Comparing iterative methods to compute the overlap Dirac operator at nonzero chemical potential

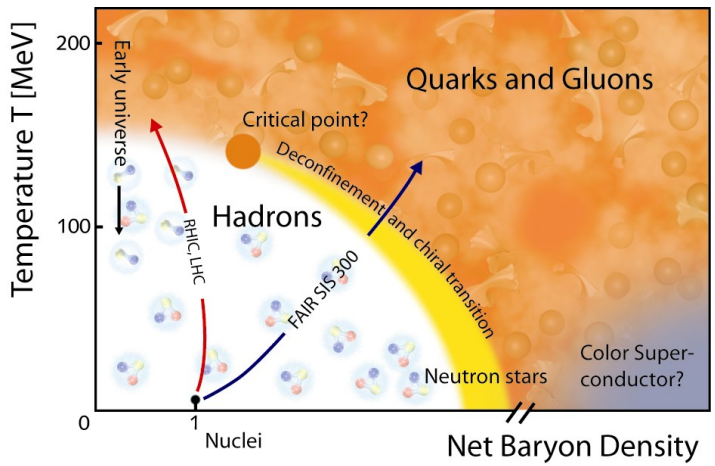
Jacques Bloch, Tobias Breu, and Tilo Wettig

Institute for Theoretical Physics, University of Regensburg

Lattice 2008, Williamsburg, Virginia, USA, 14-19 July 2008

Outline

1. Lattice implementation of chiral symmetry at nonzero quark density
   - Overlap Dirac operator at nonzero quark chemical potential
   - Sign function of non-Hermitian matrix
2. Iterative methods for function of non-Hermitian matrix
   - Arnoldi approximation
   - Two-sided Lanczos approximation
   - Deflation
   - Results



Chiral symmetry on the lattice

- chiral symmetry on the lattice – Ginsparg-Wilson relation:

$$\{D, \gamma_5\} = aD\gamma_5 D$$

- massless overlap Dirac-operator (Neuberger-Narayanan)

$$D_{\text{ov}} = \mathbb{1} + \gamma_5 \,\text{sign}(\gamma_5 D_W)$$

- D_{ov} satisfies GWR because $\text{sign}^2(A) = \mathbb{1}$
- kernel $\gamma_5 D_W$ Hermitian $\rightarrow \gamma_5 D_{\text{ov}} = D_{\text{ov}}^\dagger \gamma_5$ (γ_5-Hermiticity)
- D_{ov} has exact zero modes with definite chirality $\langle \gamma_5 \rangle = \pm 1$ reflecting topological charge of gauge configuration (Atiyah-Singer index theorem)

Chiral symmetry on the lattice at nonzero quark density

Generalize overlap Dirac operator to nonzero quark chemical potential

- replace D_W by $D_W(\mu)$ in overlap definition:

Overlap operator at $\mu \neq 0$

$$D_{\text{ov}}(\mu) = \mathbb{1} + \gamma_5 \,\text{sign}(\gamma_5 D_W(\mu))$$

JB, Wettig PRL97(012003) 2006

Wilson-Dirac operator at $\mu \neq 0$

$$D_W(\mu) = 1 - \kappa \sum_{i=1}^{3} (T_i^+ + T_i^-) - \kappa (e^{\mu} T_4^+ + e^{-\mu} T_4^-)$$

with $\quad (T_\nu^\pm)_{yx} = (1 \pm \gamma_\nu) U_{x,\pm\nu} \delta_{y,x\pm\hat{\nu}}$

Hasenfratz-Karsch 1983, Kogut et al. 1983

- kernel $\gamma_5 D_W(\mu)$ no longer Hermitian:

$D_{\text{ov}}(\mu)$ requires definition of sign of a non-Hermitian matrix

Function of a matrix

- spectral definition of matrix function:
  - if A diagonalizable: $A = U\,\mathrm{diag}\{\lambda_i\}U^{-1}$

