors. Thus the subspace

$$Span\{y_1, y_2, \ldots, y_k, v_{m+1}, Av_{m+1}, A^2 v_{m+1}, A^3 v_{m+1}, \ldots, A^{m-k-1} v_{m+1}\} \tag{2.2}$$

is the IRA subspace and is a Krylov subspace. Note that v_{m+1} is the last Arnoldi vector from the previous cycle of Arnoldi but from [40] it is also the $k+1$ Arnoldi vector in the new cycle. It is also shown in [24] that subspace (2.2) is equivalent to

$$Span\{y_1, y_2, \ldots, y_k, Ay_i, A^2 y_i, A^3 y_i, \ldots, A^{m-k} y_i\} \tag{2.3}$$

for each i such that $1 \leq i \leq k$. Thus subspace (2.2) contains Krylov subspaces with each Ritz vector as the starting vector.

In a restarted GMRES method, let r_0 be the residual vector from the previous cycle or, equivalently, the right-hand side for the new cycle. The subspace used in GMRES-E [23] is

$$Span\{r_0, Ar_0, A^2 r_0, A^3 r_0, \ldots, A^{m-k-1} r_0, \tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_k\}. \tag{2.4}$$

Thus approximate eigenvectors in the form of harmonic Ritz vectors are tacked on at the end of the Krylov subspace. It appears that putting them at the beginning would destroy the Krylov subspace. (If r_0 is orthogonalized against the harmonic Ritz vectors, then the next step of multiplying that vector by A appears to give an entirely different vector than just Ar_0.) However, as shown in [25] (see also [12]), the approximate eigenvectors can go first. This was implemented in GMRES-IR (following the approach for IRA). The approximate eigenvectors are combined in the right way so that there is an Arnoldi recurrence that can then be extended. In fact, subspace (2.4) is a Krylov subspace, though not with r_0 as the starting vector. The key is that the approximate eigenvectors are correctly chosen to be harmonic Ritz vectors. Subspace (2.4) is equivalent to

$$Span\{\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_k, A\tilde{y}_i, A^2 \tilde{y}_i, A^3 \tilde{y}_i, \ldots, A^{m-k} \tilde{y}_i\}, \tag{2.5}$$

for $1 \leq i \leq k$, so it contains Krylov subspaces with each of the harmonic Ritz vectors as starting vectors.

In Wu and Simon's thick-restart Lanczos method [48], the Ritz vectors are put in front in a simpler way. They are not combined and are not part of an Arnoldi iteration. However, they can still be extended into the Krylov subspace (2.2). The first k orthonormal basis vectors are different, but the whole subspace is the same.

3. GMRES with deflated restarting.

We look at Wu and Simon's approach to restarting [48], but for the nonsymmetric case, and we adapt it for solving linear equations. The new approach is called GMRES-DR, for GMRES with deflated restarting. We felt that name best describes what the method is trying to accomplish, although actually the term "deflated restarting" could be applied to all the approaches mentioned in subsection 2.2. The FOM version will be called FOM-DR, and it computes regular Ritz values while solving linear equations.

3.1. GMRES-DR.

The first cycle of GMRES-DR is standard GMRES with r_0 being the residual vector computed. At the end of the cycle, the k desired harmonic Ritz vectors are computed. We let V be the orthonormal matrix whose columns span the subspace. For the second cycle, the first k columns of V are formed by orthonormalizing the harmonic Ritz vectors. Then r_0 is orthogonalized against them to form v_{k+1}. From there, the rest of V can be generated with the usual Arnoldi approach.

Note that this procedure does generate the Krylov subspace (2.4); see subsection 3.3. GMRES-DR gives the same results as GMRES-IR at every iteration (not counting forming the first k columns of V), and it is mathematically equivalent to GMRES-E at the end of each cycle. We next give the algorithm. Note that because the first $k+1$ vectors of the new V are formed from the previous subspace, the orthonormalization can be done with short vectors of length m or $m+1$. However, it has been noticed that for numerical reasons, v_{k+1} needs to be reorthogonalized. We have tested this successfully with no further reorthogonalization, but it seems likely that there are cases where more reorthogonalization is needed. In the algorithm that follows, we assume that the harmonic Ritz values are distinct. (See [25] for a little discussion of the nondistinct harmonic Ritz value case.) We also assume there are at least k finite harmonic Ritz values.

GMRES-DR

1. *Start.* Choose m, the maximum size of the subspace, and k, the desired number of approximate eigenvectors. Choose an initial guess x_0 and compute $r_0 = b - Ax_0$. The recast problem is $A(x - x_0) = r_0$. Let $v_1 = r_0/||r_0||$ and $\beta = ||r_0||$.
2. *First cycle.* Apply standard GMRES(m): generate V_{m+1} and $\bar{H}_m$ with the Arnoldi iteration, solve $\min||c - \bar{H}_m d||$ for d, where $c = \beta e_1$, and form the new approximate solution $x_m = x_0 + V_m d$. Let $\beta = h_{m+1,m}$, $x_0 = x_m$, and $r_0 = b - Ax_m$. Then compute the k smallest (or others, if desired) eigenpairs $(\tilde{\theta}_i, \tilde{g}_i)$ of $H_m + \beta^2 H_m^{-T} e_m e_m^T$. (The $\tilde{\theta}_i$ are harmonic Ritz values; see [31] or [27, p. 40] for this formula.)
3. *Orthonormalization of first k vectors.* Orthonormalize $\tilde{g}_i$'s, first separating into real and imaginary parts if complex, in order to form an m by k matrix P_k. (It may be necessary to adjust k in order to make sure both parts of complex vectors are included.)
4. *Orthonormalization of $k+1$ vector.* First extend $p_1, \ldots, p_k$ (the columns of

P_k) to length $m+1$ by appending a zero entry to each. Then orthonormalize the vector $c - \bar{H}_m d$ against them to form p_{k+1}. Note $c - \bar{H}_m d$ is the length $m+1$ vector corresponding to the GMRES residual vector. P_{k+1} is $m+1$ by $k+1$.

5. *Form portions of new H and V using the old H and V.* Let $\bar{H}_k^{new} = P_{k+1}^T \bar{H}_m P_k$ and $V_{k+1}^{new} = V_{m+1} P_{k+1}$. Then let $\bar{H}_k = \bar{H}_k^{new}$ and $V_{k+1} = V_{k+1}^{new}$.
6. *Reorthogonalization of $k+1$ vector.* Orthogonalize v_{k+1} against the earlier columns of the new V_{k+1}.
7. *Arnoldi iteration.* Apply the Arnoldi iteration from this point to form the rest of V_{m+1} and $\bar{H}_m$. Let $\beta = h_{m+1,m}$.
8. *Form the approximate solution.* Let $c = V_{m+1}^T r_0$ and solve $\min ||c - \bar{H}_m d||$ for d. Let $x_m = x_0 + V_m d$. Compute the residual vector $r = b - Ax_m = V_{m+1}(c - \bar{H}_m d)$. Check $||r|| = ||c - \bar{H}_m d||$ for convergence and proceed if not satisfied.
9. *Eigenvalue computations.* Compute the k smallest (or others, if desired) eigenpairs $(\tilde{\theta}_i, \tilde{g}_i)$ of $H_m + \beta^2 H_m^{-T} e_m e_m^T$.
10. *Restart.* Let $x_0 = x_m$ and $r_0 = r$. Go to 3.

At each cycle after the first, a recurrence somewhat similar to the Arnoldi recurrence (2.1) is generated:

$$AV_m = V_{m+1} \bar{H}_m, \tag{3.1}$$

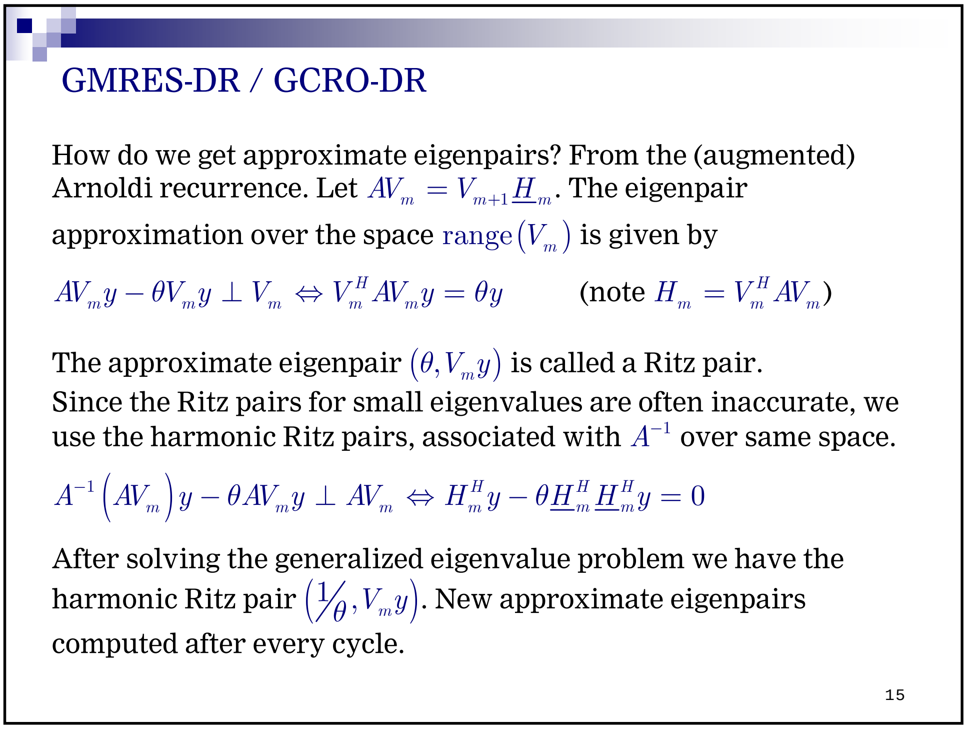
where $\bar{H}_m$ is upper-Hessenberg, *except* for a full leading $k+1$ by $k+1$ portion. Note that Schur vectors can be computed in steps 2 and 9 instead of eigenvectors.

We now look briefly at how the expense and storage of GMRES-DR compares to some previous methods. The main potential advantage of GMRES-DR compared to regular restarted GMRES is in the convergence, but it also does need only $m-k$ matrix-vector products per cycle while GMRES(m) uses m. GMRES-DR can be implemented with about the same length n storage as GMRES(m). GMRES-E is a little higher in both expense and storage than GMRES-DR. About k extra vectors of length n are normally used for GMRES-E.

GMRES-DR has about the same storage and expense requirements as GMRES-IR. The advantage of GMRES-DR is in the simplicity of the algorithm, compared to GMRES-IR. There is no QR iteration and no need for locking and purging to maintain stability, as is done for IRA and in Le Calvez and Molina's version of implicitly restarted GMRES [6]. Experiments are given in section 5 showing potential problems for GMRES-IR without lock and purge. GMRES-DR has no difficulties on the same examples. For more, see [44] in which Stewart shows stability of related eigenvalue methods.

3.2. FOM-DR. The main changes for an FOM version are that the small system of linear equations $H_m d = c$, with $c = V_m^T r_0 = \beta e_{k+1}$, is solved in step 8 instead of the small least squares problem, and the eigenvectors of H_m are computed in step 9. (This gives regular Ritz vectors instead of harmonic ones.) Step 2 is similarly changed. Also, the $k+1$ column of P_{k+1} in step 4 is just e_{m+1} with no orthonormalization. (The reorthogonalization in step 6 is still needed.)

3.3. The whole subspace is a Krylov subspace. As mentioned in subsection 2.3, it has been shown that the subspaces for GMRES-DR and FOM-DR are Krylov subspaces [25]. However, the proofs involved implicit restarting. Here we give more direct proofs.