$$f(A) = U\,\mathrm{diag}\{f(\lambda_i)\}U^{-1}$$

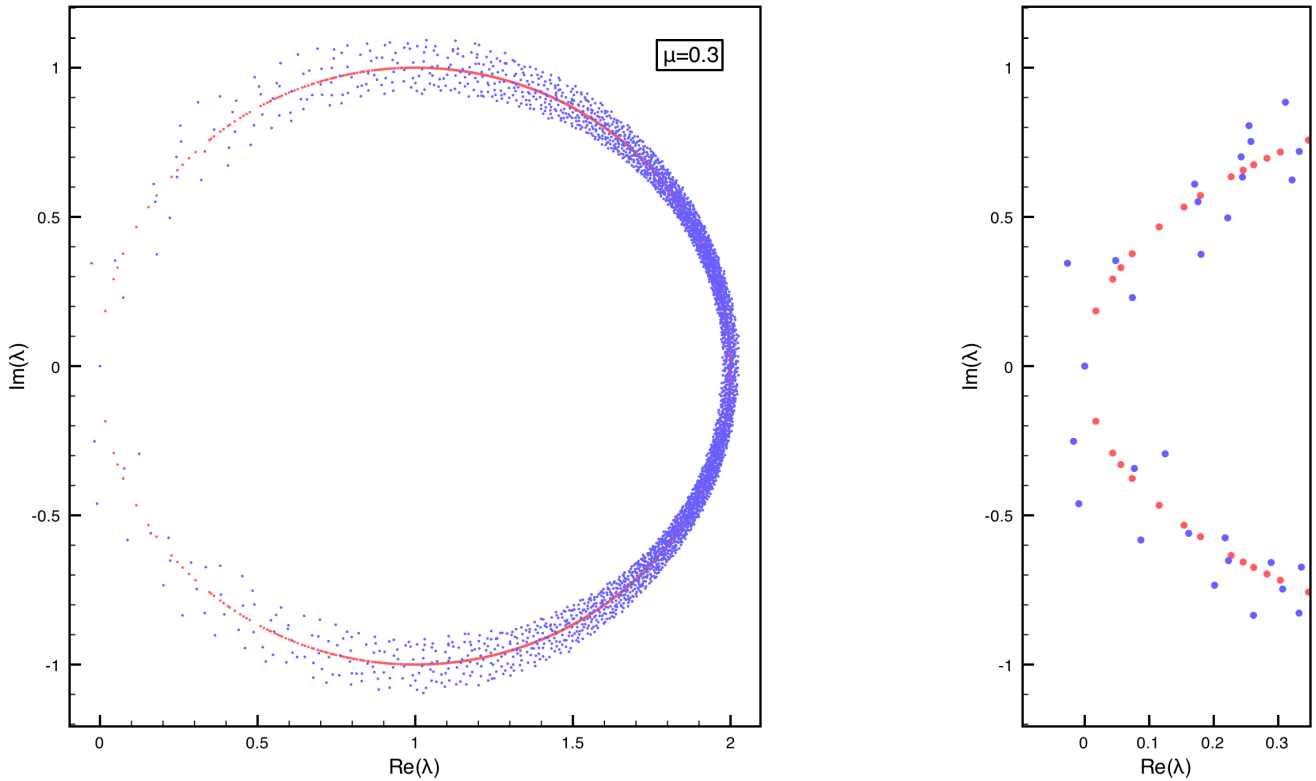
    with complex eigenvalues $\lambda_1, \ldots, \lambda_N$ and eigenvector matrix U
  - if A not diagonalizable: spectral definition using Jordan canonical form
- Sign function of non-Hermitian matrix requires sign of complex number:

$$\mathrm{sign}(z) = \frac{z}{\sqrt{z^2}} = \mathrm{sign}(\mathrm{Re}\, z)$$

  - ensures $\mathrm{sign}^2(z) = 1$
  - gives correct result for $z \in \mathbb{R}$
- definition ensures $\mathrm{sign}^2(A) = \mathbb{1}$

Typical spectrum ($V = 4^4, \beta = 5.1, m_W = -2$)

$$\mu = 0.3$$



- $D_{\text{ov}}(\mu)$ satisfies Ginsparg-Wilson relation → lattice chiral symmetry
- exact zero modes with definite chirality
- naturally violates γ_5-Hermiticity → spectrum no longer on circle

Iterative method for function of non-Hermitian matrix

- exact computation of $\mathrm{sign}(A)$ only possible for small volumes
  - memory requirements (store full matrix)
  - computation time (compute full diagonalization)
- develop iterative method to compute $f(A)b$ for non-Hermitian A
- exact statement: for the unique polynomial $P_K(z)$ which interpolates $f(z)$ at all eigenvalues of A,

$$P_K(A)b = f(A)b \quad \text{for any vector } b$$

- approximation method for $y = f(A)b$:

  construct good low degree polynomial approximation to f on $\lambda(A)$ wrt b
  - depends on spectrum of A
  - depends on decomposition of b in eigenvectors of A