GMRES-DR / GCRO-DR

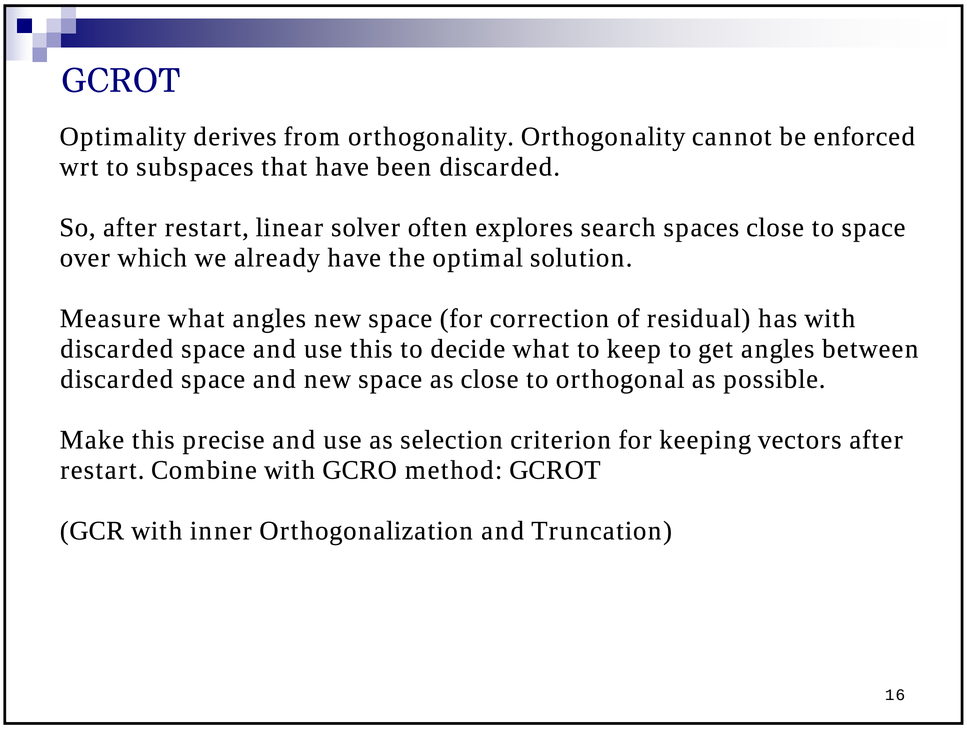
How do we get approximate eigenpairs? From the (augmented) Arnoldi recurrence. Let $AV_m = V_{m+1}\underline{H}_m$. The eigenpair approximation over the space $\mathrm{range}(V_m)$ is given by

$$AV_m y - \theta V_m y \perp V_m \Leftrightarrow V_m^H AV_m y = \theta y \qquad \text{(note } H_m = V_m^H AV_m\text{)}$$

The approximate eigenpair $(\theta, V_m y)$ is called a Ritz pair.
Since the Ritz pairs for small eigenvalues are often inaccurate, we use the harmonic Ritz pairs, associated with A^{-1} over same space.

$$A^{-1}\left(AV_m\right)y - \theta AV_m y \perp AV_m \Leftrightarrow H_m^H y - \theta \underline{H}_m^H \underline{H}_m y = 0$$

After solving the generalized eigenvalue problem we have the harmonic Ritz pair $\left(1/\theta, V_m y\right)$. New approximate eigenpairs computed after every cycle.



GCROT

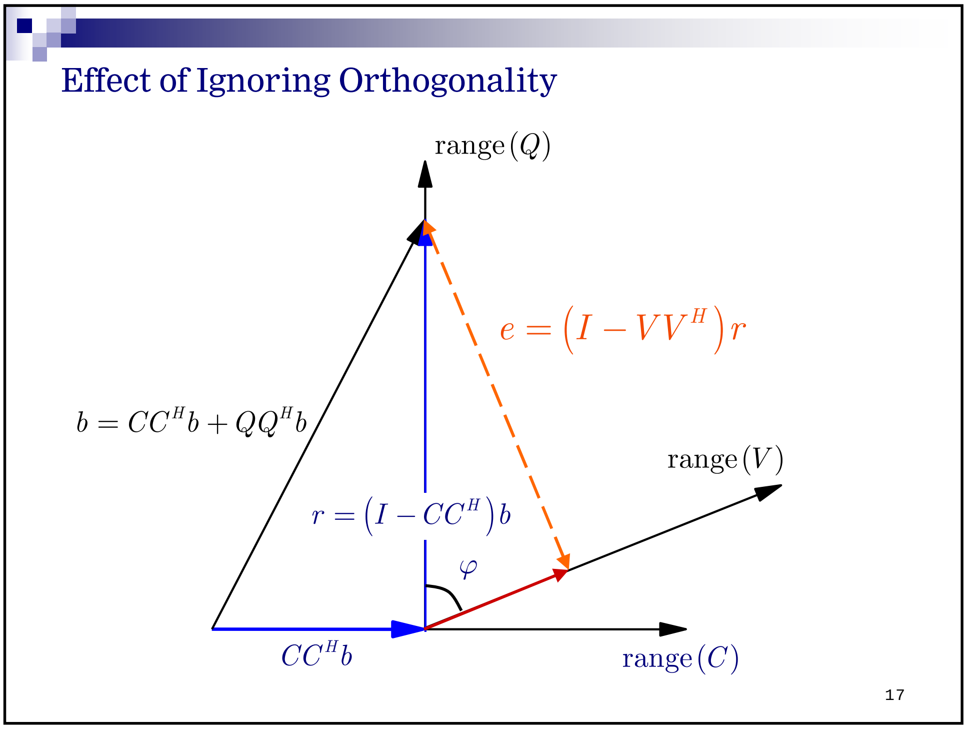
Optimality derives from orthogonality. Orthogonality cannot be enforced wrt to subspaces that have been discarded.

So, after restart, linear solver often explores search spaces close to space over which we already have the optimal solution.

Measure what angles new space (for correction of residual) has with discarded space and use this to decide what to keep to get angles between discarded space and new space as close to orthogonal as possible.

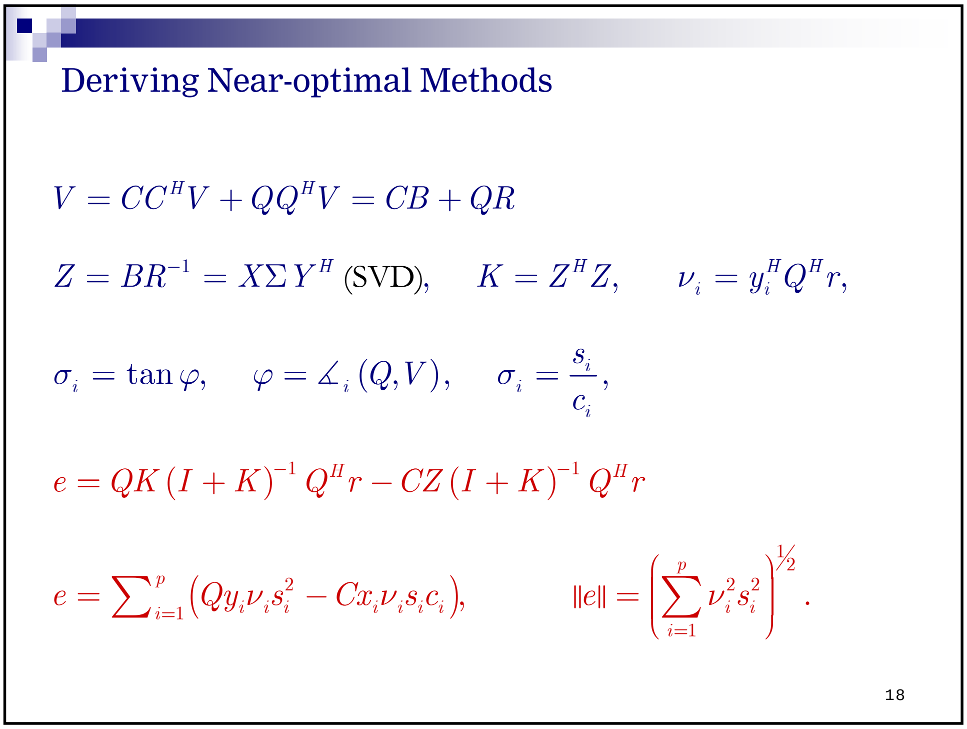
Make this precise and use as selection criterion for keeping vectors after restart. Combine with GCRO method: GCROT

(GCR with inner Orthogonalization and Truncation)



Effect of Ignoring Orthogonality

17



Deriving Near-optimal Methods

$$V = CC^H V + QQ^H V = CB + QR$$

$$Z = BR^{-1} = X \Sigma Y^H \text{ (SVD)}, \qquad K = Z^H Z, \qquad \nu_i = y_i^H Q^H r,$$

$$\sigma_i = \tan \varphi, \qquad \varphi = \measuredangle_i (Q, V), \qquad \sigma_i = \frac{s_i}{c_i},$$

$$e = QK (I + K)^{-1} Q^H r - CZ (I + K)^{-1} Q^H r$$

$$e = \sum_{i=1}^{p} \left(Q y_i \nu_i s_i^2 - C x_i \nu_i s_i c_i \right), \qquad \|e\| = \left(\sum_{i=1}^{p} \nu_i^2 s_i^2 \right)^{1/2}.$$

18

Deriving Near-optimal Methods

Let $[T \mid T_c]$ unitary and $\operatorname{rank}\left([T \,|T_c]\right) = \operatorname{rank}(\mathrm{C})$, $\hat{V} = [CT \mid V]$ and $\widehat{C} = [CT_c]$.

$$\widehat{Z} = \widehat{B}\widehat{R}^{-1} = \left[0 \middle| T_c^H Z\right] = \left[0 \middle| T_c^H X \Sigma Y^H\right]$$

Select T_c^H so that singular values of $\widehat{Z}$ minimal:

$$\min_{\dim S = p-k} \max_{u \in S} \frac{\left\|u^H Z\right\|_2}{\|u\|_2} = \sigma_{k+1} \text{ and } \qquad S = \operatorname{span}\left\{x_{k+1}, x_{k+2}, \ldots, x_p\right\}$$

$$e = \sum\nolimits_{i=k+1}^{p} \left(Q y_i \nu_i s_i^2 - C x_i \nu_i s_i c_i\right), \qquad \|e\| = \left(\sum_{i=k+1}^{p} \nu_i^2 s_i^2\right)^{1/2}.$$

GCROT

Define

$\nu = Y_m^* Q_m^* r$ where Y_m represents a change of basis

$\varphi = \left[\varphi_1 \cdots \varphi_m\right]$ principal angles $\operatorname{range}(Q_m)$ and $\operatorname{range}(V_m)$

$s = \left[\sin\varphi_1 \, \sin\varphi_2 \cdots \sin\varphi_m\right]$

Then $\|e\| = \left(\sum_{i=1}^{m} \nu_i^2 s_i^2\right)^{1/2}$

Keeping k dimensional subspace of $\operatorname{range}(C)$ can reduce this to

$$\|e\| = \left(\sum_{i=k+1}^{m} \nu_i^2 s_i^2\right)^{1/2} \qquad \text{(optimal)}$$

2. Neglecting orthogonality and optimal truncation.

We will now derive a set of equations for the residual error.

Consider the following case. We have already computed the optimal approximation to b in range(C), and the new residual is given by $r = b - CC^H b$, where C is a matrix with orthonormal columns. Now let F be a matrix with full column rank, and let $\dim(\text{range}(C) \oplus \text{range}(F)) = \dim(\text{range}(C)) + \dim(\text{range}(F))$. Then the QR-decomposition $QR = F - CC^H F$ yields the best approximation to r in the space range$(C) \oplus$ range(F), $QQ^H r$. On the other hand, the QR-decomposition $F = WS$ yields the best approximation to r in the subspace range(F), $WW^H r$. The difference between these two approximations is the *residual error* e. The residual error depends on the *principal angles* [2], [13, pp. 584–585] between the subspaces range(C) and range(F), and one way to analyze the consequences of neglecting orthogonality is to study these principal angles. However, we follow a slightly different, but equivalent, strategy. Instead of looking at the principal angles we look at the (length of the) residual error e. This approach is cheaper.[1]

DEFINITION 2.1. *Let the matrices $C \in \mathbb{C}^{n\times k}$, $F \in \mathbb{C}^{n\times m}$, and the vector $r \in \mathbb{C}^n$ be given, such that*

$$C^H C = I_k, \tag{2.1}$$
$$C^H r = 0, \tag{2.2}$$
$$\mathrm{rank}(F) = m, \tag{2.3}$$
$$\mathrm{rank}(F - CC^H F) = m. \tag{2.4}$$

Furthermore, let

$$F = CB + QR, \tag{2.5}$$
$$Q^H Q = I_m, \tag{2.6}$$

where $B = C^H F$ and R is upper-triangular. Using B and R we define

$$Z = BR^{-1} \quad = \quad (C^H F)(Q^H F)^{-1}, \tag{2.7}$$
$$K = Z^H Z, \tag{2.8}$$

and we denote the singular value decomposition of Z by

$$Z = Y_Z \Sigma_Z V_Z^H. \tag{2.9}$$

$Y_Z = [y_1\ y_2\ \dots\ y_k]$ and $V_Z = [v_1\ v_2\ \dots\ v_m]$ are ordered so as to follow the convention that

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p,$$

where $p = \min(k, m)$. Also let

$$F = WS, \tag{2.10}$$
$$W^H W = I_m, \tag{2.11}$$

and S be upper-triangular. Finally, let

$$r_1 = (I - QQ^H)r, \tag{2.12}$$
$$r_2 = (I - WW^H)r, \tag{2.13}$$

and let the residual error be $e = r_2 - r_1$.