Constructing an Arnoldi basis

- Krylov subspace: $K_k(A, b) = \text{span}(b, Ab, A^2 b, \ldots, A^{k-1} b)$.
  - contains all vectors resulting from action of arbitrary polynomial of degree $\leq k-1$ in A on vector b
  - one of these vectors minimizes $||P_{k-1}(A)b - f(A)b||$ over all polynomials of degree $\leq k-1 \rightarrow$ namely, the projection of $f(A)b$ on the Krylov subspace

- Arnoldi method uses the recursive scheme

  with
  $$AV_k = V_k H_k + \beta_k v_{k+1} e_k^T$$

  $$V_k^\dagger A V_k = H_k$$

  to build an orthonormal basis $V_k = (v_1, \ldots, v_k)$ in $K_k(A, b)$, where:
  - H_k is a $k \times k$ Hessenberg matrix (upper triangular + one subdiagonal)
  - eigenvalues of H_k are Ritz values of A w.r.t. $K_k(A, b)$
  - $v_1 = b/|b|$
  - β_k: normalization of v_{k+1}, e_k is the k-th basis vector in $\mathbb{C}^k$.

Arnoldi approximation for function of non-Hermitian matrix

- projection of $y = f(A)b$ on $K_k(A, b)$:

$$y \approx y_{\text{proj}} = V_k V_k^\dagger f(A) b = V_k \underbrace{V_k^\dagger f(A) V_k}_{\approx f(H_k) \rightarrow \text{Ritz approximation}} V_k^\dagger b$$

- approximation to y_{proj} using $V_k^\dagger f(A) V_k \approx f(H_k)$

$$y_{\text{proj}} \approx \tilde{y} = |b| V_k f(H_k) e_1$$

  - $\tilde{y} \in K_k(A, b)$
  - $f(x)$ interpolated at Ritz values of A wrt $K_k(A, b)$
  - problem reduced to computation of $f(H_k)$ ($\dim H_k \ll \dim A$)

- $f(H_k)$ computed with suitable method
  - exactly with spectral decomposition
  - suitable approximation method
    e.g., for sign function use Roberts' matrix-iterative method:

$$S^{n+1} = \frac{1}{2}\left[S^n + (S^n)^{-1}\right] \quad , \quad \text{with} \quad S^0 = A$$

Sign function and deflation – Hermitian case

- problem: need large Krylov space if A has small eigenvalues
- reason: in region of $\mathbb{C}$ where f changes rapidly
  $\rightarrow$ hard to approximate f by low-degree polynomial
- solution: improve efficiency by using exact value of f for critical eigenvalues of A
- Hermitian case: deflation straightforward because any # eigenvectors form subspace orthonormal to remaining eigenvectors:

$$f(A)b = Uf(\Lambda)U^\dagger b = \sum_{i=1}^{m} \underbrace{f(\lambda_i)(u_i^\dagger b)\,u_i}_{\text{exact}} + \underbrace{f(A)b_\perp}_{\text{approximation}}$$

  - u_i eigenvector corresp. to λ_i, and $b_\perp = b - \sum_{i=1}^{m}(u_i^\dagger b)u_i$
  - compute eigenvalues and eigenvectors needed for deflation once $\forall\, b$
  - approximation for $f(A)b_\perp$ in space $\perp \operatorname{span}(u_1, \ldots, u_m)$
- simple decomposition does not work in the non-Hermitian case since eigenvectors of A are not orthonormal

Non-Hermitian case: Left-Right deflation

- use left and right eigenvectors belonging to m critical eigenvectors

$$AR_m = R_m \Lambda_m$$

$$L_m^\dagger A = \Lambda_m L_m^\dagger$$

  - Λ_m is the diagonal eigenvalue matrix for the m critical eigenvalues
  - $R_m = (r_1, \ldots, r_m)$ is the matrix of right eigenvectors
  - $L_m = (\ell_1, \ldots, \ell_m)$ is the matrix containing the left eigenvectors
  - $L_m^\dagger R_m = I_m$, and $R_m L_m^\dagger$ is oblique projector on the subspace Ω_m

- decompose b as

$$b = b_\parallel + b_\ominus$$

where $b_\parallel = R_m L_m^\dagger b$ is oblique projection of b on Ω_m and $b_\ominus = b - b_\parallel$

$$f(A)b = f(A)R_m L_m^\dagger b + f(A)b_\ominus$$

Left-Right deflation – the approximation

- 1st term: exact contribution

$$f(A)R_m L_m^\dagger b = R_m f(\Lambda_m) L_m^\dagger b$$

- 2nd term: Arnoldi method in the Krylov subspace $K_k(A, b_\ominus)$

$$AV_k = V_k H_k + \beta_k v_{k+1} e_k^T$$

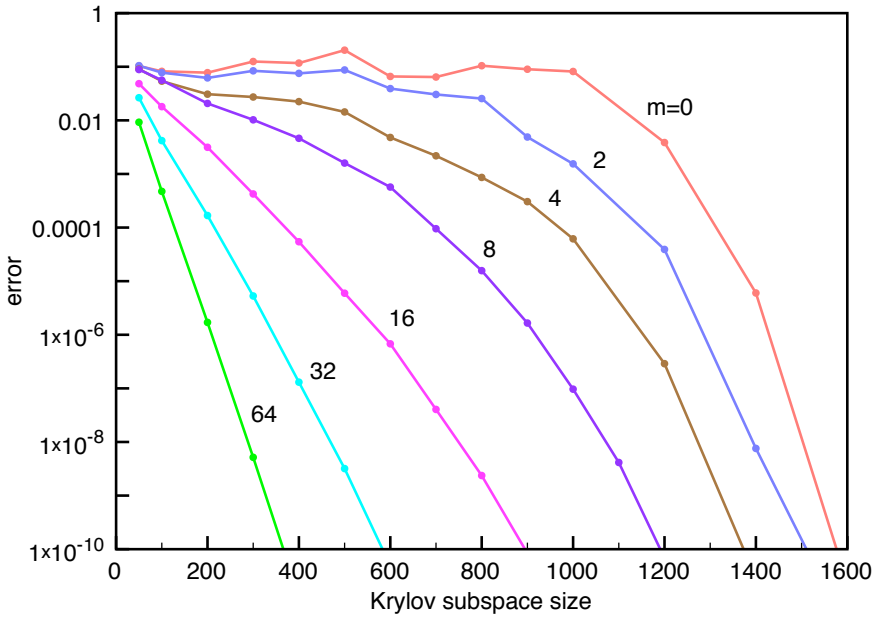
- Finally,

$$f(A)b \approx R_m f(\Lambda_m) L_m^\dagger b + |b_\ominus| V_k f(H_k) e_1$$

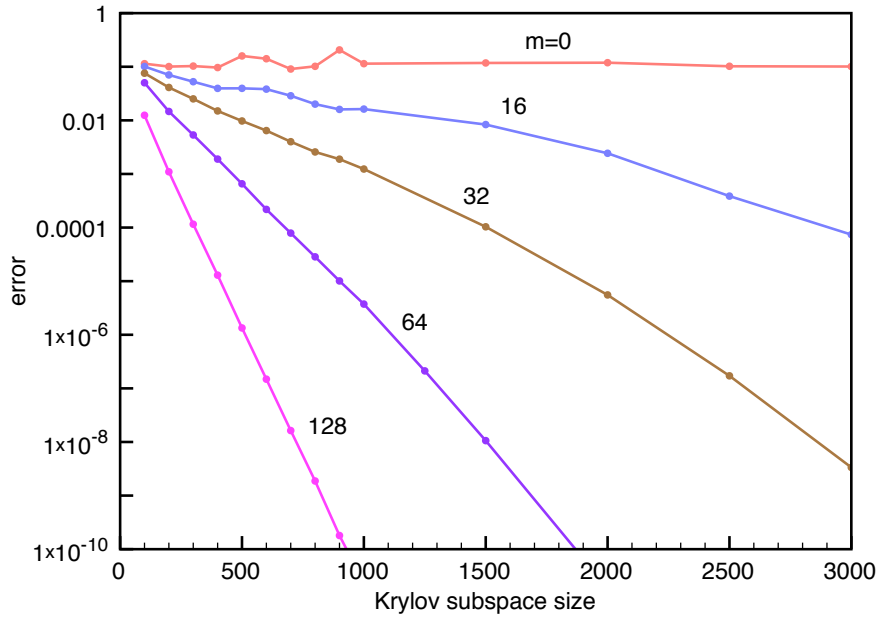
  - compute $f(H_k)$ with suitable method
  - only needs first column of $f(H_k)$
  - requires left and right critical eigenvectors

Deflation and convergence for $D_{\mathrm{ov}}(\mu)b$

4^4 lattice (dim=3072)