In Definition 2.1, r_1 is the residual corresponding to the best approximation to r in the space range$(C) \oplus$ range(F), whereas r_2 is the residual corresponding to the best approximation to r in the space range(F) ignoring the orthogonality to range(C). The residual error is the difference between r_2 and r_1. Note that $C^H Q = O$ and that from (2.3)–(2.4) we know that R and S are nonsingular.

THEOREM 2.2. *The residual error e is given by*

$$r_2 - r_1 = \sum_{i=1}^{p} \left(\frac{\nu_i \sigma_i^2}{1+\sigma_i^2} Q v_i - \frac{\nu_i \sigma_i}{1+\sigma_i^2} C y_i \right), \tag{2.14}$$

[1] In practice we often do not have F explicitly available, and in our approach we do not need to orthogonalize F.

where $\nu_i = v_i^H Q^H r$, *and the norm of the residual error is given by*

$$\|r_2 - r_1\|_2 = \left(\sum_{i=1}^{p} \frac{|\nu_i|^2 \sigma_i^2}{1+\sigma_i^2} \right)^{1/2}. \tag{2.15}$$

Proof. Equations (2.12)–(2.13) give

$$r_2 - r_1 = QQ^H r - WW^H r. \tag{2.16}$$

From (2.5) and (2.10) we can derive

$$W = CBS^{-1} + QRS^{-1},$$

which leads to

$$WW^H r = CB(S^H S)^{-1} R^H Q^H r + QR(S^H S)^{-1} R^H Q^H r, \tag{2.17}$$

using $C^H r = 0$ from Definition 2.1. Again using (2.5)–(2.6) and (2.10)–(2.11), we see that

$$\begin{aligned} (WS)^H (WS) = (CB + QR)^H (CB + QR) \quad &\Leftrightarrow \\ S^H S = B^H B + R^H R, \end{aligned}$$

so that

$$\begin{aligned} R(S^H S)^{-1} R^H &= (R^{-H}(S^H S) R^{-1})^{-1} \\ &= (R^{-H} B^H B R^{-1} + I)^{-1} \\ &= (Z^H Z + I)^{-1} \\ &= (K + I)^{-1}. \end{aligned} \tag{2.18}$$

Note that all inverses above are well defined since R and S are nonsingular by definition. Substituting (2.18) into (2.17) gives

$$\begin{aligned} WW^H r &= CBR^{-1} R(S^H S)^{-1} R^H Q^H r + QR(S^H S)^{-1} R^H Q^H r \\ &= CZ(I + K)^{-1} Q^H r + Q(I + K)^{-1} Q^H r, \end{aligned} \tag{2.19}$$

and then substituting (2.19) into (2.16) gives

$$\begin{aligned} r_2 - r_1 &= QQ^H r - CZ(I + K)^{-1} Q^H r - Q(I + K)^{-1} Q^H r \\ &= QK(I + K)^{-1} Q^H r - CZ(I + K)^{-1} Q^H r. \end{aligned} \tag{2.20}$$

Using (2.9) we can rewrite (2.20) as

$$r_2 - r_1 = QV_Z(\Sigma_Z^H \Sigma_Z)(I + \Sigma_Z^H \Sigma_Z)^{-1} V_Z^H Q^H r - CY_Z \Sigma_Z (I + \Sigma_Z^H \Sigma_Z)^{-1} V_Z^H Q^H r. \tag{2.21}$$

From $\nu_i = v_i^H Q^H r$ we have

$$V_Z^H Q^H r = \sum_{i=1}^{m} \nu_i e_i, \tag{2.22}$$

where e_i is the ith Cartesian basis vector. Finally, the substitution of (2.22) into (2.21) gives

$$r_2 - r_1 = \sum_{i=1}^{p} \left(\frac{\nu_i \sigma_i^2}{1+\sigma_i^2} Q v_i - \frac{\nu_i \sigma_i}{1+\sigma_i^2} C y_i \right),$$

and the norm of the residual error follows immediately from the orthogonality of C and Q. $\square$

Theorem 2.2 indicates that for $\sigma_i = 0$, corresponding to a direction in range(F) orthogonal to range(C), the associated component in the residual error is zero, and for $\sigma_i \to \infty$, corresponding to a direction in range(F) that becomes dependent with range(C), the associated component in the residual error equals the optimal correction. Thus, no correction is made in that direction.

We have derived equations for the residual error that show the consequences of neglecting the orthogonality to range(C), that is, the consequences of discarding range(C) by truncation or restart. In the next subsection we show how Theorem 2.2 can be used to obtain the (components of the) residual error in the case of discarding an arbitrary subspace of range(C). This will then be used to select subspaces to discard or to keep in order to maintain good convergence at low cost.

Optimal truncation. We will now use the results of the previous subsection to determine which subspace of range(C) should be kept and what can be discarded. We consider computing the residual r_3 while maintaining orthogonality to the subspace range(CT) and neglecting orthogonality to the subspace range(CT_c). By defining T appropriately we can select arbitrary subspaces of range(C). T_c is the complement of T (see below). In the following we use the notation $[X|Y]$ to indicate the matrix that is formed by appending the columns of Y to the matrix X; we will use a similar notation for appending rows to a matrix.

DEFINITION 2.3. *Let C, F, Q, and r be as in Definition* 2.1, *and let the matrix $[T|T_c]$ be a square, unitary matrix such that* rank$([T|T_c])$ = rank(C), *and $T \in \mathbb{C}^{k\times l}$. Now let $\bar{F} = [CT|F]$, $\bar{C} = CT_c$, and $\bar{Q} = [CT|Q]$, and let*

$$\bar{B} = \bar{C}^H \bar{F} = [0|T_c^H B], \tag{2.23}$$

$$\bar{R} = \bar{Q}^H \bar{F} = \left[\begin{array}{c|c} I & T^H B \\ \hline 0 & R \end{array} \right]. \tag{2.24}$$

Using $\bar{B}$ and $\bar{R}$, we define

$$\bar{Z} = \bar{B}\bar{R}^{-1} \quad = \quad [0|T_c^H Z] \tag{2.25}$$

(2.26)

and $\bar{K} = \bar{Z}^H \bar{Z}$. We denote the singular value decomposition of $\bar{Z}$ by

$$\bar{Z} = Y_{\bar{Z}} \Sigma_{\bar{Z}} V_{\bar{Z}}^H, \tag{2.27}$$

where $Y_{\bar{Z}} = [\bar{y}_1\ \bar{y}_2\ \ldots\ \bar{y}_{k-l}]$ and $V_{\bar{Z}} = [\bar{v}_1\ \bar{v}_2\ \ldots\ \bar{v}_{m+l}]$ are ordered such as to follow the convention that

$$\bar{\sigma}_1 \geq \bar{\sigma}_2 \geq \cdots \geq \bar{\sigma}_{\min(k-l,m+l)}.$$

Furthermore, let

$$\bar{F} = \bar{W}\bar{S}, \tag{2.28}$$

$$\bar{W}^H \bar{W} = I_{m+l}, \tag{2.29}$$

where $\bar{S}$ is upper-triangular, and let

$$r_3 = (I - \bar{W}\bar{W}^H)r. \tag{2.30}$$

Finally, let the residual error from discarding range(CT_c) *be given by $\bar{e} = r_3 - r_1$.*

We say that the subspace range(CT) is kept and that the subspace range(CT_c) is discarded. Note that $\bar{F} = \bar{C}\bar{B} + \bar{Q}\bar{R}$ (cf. Definition 2.1). The residual r_3 corresponds to the best approximation to r in the space range$(CT) \oplus$ range(F). Following Theorem 2.2 we can derive an equation for the residual error $\bar{e}$ depending on T. From this equation we can derive what the best choice for T is.

THEOREM 2.4. *The residual error $\bar{e}$ is given by*

$$r_3 - r_1 = \sum_{i=1}^{\min(k-l,m+l)} \left(\frac{\bar{\nu}_i \bar{\sigma}_i^2}{1+\bar{\sigma}_i^2} \bar{Q}\bar{v}_i - \frac{\bar{\nu}_i \bar{\sigma}_i}{1+\bar{\sigma}_i^2} \bar{C}\bar{y}_i \right), \tag{2.31}$$

where $\bar{\nu}_i = \bar{v}_i^H \bar{Q}^H r$, and its norm is given by

$$\|r_3 - r_1\|_2 = \left(\sum_{i=1}^{\min(k-l,m+l)} \frac{|\bar{\nu}_i|^2 \bar{\sigma}_i^2}{1+\bar{\sigma}_i^2} \right)^{1/2}. \tag{2.32}$$

Proof. The proof follows immediately from Theorem 2.2. □

Analogous to (2.15) the norm of the residual error is determined by the singular values of $\bar{Z}$ and by the values $\bar{\nu}_i$. The smaller the singular values are, the smaller the residual error will be. If we can bound the singular values from above by some small value, then the error cannot be large. Likewise, if we want to maintain orthogonality to a subspace of dimension l, then we should truncate such that the l largest singular values from Z are removed in $\bar{Z}$. For the moment we ignore the fact that the coefficient $\bar{\nu}_i$ may be very small, in which case the size of $\bar{\sigma}_i$ does not matter.

So, we want to choose T_c (and hence T) such that the maximum singular value of $\bar{Z}$ is minimized. How to achieve this is indicated by the following min-max theorem, which is an obvious variant of Theorems 3.1.2 and 3.3.15 in [15, p. 148 and pp. 177–178].

THEOREM 2.5. *Let Z and its singular value decomposition be as in Definition* 2.1, *and let T_c be as in Definition* 2.3. *Then*

$$\min_{\substack{S \subset \mathbb{C}^k \\ \dim(S) = k-l}} \max_{\substack{x \in S \\ \|x\|_2 = 1}} \|x^H Z\|_2 = \begin{cases} \sigma_{l+1} & \text{if } l+1 \le p \\ 0 & \text{if } l+1 > p, \end{cases} \tag{2.33}$$

and the minimum is found for

$$S = \text{span}\{y_{l+1}, y_{l+2}, \ldots, y_k\}. \tag{2.34}$$

This is equivalent to

$$\min_{\substack{T_c \in \mathbb{C}^{k\times(k-l)} \\ T_c^H T_c = I_{k-l}}} \max_{\substack{\xi \in \mathbb{C}^{k-l} \\ \|\xi\|_2 = 1}} \|(T_c\xi)^H Z\|_2 = \begin{cases} \sigma_{l+1} & \text{if } l+1 \le p \\ 0 & \text{if } l+1 > p, \end{cases} \tag{2.35}$$

and the minimum is found for T_c*, such that*

$$\text{range}(T_c) = S = \text{span}\{y_{l+1},\, y_{l+2},\, \ldots,\, y_k\}, \tag{2.36}$$

$$x = T_c\xi. \tag{2.37}$$

Proof. The proof is a variant of the proofs of Theorems 3.1.2 and 3.3.15 in [15, p. 148 and pp. 177–178]. □

Since $\bar{Z} = [0|T_c^H Z]$ and $\|(T_c\xi)^H Z\|_2 = \|\xi^H(T_c^H Z)\|_2 = \|\xi^H \bar{Z}\|_2$, it is clear that the choice for T_c in (2.36) minimizes the maximum singular value of $\bar{Z}$.

Now from Theorem 2.5 the most obvious choices for the optimal truncation T and its complement T_c are

$$T = [y_1\, y_2\, \ldots\, y_l], \tag{2.38}$$

$$T_c = [y_{l+1}\, y_{l+2}\, \ldots\, y_k]. \tag{2.39}$$

For this particular choice of T and T_c, we can derive the singular value decomposition of $\bar{Z}$ immediately from the singular value decomposition of Z. From (2.25) and the singular value decomposition of Z (Definition 2.1) we get $\bar{Z} = Y_{\bar{Z}}\Sigma_{\bar{Z}}V_{\bar{Z}}^H$, where

$$Y_{\bar{Z}} = [e_1\, e_2\, \ldots\, e_{p-l}|\star], \tag{2.40}$$

$$\Sigma_{\bar{Z}} = \text{diag}(\sigma_{l+1}, \sigma_{l+2}, \ldots, \sigma_p, 0, \ldots, 0)_{\text{(k-l)} \times \text{(m+l)}}, \tag{2.41}$$

$$V_{\bar{Z}} = \left[\begin{array}{c|c|c} 0 & v_{l+1}\, v_{l+2}\, \ldots\, v_p & \star \\ \hline \star & 0 & 0 \end{array}\right] \tag{2.42}$$

(see [8]). Here the $\star$ symbols denote any submatrices that satisfy the respective conditions that $Y_{\bar{Z}}$ and $V_{\bar{Z}}$ be unitary matrices.

We can now give the following theorem about the residual error $\bar{e} = r_3 - r_1$ and its norm.

THEOREM 2.6. *The residual error* $\bar{e} = r_3 - r_1$ *that results from the truncation defined by the matrix* T *from* (2.38) *is given by*

$$r_3 - r_1 = \sum_{i=l+1}^{p} \left(\frac{\nu_i \sigma_i^2}{1+\sigma_i^2} Q v_i - \frac{\nu_i \sigma_i}{1+\sigma_i^2} C y_i \right), \tag{2.43}$$

and its norm is given by

$$\|r_3 - r_1\|_2 = \left(\sum_{i=l+1}^{p} \frac{|\nu_i|^2 \sigma_i^2}{1+\sigma_i^2} \right)^{1/2}. \tag{2.44}$$

Proof. The proof follows from Theorems 2.2 and 2.4, using $\bar{Q}$, $V_{\bar{Z}}$, $\Sigma_{\bar{Z}}$, $\bar{C}$, and $Y_{\bar{Z}}$ from Definition 2.3, and (2.40)–(2.42). We also use the possibility to take $Y_{\bar{Z}} = I$; see (2.40). For details we refer to [8]. □

The effect of restarting GMRES and selecting the subspace to keep. We will now analyze the residual error that results from restarting GMRES, and select the subspace to keep after m iterations of GMRES. The implementation will be given in the next section.

Given some iteration $s < m$, our analysis gives the following information. First, how much worse the convergence would have been after m iterations, if we had restarted after s iterations, that is, discarded range(AW_s). Second, which subspace from the first s iterations we should have kept, in order to have in the remaining $(m-s)$-iterations convergence as close as possible to that of the full m GMRES iterations.

After m iterations of GMRES, starting with $w_1 = r_0/\|r_0\|$, we have from (1.12) $AW_m = W_{m+1}\bar{H}_m$, and (1.15) gives the orthonormal basis for range(AW_m):

$$W_{m+1}\bar{Q}_m = W_{m+1}[q_1\, q_2\, \cdots\, q_m]. \tag{2.45}$$

Furthermore, we have $AW_s = W_{s+1}\bar{H}_s$, and the residual r_s is given by (1.16):

$$\begin{aligned} r_s &= r_0 - W_{s+1}\bar{Q}_s\bar{Q}_s^H W_{s+1}^H r_0 \\ &= W_{s+1}(I - \bar{Q}_s\bar{Q}_s^H)\|r_0\|_2 e_1 \\ &= W_{s+1}\tilde{q}_{s+1}\tilde{q}_{s+1}^H\|r_0\|_2 e_1. {}^{2} \end{aligned} \tag{2.46}$$

We define $\rho_s = (\tilde{q}_{s+1}^H e_1)\|r_0\|_2\, \tilde{q}_{s+1}$. Then $r_s = W_{s+1}\rho_s$, which we can also write as $r_s = W_{m+1}\rho_s$ with some abuse of notation.[3]

Now, consider a restart of GMRES with r_s as initial residual and making $m-s$ iterations. Clearly, range$(AW_m) =$ range$(AW_s) \oplus AK^{m-s}(A, r_s)$. Using (2.45)–(2.46) and following the notation of Definition 2.1, we take

$$C = W_{s+1}\bar{Q}_s = W_{m+1}[q_1\, \cdots\, q_s], \tag{2.47}$$

$$Q = W_{m+1}[q_{s+1}\, \ldots\, q_m]. \tag{2.48}$$

For F we can take any basis of $AK^{m-s}(A, r_s)$, because any matrix whose columns form a basis for range(F) gives the same matrix Z. Let $F = MS$, with S invertible; then we have

$$Z = (C^H F)\,(Q^H F)^{-1} = (C^H M)S\,S^{-1}(Q^H M)^{-1} = (C^H M)\,(Q^H M)^{-1}.$$

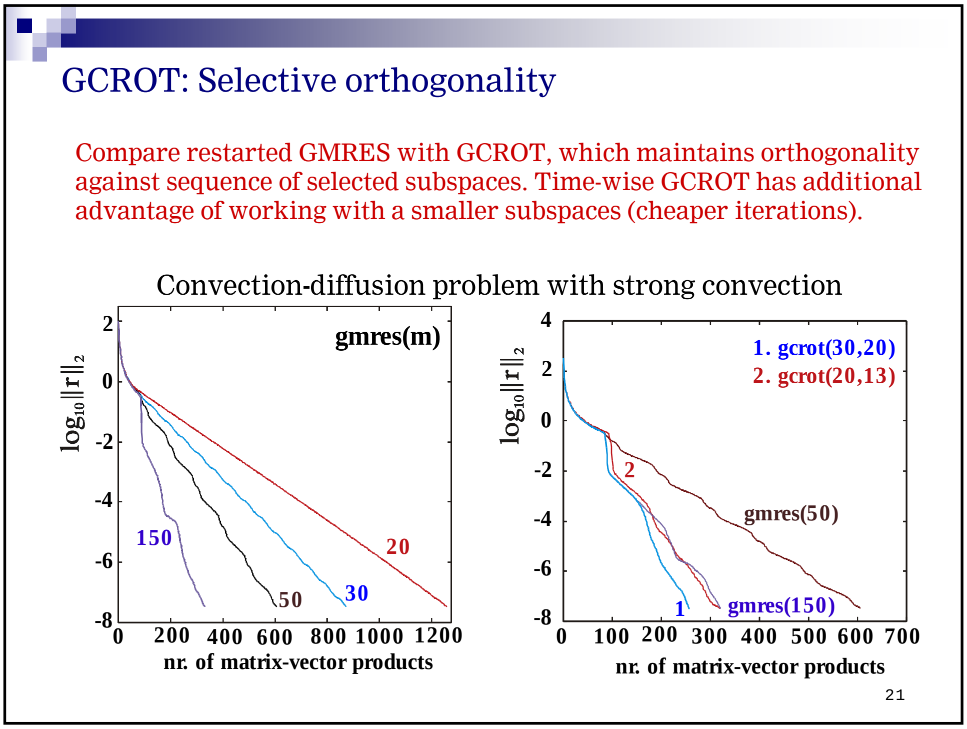
So F can be represented implicitly by

$$F = W_{m+1}[(\bar{H}_{s+1}\rho_s)\,(\bar{H}_{s+2}\bar{H}_{s+1}\rho_s)\,\ldots\,(\bar{H}_m\cdots\bar{H}_{s+2}\bar{H}_{s+1}\rho_s)].$$

In practice we generate F by an Arnoldi iteration with $\bar{H}_m$ and ρ_s. Now, following Definition 2.1, we compute $B = C^H F$ and $R = Q^H F$. The optimal residual after m GMRES iterations is given by r_1 in Definition 2.1. The residual after making first s GMRES iterations, restarting, and then making another $m-s$ GMRES iterations is given by r_2 in Definition 2.1. The difference between the two residuals, the residual error e, is given by Theorem 2.2, where $r = r_s$, and $Z = BR^{-1}$, with B and R computed as above. The singular value decomposition of Z not only describes the loss of convergence because of restarting (discarding C), according to Theorem 2.2, but it also indicates which vectors from the first s iterations we should have kept for good convergence in the remaining $m-s$ iterations, according to Theorems 2.5 and 2.6. Of course, we can choose any $s < m$, and we can do the analysis for several values of s if we want.

[2] $\tilde{q}_{s+1}$ is the last column of the matrix generated by the first s Givens rotations; $\tilde{q}_{s+1} \neq q_{s+1}$ since $\tilde{q}_{s+1}$ will be changed by the next Givens rotation.

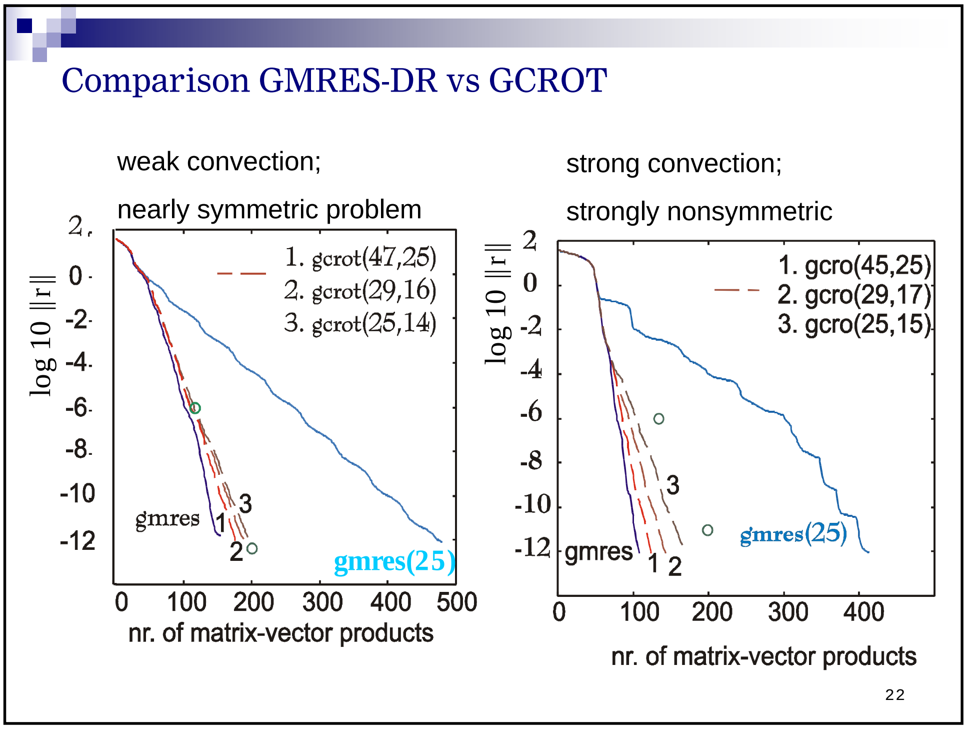
[3] We will assume every vector suitably adjusted by adding zeros at the end. Likewise, we will assume every matrix suitably adjusted by adding zero rows at the bottom.



GCROT: Selective orthogonality

Compare restarted GMRES with GCROT, which maintains orthogonality against sequence of selected subspaces. Time-wise GCROT has additional advantage of working with a smaller subspaces (cheaper iterations).