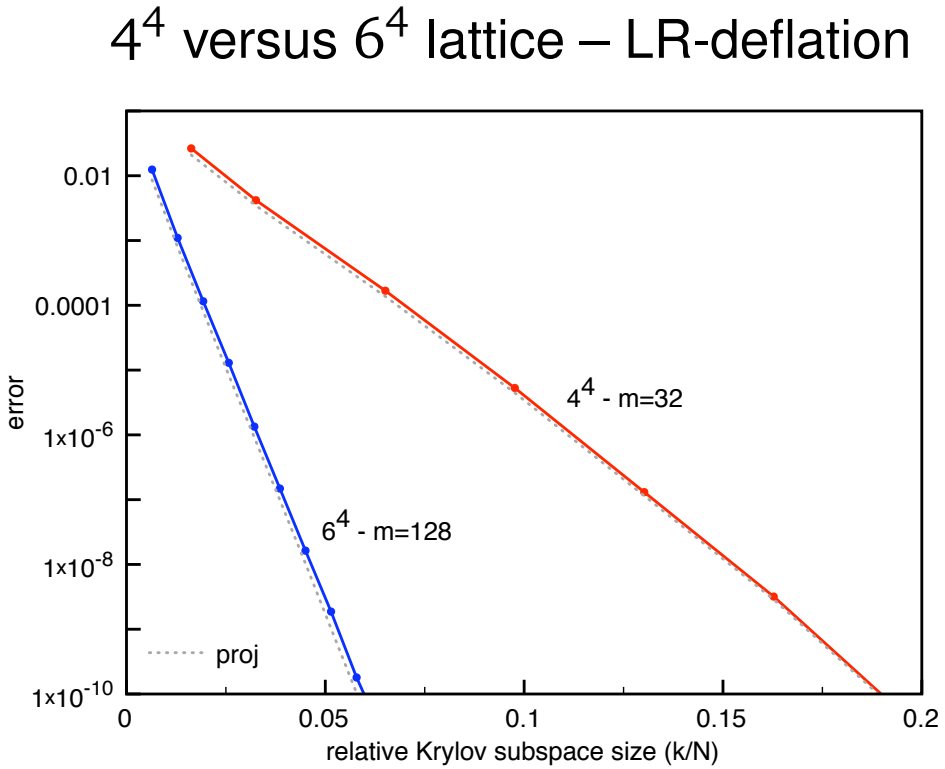
6^4 lattice (dim=15552)



- initialization phase: determine right and left eigenvectors of $\gamma_5 D_{\mathrm{w}}(\mu)$ corresponding to eigenvalues with smallest magnitude using ARPACK
- trade-off between # of deflated eigenvalues and Krylov subspace size

deflation is essential to reach satisfying efficiency

Deflation efficiency for increasing volume



deflation efficiency grows with increasing lattice volume

Two-sided Lanczos method

Arnoldi method suffers from long recurrences

$\longrightarrow$ Two-sided Lanczos: short recurrences but only bi-orthogonal

- Consider two Krylov subspaces $K_k(A, v_1)$ and $K_k(A^\dagger, w_1)$
- Construct bi-orthogonal bases V_k and W_k such that
  - $W_k^\dagger V_k = I_k$
  - $G_k \equiv W_k^\dagger A V_k$ is tridiagonal

$$G_k \equiv W_k^\dagger A V_k = \begin{pmatrix} \alpha_1 & \gamma_1 & 0 & \cdots & 0 \\ \beta_1 & \alpha_2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \gamma_{k-1} \\ 0 & \cdots & 0 & \beta_{k-1} & \alpha_k \end{pmatrix}$$

  - V_k and W_k can be built with short recurrence relations:

$$\begin{cases} \beta_i v_{i+1} = (A - \alpha_i) v_i - \gamma_{i-1} v_{i-1}, \\ \gamma_i^* w_{i+1} = (A^\dagger - \alpha_i^*) w_i - \beta_{i-1}^* w_{i-1}, \end{cases}$$

where

$$\alpha_i = w_i^\dagger A v_i \quad \text{and} \quad \beta_i, \gamma_i \ \text{ from } \ w_{i+1}^\dagger v_{i+1} = 1$$

Two-sided Lanczos approximation + deflation

- $V_k W_k^\dagger$ is oblique projector on $K_k(A, v_1)$
- oblique projection of $f(A)b$ on $K_k(A, b)$:

$$y \approx y_{\text{obl}} = V_k \underbrace{W_k^\dagger f(A) V_k}\, W_k^\dagger b$$