Convection-diffusion problem with strong convection



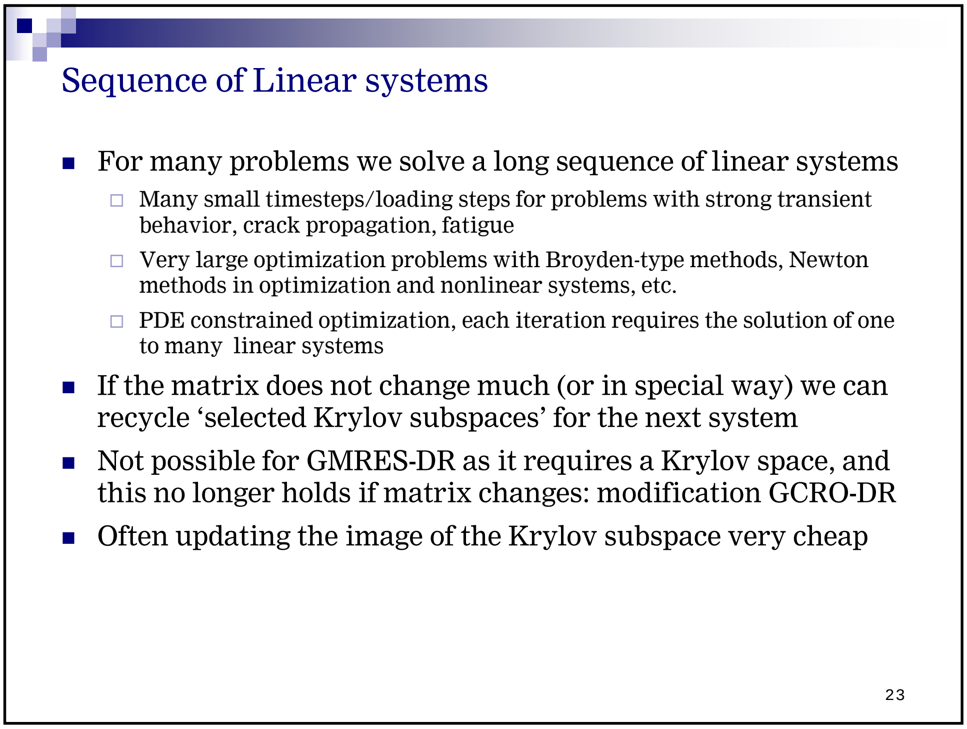
Comparison GMRES-DR vs GCROT

weak convection;

nearly symmetric problem

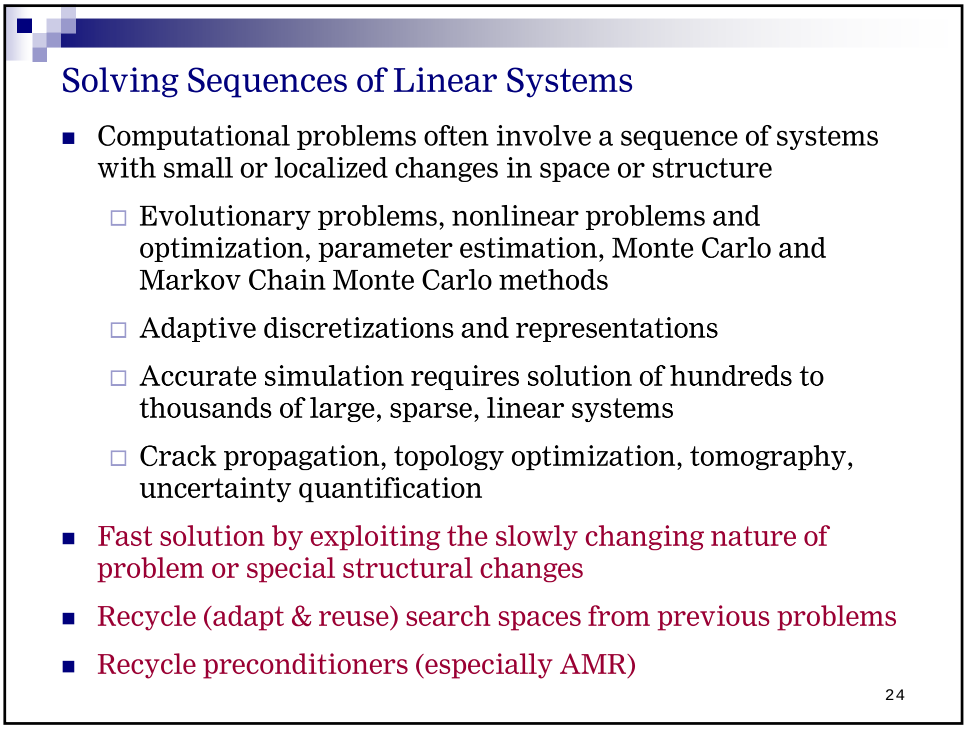
strong convection;

strongly nonsymmetric



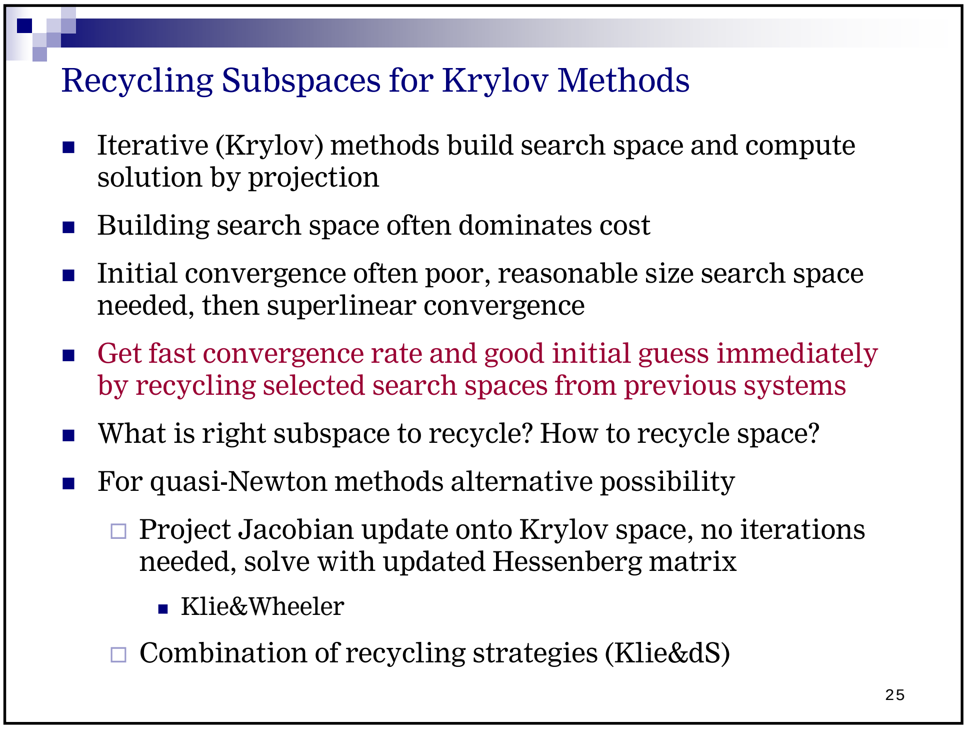
Sequence of Linear systems

- For many problems we solve a long sequence of linear systems
  - □ Many small timesteps/loading steps for problems with strong transient behavior, crack propagation, fatigue
  - □ Very large optimization problems with Broyden-type methods, Newton methods in optimization and nonlinear systems, etc.
  - □ PDE constrained optimization, each iteration requires the solution of one to many linear systems
- If the matrix does not change much (or in special way) we can recycle ‘selected Krylov subspaces’ for the next system
- Not possible for GMRES-DR as it requires a Krylov space, and this no longer holds if matrix changes: modification GCRO-DR
- Often updating the image of the Krylov subspace very cheap



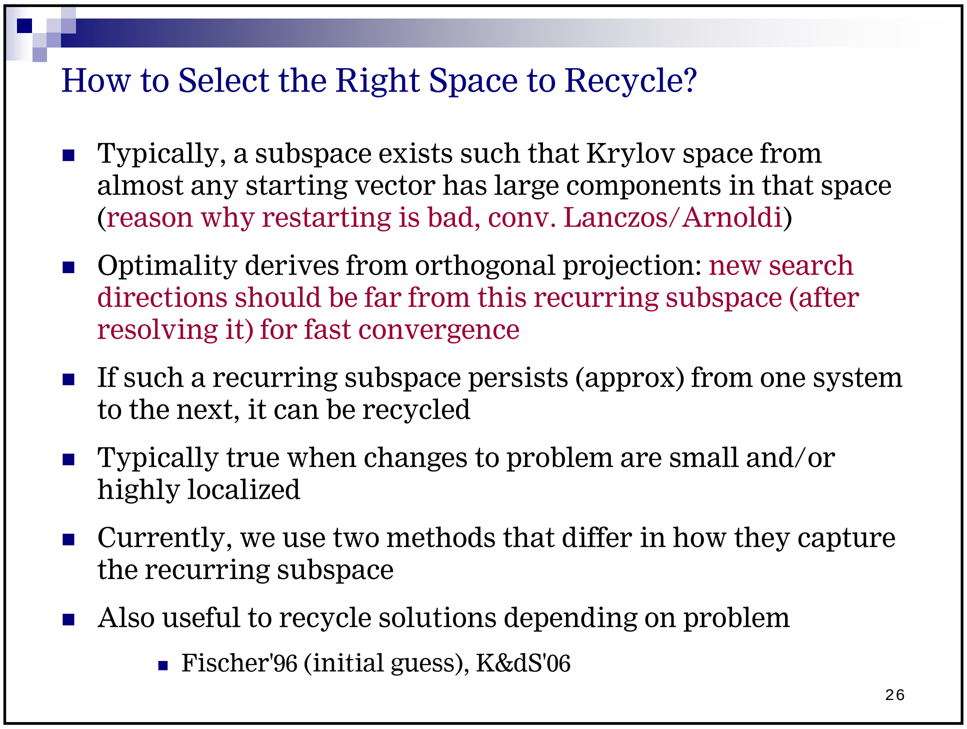
Solving Sequences of Linear Systems

- Computational problems often involve a sequence of systems with small or localized changes in space or structure
  - □ Evolutionary problems, nonlinear problems and optimization, parameter estimation, Monte Carlo and Markov Chain Monte Carlo methods
  - □ Adaptive discretizations and representations
  - □ Accurate simulation requires solution of hundreds to thousands of large, sparse, linear systems
  - □ Crack propagation, topology optimization, tomography, uncertainty quantification
- Fast solution by exploiting the slowly changing nature of problem or special structural changes
- Recycle (adapt & reuse) search spaces from previous problems
- Recycle preconditioners (especially AMR)



Recycling Subspaces for Krylov Methods

- Iterative (Krylov) methods build search space and compute solution by projection
- Building search space often dominates cost
- Initial convergence often poor, reasonable size search space needed, then superlinear convergence
- Get fast convergence rate and good initial guess immediately by recycling selected search spaces from previous systems
- What is right subspace to recycle? How to recycle space?
- For quasi-Newton methods alternative possibility
  - □ Project Jacobian update onto Krylov space, no iterations needed, solve with updated Hessenberg matrix
    - Klie&Wheeler
  - □ Combination of recycling strategies (Klie&dS)



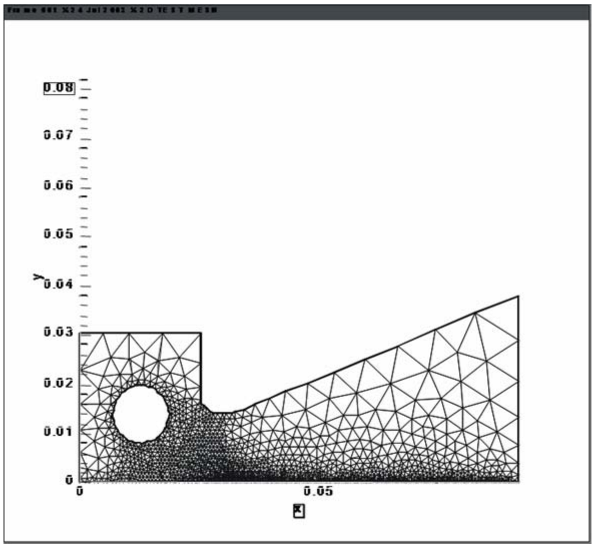
How to Select the Right Space to Recycle?

- Typically, a subspace exists such that Krylov space from almost any starting vector has large components in that space (reason why restarting is bad, conv. Lanczos/Arnoldi)
- Optimality derives from orthogonal projection: new search directions should be far from this recurring subspace (after resolving it) for fast convergence
- If such a recurring subspace persists (approx) from one system to the next, it can be recycled
- Typically true when changes to problem are small and/or highly localized
- Currently, we use two methods that differ in how they capture the recurring subspace
- Also useful to recycle solutions depending on problem
    - Fischer'96 (initial guess), K&dS'06

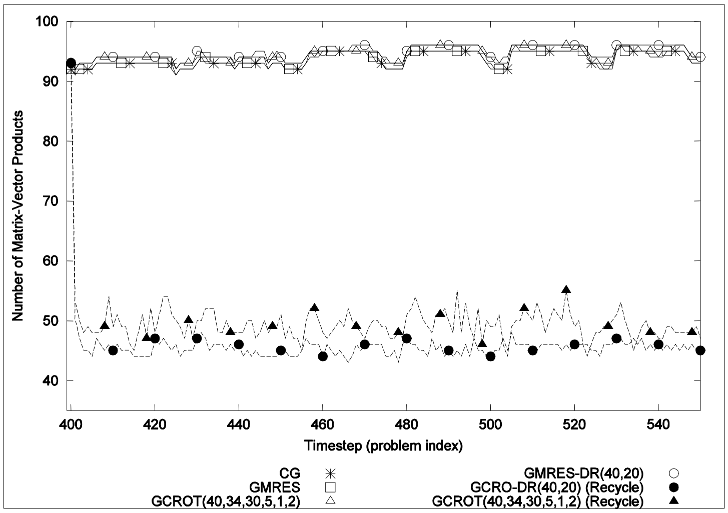
How to Select the Right Space to Recycle?

- Compute canonical angles between successive spaces – measure recurring subspace
  - □ GCROT (dS'99), with recycling (P&dS'06)
  - □ Recycle recurring subspace
    - in current and subseq. systems – non-Hermitian systems
    - subsequent systems – Hermitian systems
- Invariant subspace from small eigenvalues (or large, or both)
  - □ GMRESDR (Morgan'04, '95), with recycling: GCRODR (P&dS'06), RMINRES (WdS&P'07)
- Subspace from previous solutions
  - □ for initial guess (Fischer'96),
  - □ for recycling/combined with invariant subspace (K&dS'06)
- Keep previous space & constrain Jacobian update in quasi-Newton methods
  - □ Klie&Wheeler'05
  - □ Combined with recycling (Klie&dS in prep)

Example: Crack Propagation



Results for Crack Propagation



Fast Solution of Sequences: Issues

- How do relevant subspaces change under changes in the matrix? (invariant subspaces, solution subspaces, …)
- Analysis of convergence of recycle method when keeping approximate (invariant) subspace
- Use application details to tune the recycling (subspace)
- Type of matrix update, problem and algorithm dependent
  - timesteps, rank-k update in quasi-Newton method, localized nonlinear behavior (crack propagation), line search
- Nature of PDE and changes in parameters
- Varying behavior over time: certain modes stationary while others still change (does the method learn?)
- Perturbation of (e.g.) invariant subspaces under specific changes in the matrix
- Multiple parameterized matrices and right hand sides

Example: Tomography

- Reconstruct medium by measuring how signals propagate
- Parameterize medium and optimize parameters by matching measured signal with computed signal (at receivers)
- Forward problem $-\nabla\bullet\left(a(x,\omega;p)\nabla u_j\right) + m(x,\omega;p)\,u = f_j$
- Have to solve (forward) problem many times (optimization)
- Problem is Hermitian for zero frequency and nonzero frequency gives imaginary shift
- Multiple sources give multiple right hand sides
- Nonlinear least squares/Gauss-Newton with line search
- First few steps fix background parameters, later steps mainly change shape of tumor: ‘diffusion’ jump in small region
- Change in matrix concentrated in high frequency modes
- Lot of opportunity to exploit structure

Perturbation of Invariant Subspace

Consider $A = \begin{pmatrix} 1 & 0 \\ 0 & 1+\varepsilon_1 \end{pmatrix} \quad \to \quad x_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ x_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$

Consider perturbation E with $\|E\| = \varepsilon_1 + \varepsilon_2$, ε_2 arb. small.

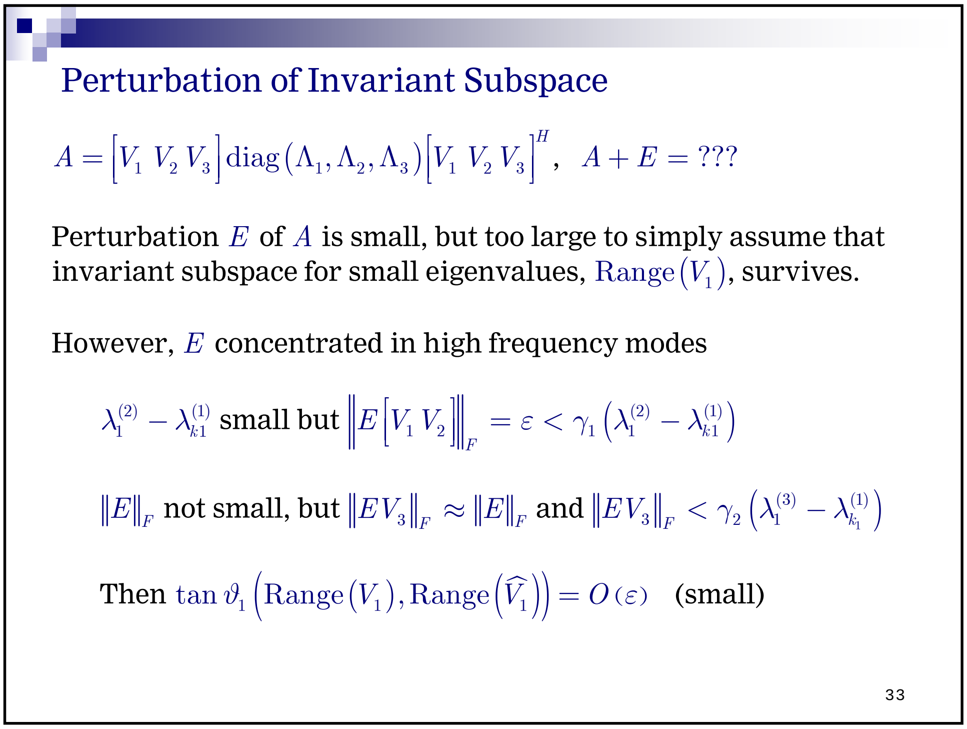
Enough to give A *any eigenvectors* (different from x_1 and x_2).

Let $E = E_1 + E_2$ and $\hat{X} = \left[\hat{x}_1\ \hat{x}_2\right]$ (unitary)

Let $E_1 = \begin{pmatrix} 0 & 0 \\ 0 & -\varepsilon_1 \end{pmatrix}$ and $E_2 = \hat{X}\begin{pmatrix} 0 & 0 \\ 0 & \varepsilon_2 \end{pmatrix}\hat{X}^*$.

$A + E_1 = I$ (all nonzero vectors are eigenvectors)

$$A + E = I + \hat{X}\begin{pmatrix} 0 & 0 \\ 0 & \varepsilon_2 \end{pmatrix}\hat{X}^* = \hat{X}\begin{pmatrix} 1 & 0 \\ 0 & 1+\varepsilon_2 \end{pmatrix}\hat{X}^*$$



Perturbation of Invariant Subspace

$$A = \left[V_1 \; V_2 \; V_3\right] \operatorname{diag}\left(\Lambda_1, \Lambda_2, \Lambda_3\right)\left[V_1 \; V_2 \; V_3\right]^H, \quad A + E = ???$$

Perturbation E of A is small, but too large to simply assume that invariant subspace for small eigenvalues, $\operatorname{Range}\left(V_1\right)$, survives.

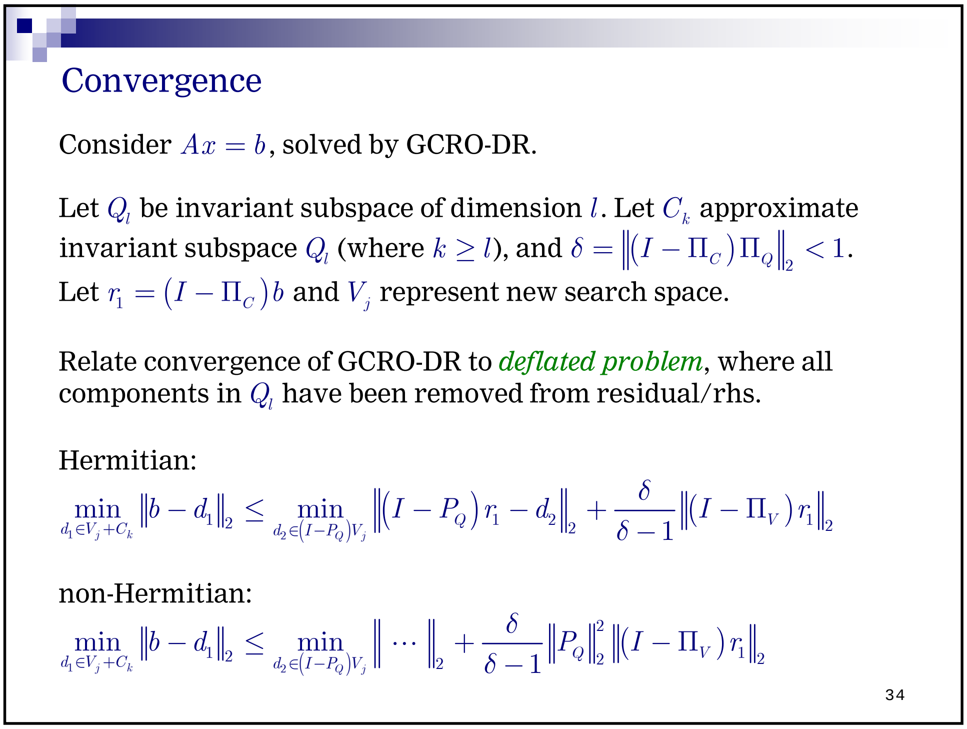
However, E concentrated in high frequency modes

$\lambda_1^{(2)} - \lambda_{k1}^{(1)}$ small but $\left\| E \left[V_1 \; V_2\right] \right\|_F = \varepsilon < \gamma_1 \left(\lambda_1^{(2)} - \lambda_{k1}^{(1)}\right)$

$\|E\|_F$ not small, but $\|EV_3\|_F \approx \|E\|_F$ and $\|EV_3\|_F < \gamma_2 \left(\lambda_1^{(3)} - \lambda_{k_1}^{(1)}\right)$

Then $\tan \vartheta_1 \left(\operatorname{Range}\left(V_1\right), \operatorname{Range}\left(\widehat{V}_1\right)\right) = O\left(\varepsilon\right)$ (small)

33



Convergence

Consider $Ax = b$, solved by GCRO-DR.

Let Q_l be invariant subspace of dimension l. Let C_k approximate invariant subspace Q_l (where $k \geq l$), and $\delta = \left\|\left(I - \Pi_C\right)\Pi_Q\right\|_2 < 1$. Let $r_1 = \left(I - \Pi_C\right)b$ and V_j represent new search space.

Relate convergence of GCRO-DR to *deflated problem*, where all components in Q_l have been removed from residual/rhs.