- approximation to y_{obl} using $W_k^\dagger f(A) V_k \approx f(G_k)$

$$y_{\text{obl}} \approx \tilde{y} = |b| V_k f(G_k) e_1$$

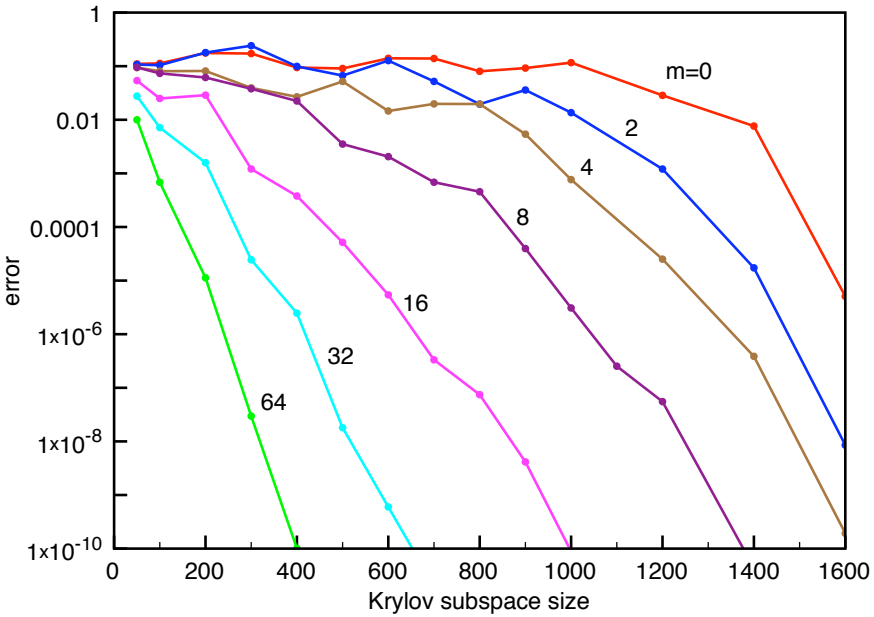
  - $\tilde{y} \in K_k(A, b)$
  - problem reduced to computation of $f(G_k)$ ($\dim G_k \ll \dim A$)

- Enhance with LR-deflation:
  construct bi-orthogonal bases V_k and W_k in $K_k(A, b_\ominus^R)$ and $K_k(A^\dagger, b_\ominus^L)$, where directions along R_m, resp. L_m, have been fully deflated from b: $b_\ominus^R = (1 - R_m L_m^\dagger) b$ and $b_\ominus^L = (1 - L_m R_m^\dagger) b$.
  Function approximation:

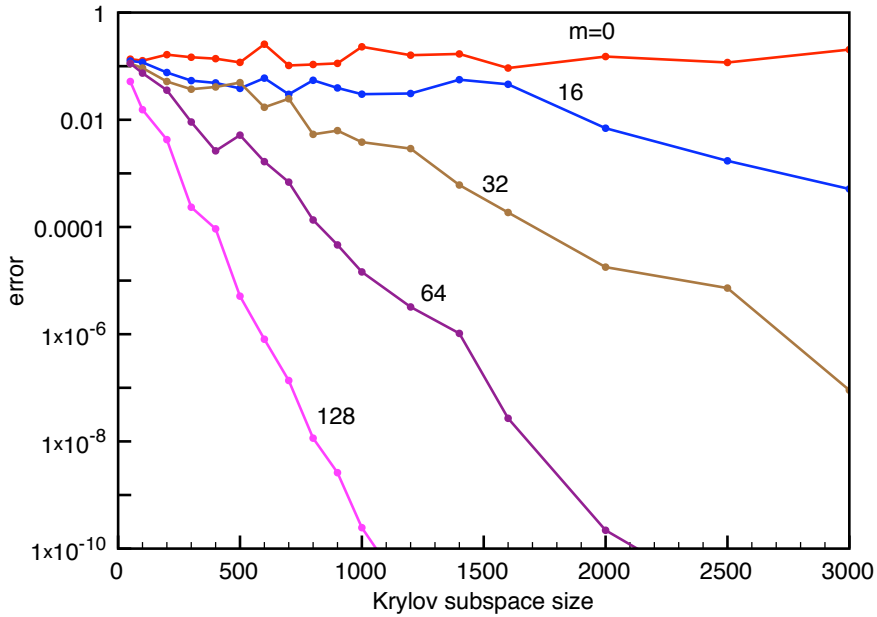
$$f(A)b \approx R_m f(\Lambda_m) L_m^\dagger b + |b_\ominus^R| V_k f(G_k) e_1$$