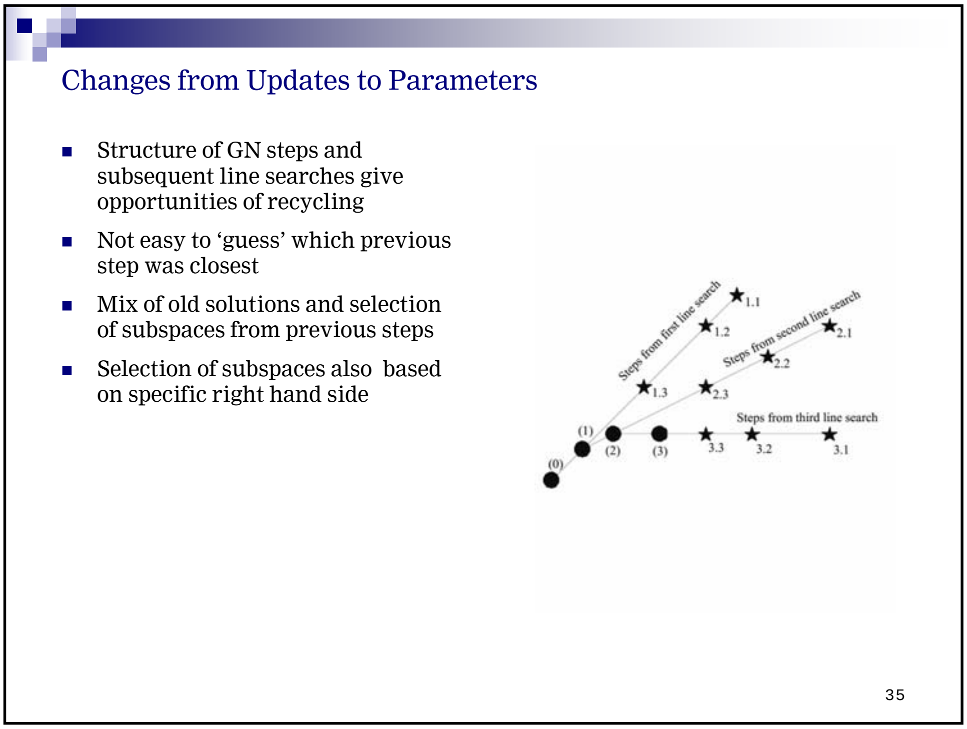
Hermitian:

$$\min_{d_1 \in V_j + C_k} \|b - d_1\|_2 \leq \min_{d_2 \in \left(I - P_Q\right)V_j} \left\|\left(I - P_Q\right) r_1 - d_2\right\|_2 + \frac{\delta}{\delta - 1}\left\|\left(I - \Pi_V\right) r_1\right\|_2$$

non-Hermitian:

$$\min_{d_1 \in V_j + C_k} \|b - d_1\|_2 \leq \min_{d_2 \in \left(I - P_Q\right)V_j} \left\| \cdots \right\|_2 + \frac{\delta}{\delta - 1}\|P_Q\|_2^2 \left\|\left(I - \Pi_V\right) r_1\right\|_2$$

34

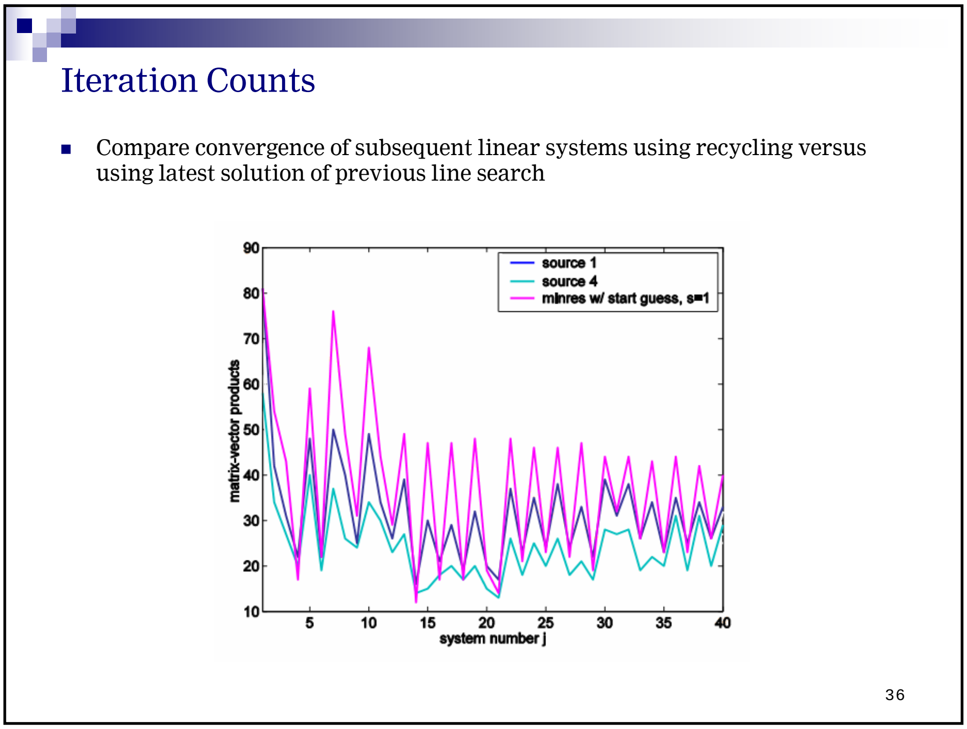


Changes from Updates to Parameters

- Structure of GN steps and subsequent line searches give opportunities of recycling
- Not easy to ‘guess’ which previous step was closest
- Mix of old solutions and selection of subspaces from previous steps
- Selection of subspaces also based on specific right hand side

Steps from first line search
1.1
1.2
1.3
Steps from second line search
2.1
2.2
2.3
Steps from third line search
3.3
3.2
3.1
(0)
(1)
(2)
(3)

35

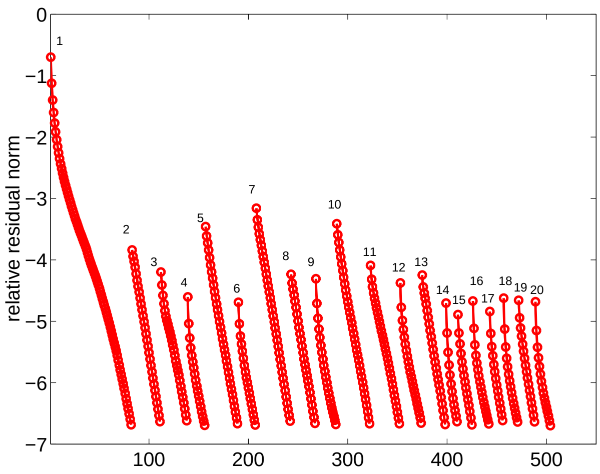


Iteration Counts

- Compare convergence of subsequent linear systems using recycling versus using latest solution of previous line search

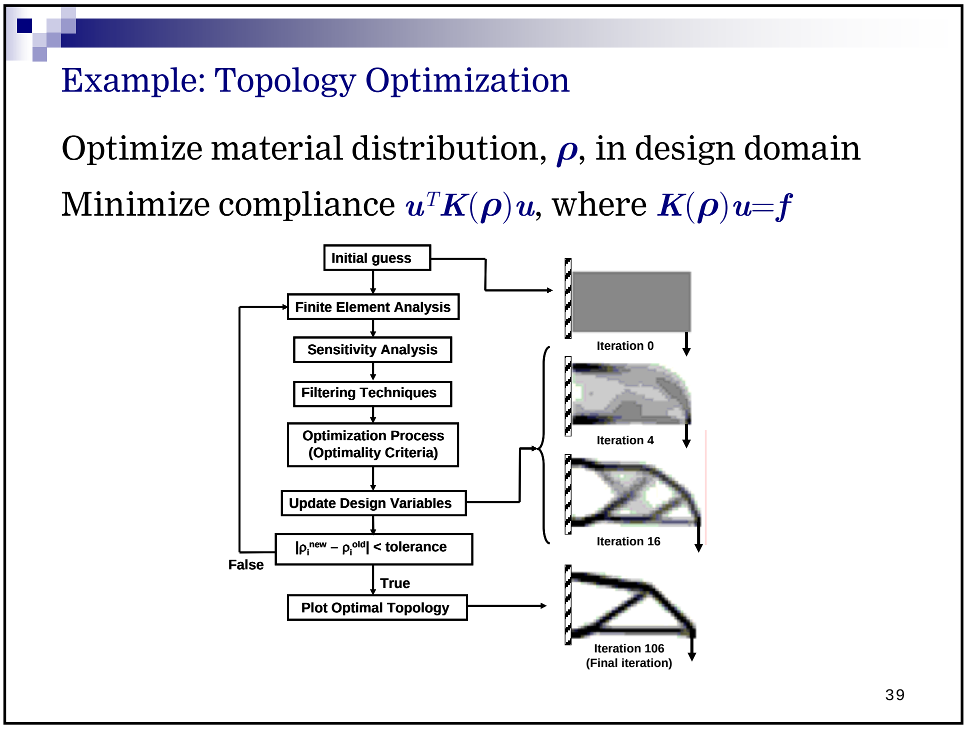
36

Iteration Counts



Conclusions and Future Work

- Recycling Krylov spaces can greatly reduce iteration counts for sequence of problems
- Significant opportunity to tune recyling for specific algorithms and applications (with benefits)
- Further convergence analysis (GCROT)
- Refined perturbation analysis
- Applications: Tomography (with Misha Kilmer), QCD, electronic structure, topology optimization (Paulino), PDE constrained optimization
- Nonlinear and optimization algorithms: quasi-Newton methods (updated ‘Hessian’ or ‘Jacobian’)
- Probably interesting links with limited memory matrix methods



Example: Topology Optimization

Optimize material distribution, $\boldsymbol{\rho}$, in design domain

Minimize compliance $\boldsymbol{u}^T\boldsymbol{K}(\boldsymbol{\rho})\boldsymbol{u}$, where $\boldsymbol{K}(\boldsymbol{\rho})\boldsymbol{u}=\boldsymbol{f}$

Initial guess

Finite Element Analysis

Sensitivity Analysis

Filtering Techniques

Optimization Process (Optimality Criteria)

Update Design Variables

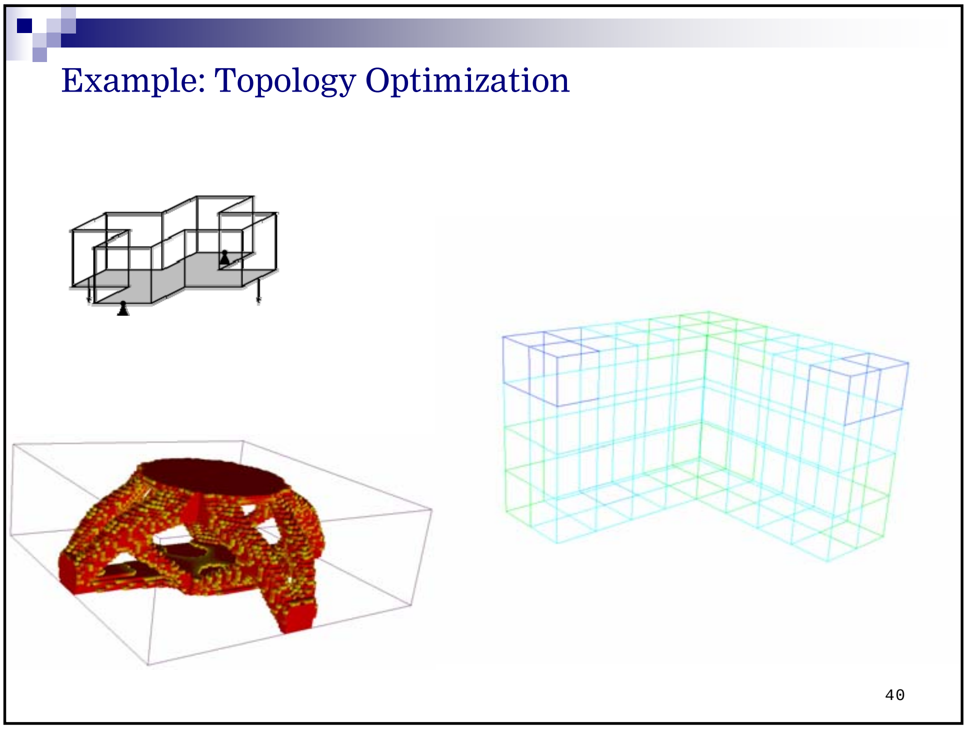
$|\rho_i^{new} - \rho_i^{old}| <$ tolerance

False

True

Plot Optimal Topology

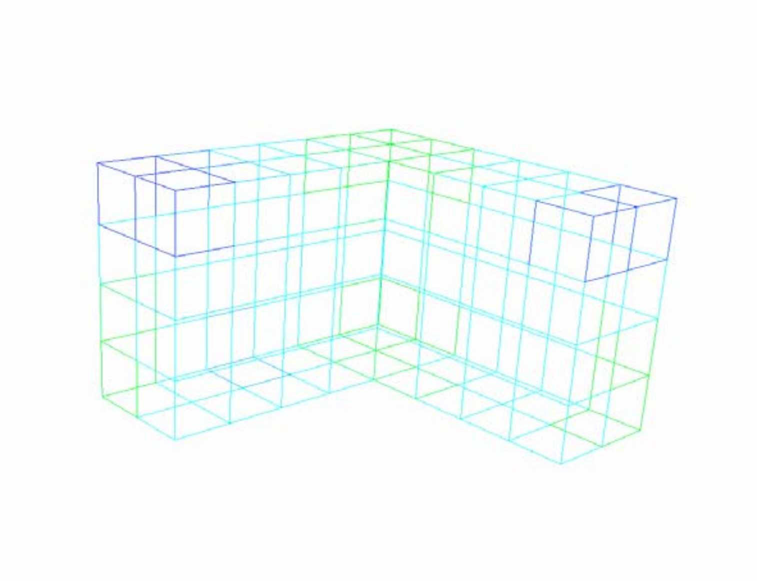
39



Example: Topology Optimization

40

Example: Topology Optimization



How to Use a Recycle Space?