Two-sided Lanczos – Deflation and convergence for $D_{\mathrm{ov}}(\mu)b$

4^4 lattice (dim=3072)



6^4 lattice (dim=15552)



Arnoldi versus two-sided Lanczos – CPU-time

4^4 lattice (dim=3072)

$m = 32$ – LR-deflation

initialization time: 27.5 s

Arnoldi

k	Arnoldi	$\mathrm{sign}(H_k)$	total
200	0.45	0.20	0.66
400	1.77	1.02	2.82
600	3.94	2.77	6.74
800	6.96	6.44	13.44
1000	10.84	12.33	23.21

two-sided Lanczos

k	2s-Lanczos	$\mathrm{sign}(G_k)$	total
200	0.11	0.19	0.31
400	0.20	0.98	1.20
600	0.31	2.82	3.15
800	0.43	6.52	6.97
1000	0.51	12.45	13.00

6^4 lattice (dim=15552)

$m = 128$ – LR-deflation

initialization time: 1713 s

Arnoldi

k	Arnoldi	$\mathrm{sign}(H_k)$	total
200	2.39	0.15	2.62
400	9.01	0.94	10.06
600	20.03	2.80	22.98
800	35.09	6.49	41.78
1000	54.74	12.36	67.34

two-sided Lanczos

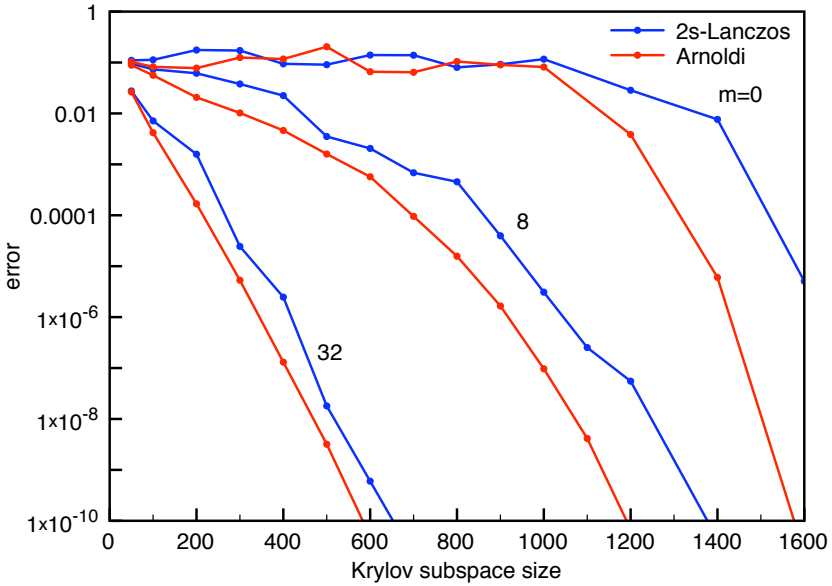
k	2s-Lanczos	$\mathrm{sign}(G_k)$	total
200	0.60	0.19	0.87
400	1.17	0.95	2.24
600	1.72	2.84	4.71
800	2.33	6.55	9.08
1000	3.02	12.45	15.69

Arnoldi basis $\sim Nk^2$

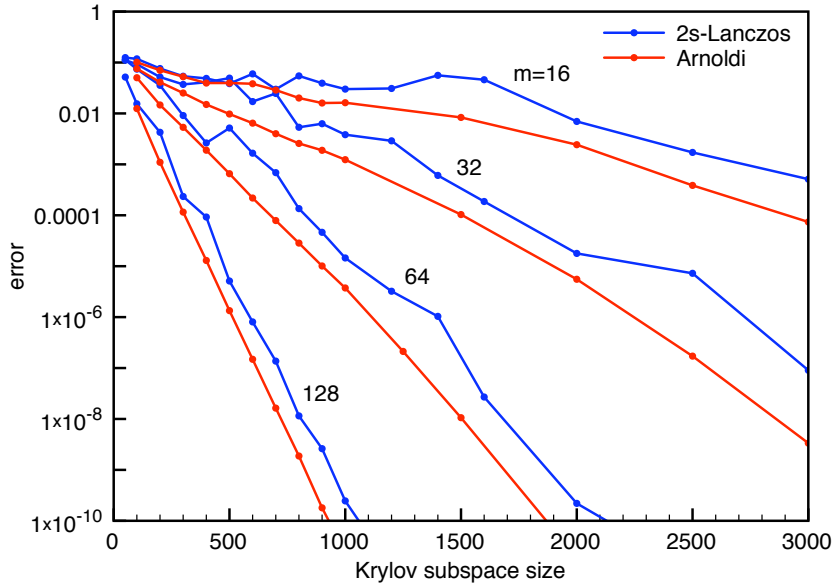
2s-Lanczos basis $\sim Nk$

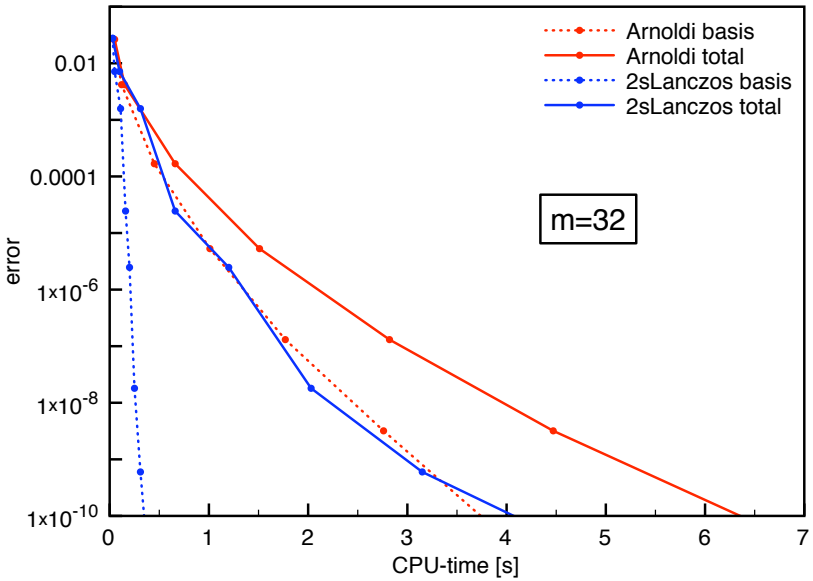
Arnoldi versus two-sided Lanczos – deflation efficiency

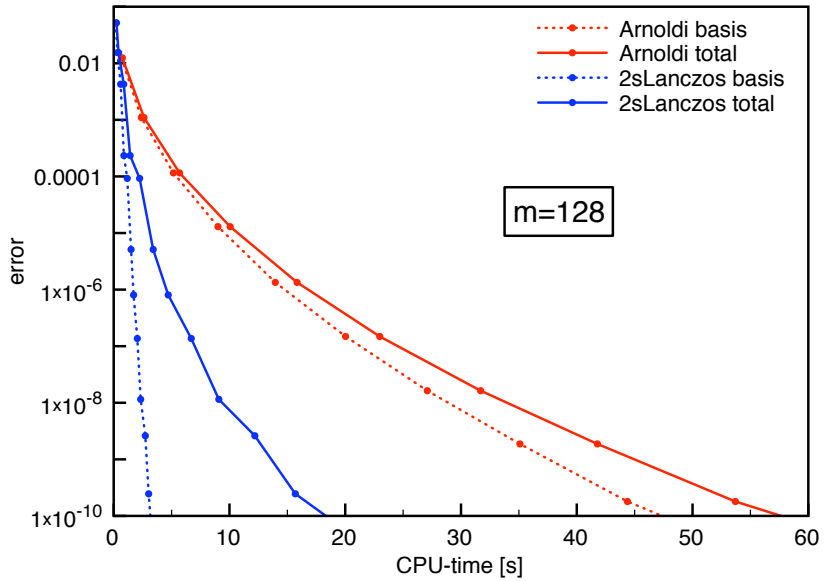
4^4 lattice (dim=3072)



6^4 lattice (dim=15552)







Outlook

- recursive Krylov subspace method for inner function computation
- alternative iterative methods:
  - restarted Arnoldi (stability problems)
  - partial fraction expansion
- improve efficiency of deflation
- apply method to larger lattices $\longrightarrow$ physics (tested for 8^4 lattice)
- use method in eigenvalue determination of overlap operator
  - test for 6^4 lattice using Arpack on Intel Core 2 Duo 2.33GHz
  - initialization: computing the 128 smallest eigenvalues of kernel $\sim$ 30 min
  - compute 16 smallest eigenvalues of overlap operator with accuracy of 10^{-4}
    - Arnoldi approximation with $k = 400 \sim$ 1h30min
    - 2S-Lanczos approximation with $k = 600 \sim$ 50min, $k = 400 \sim$ 25min
- use method in inversion of overlap operator