Solve $Ax = b$ using recycled subspace/matrix $\tilde{U}$ (for new A):

Compute $A\tilde{U} = \tilde{C}$, $CR = \tilde{C}$ (QR), $U = \tilde{U}R^{-1}$ (implicit)
Now $AU = C$ and $C^*C = I$

Set $r_0 = \left(I - CC^*\right)b$, $x_0 = UC^*b$, and $v_1 = r_0 / \|r_0\|$
Augmented Arnoldi: $AV_m = CC^*AV_m + V_{m+1}\underline{H}_m = CB + V_{m+1}\underline{H}_m$

Minimize $\left\|b - A\left(x_0 + Uz + V_m y\right)\right\| = \left\|r_0 - Cz - CBy - V_{m+1}\underline{H}_m y\right\| =$

$$\left\|V_{m+1}\left(e_1\|r_0\| - \underline{H}_m y\right) - C\left(z + By\right)\right\|$$

Solve $\underline{H}_m y \approx e_1\|r_0\|$ and set $z = -By$
$x_m = x_0 + Uz + V_m y$ and $r_m = V_{m+1}\left(e_1\|r_0\| - \underline{H}_m y\right)$ (GCRO, dS'95)

Recycling for Hermitian systems

- No need to recycle for single system (theory)
  - □ Recycle with fixed space for given system
  - □ Update recycle space for next system
- Use short recurrence and discard unnecessary vectors
- Complications; how to efficiently update recycle space
  - □ Improve using Lanczos vectors periodically (discard)
  - □ Multiple possible combinations (3 spaces)
  - □ Using only Lanczos vectors not a good idea as they may be nearly orthogonal to good eigenvectors in recycle space
  - □ We update/merge 'new' recycle space with Lanczos vectors
  - □ Other options should be evaluated
- Efficient implementation of invariant subspace computation

Updating the Recycle Subspace

Each m its: $(I - CC^*)A(I - CC^*)V_j = \underline{\overline{V}}_j \underline{\overline{T}}_j$

Update: $U_0 = U$ (recycle space) and $[U_{j-1}\ V_j] \rightarrow U_j$

$$A[U_{j-1}\ V_j] = [C\ C_{j-1} V_j]\begin{bmatrix} 0 & B_j \\ I & 0 \\ 0 & \underline{\overline{T}}_j \end{bmatrix} \rightarrow A\,W = \tilde{W}\underline{\tilde{H}}$$

Exploit orthog. relations to compute (harm) Ritz vectors efficiently

Solve $\underline{\tilde{H}}^* \tilde{W}^* \tilde{W} \underline{\tilde{H}} P = \underline{\tilde{H}}^* \tilde{W}^* W P \Theta$

Let $\tilde{U}_j = [U_{j-1}\ V_j] P$; then $A\,\tilde{U}_j = \tilde{W}\tilde{H}P$

Let $\tilde{W} \overset{\text{QR}}{=} \hat{W}K$ and $K\tilde{H}P \overset{\text{QR}}{=} QR$, and set $C_j = \hat{W}Q$ and $U_j = \tilde{U}_j R^{-1}$

Short Term Recurrences for RMINRES

Full iteration $\left(I - CC^*\right)A\left(I - CC^*\right)V_m = V_{m+1}\underline{T}_m$

Let $\underline{T}_m = \underline{G}_m F_m$ and $y_m = F_m^{-1}\underline{G}_m^* e_1 \|r_0\|$

Then $x_m = UC^*b - UBy_m + V_m y_m$ and $r_m = r_0 - CC^*r_0 - V_{m+1}\underline{T}_m y_m$

Let $\tilde{B} = UBF_m^{-1}$, $\tilde{V}_m = V_m F_m^{-1}$, and $\tilde{y}_m = F_m y_m = \underline{G}_m^* e_1 \|r_0\| = G_{(m)}^* \tilde{y}_{m-1}$

Then

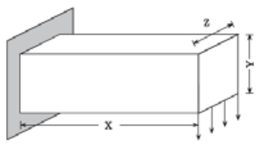
$$\tilde{v}_{m-2} f_{m-2,m} + \tilde{v}_{m-1} f_{m-1,m} + \tilde{v}_m f_{m,m} = v_m$$

$$\tilde{b}_{m-2} f_{m-2,m} + \tilde{b}_{m-1} f_{m-1,m} + \tilde{b}_m f_{m,m} = UC^*Av_m$$

$$x_m = x_0 + UC^*r_0 - \tilde{B}_m \tilde{y}_m + \tilde{V}_m \tilde{y}_m = x_{m-1} - \tilde{b}_m \tilde{y}_{m,m} + \tilde{v}_m \tilde{y}_{m,m}$$

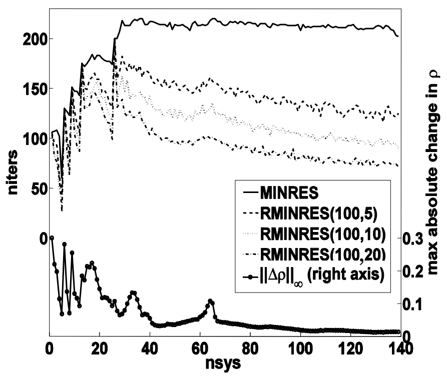
And we can drop all vectors from recurrence except the last two

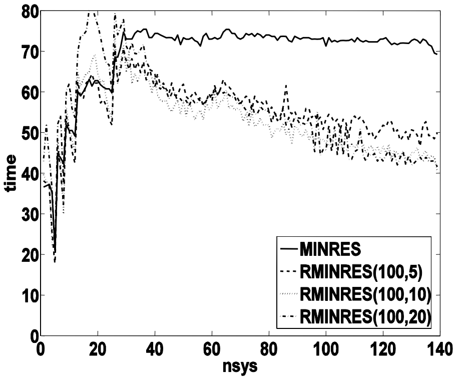
Topology Opt. Convergence results

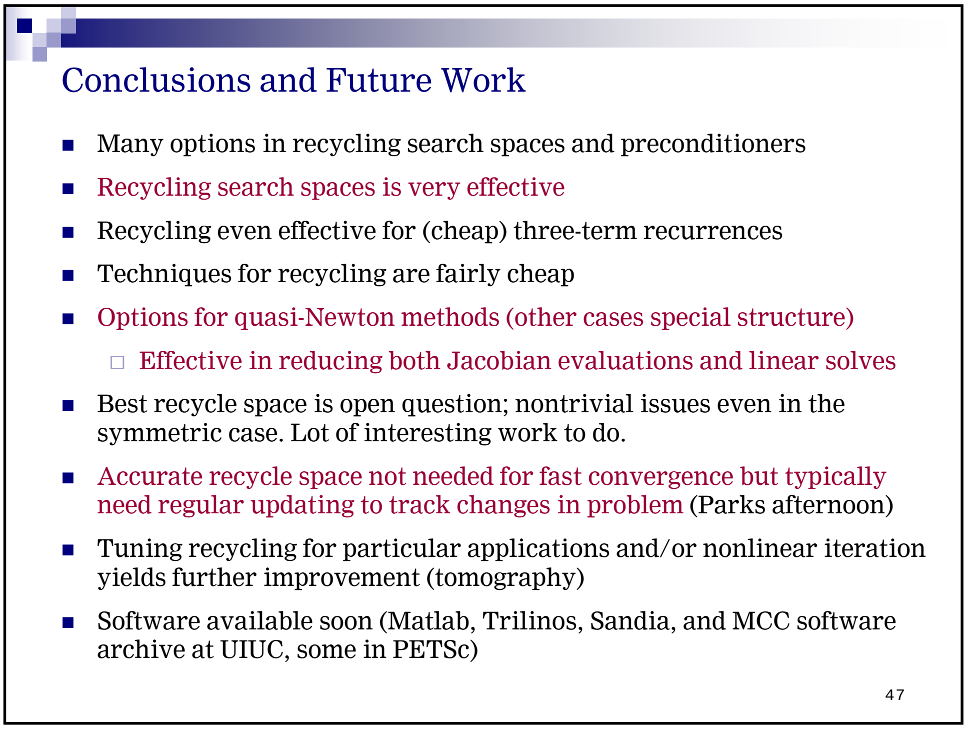


Results: 84x28x14 mesh (107K dofs) and 180x60x30 (1M dofs) (on PC)

Currently more complicated models up to 2M dof (on PC)

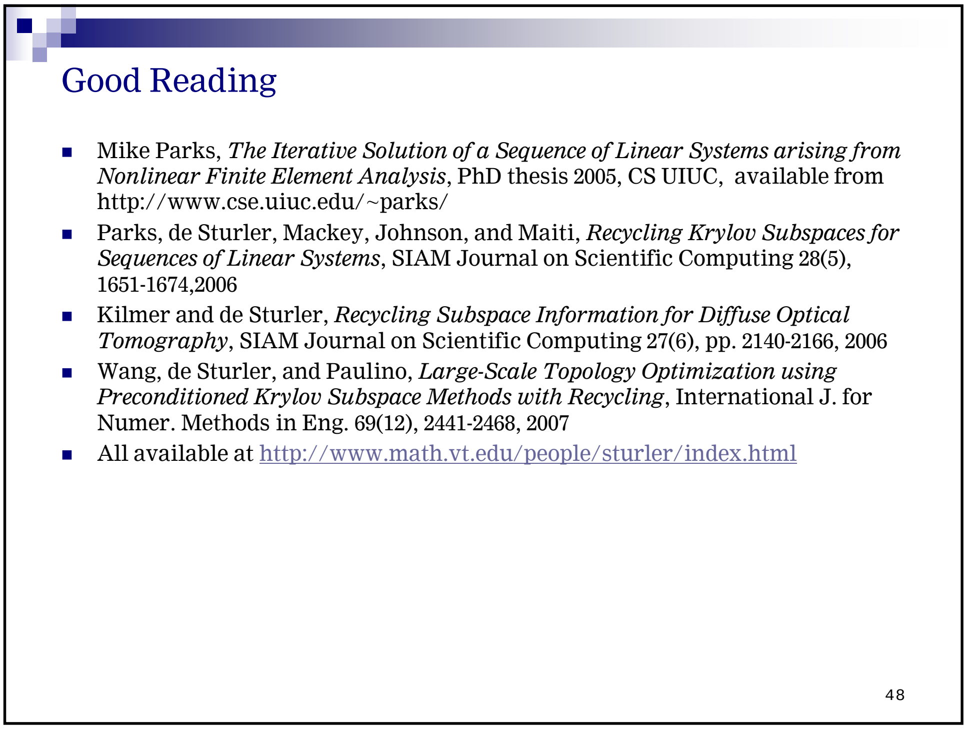






Conclusions and Future Work

- Many options in recycling search spaces and preconditioners
- Recycling search spaces is very effective
- Recycling even effective for (cheap) three-term recurrences
- Techniques for recycling are fairly cheap
- Options for quasi-Newton methods (other cases special structure)
  - Effective in reducing both Jacobian evaluations and linear solves
- Best recycle space is open question; nontrivial issues even in the symmetric case. Lot of interesting work to do.
- Accurate recycle space not needed for fast convergence but typically need regular updating to track changes in problem (Parks afternoon)
- Tuning recycling for particular applications and/or nonlinear iteration yields further improvement (tomography)
- Software available soon (Matlab, Trilinos, Sandia, and MCC software archive at UIUC, some in PETSc)



Good Reading

- Mike Parks, *The Iterative Solution of a Sequence of Linear Systems arising from Nonlinear Finite Element Analysis*, PhD thesis 2005, CS UIUC, available from http://www.cse.uiuc.edu/~parks/
- Parks, de Sturler, Mackey, Johnson, and Maiti, *Recycling Krylov Subspaces for Sequences of Linear Systems*, SIAM Journal on Scientific Computing 28(5), 1651-1674,2006
- Kilmer and de Sturler, *Recycling Subspace Information for Diffuse Optical Tomography*, SIAM Journal on Scientific Computing 27(6), pp. 2140-2166, 2006
- Wang, de Sturler, and Paulino, *Large-Scale Topology Optimization using Preconditioned Krylov Subspace Methods with Recycling*, International J. for Numer. Methods in Eng. 69(12), 2441-2468, 2007
- All available at http://www.math.vt.edu/people/sturler/index.html