A multigrid accelerated eigensolver for the Hermitian Dirac Operator

A. Frommer, K. Kahl*, M. Rottmann*, A. Strebel**

Quantum Chromodynamics (QCD) is a quantum field theory for the strong interaction of the quarks via gluons and as such part of the standard model of elementary particle physics. Deducing predictions mostly requires a discretization onto a lattice. At the heart of Lattice QCD lies the Dirac Operator D and its Hermitian version Q. Here, eigenvalues of Q are needed to improve the signal-to-noise ratio when computing pion and eta-meson correlators.
In this work, we present a Davidson type eigensolver for Q which uses our successful DD-αAMG multigrid method as a preconditioner. Within this framework, we incorporated several modifications specifically designed for Q and the multigrid method. We employed a strategy which introduces a powerful synergy between the preconditioner and the Davidson method. Numerical results show the impact of these modifications and compare with common and state-of-the-art eigensolvers, i.e. PARPACK and PRIMME. This approach can also be used to formulate a new, promising multigrid setup proceduce.

*Bergische Universität Wuppertal

SFB TRR55

Motivation

Hadronic Observables in Lattice QCD

- (Wilson) Dirac Operator D is defined by

$$(D_W\psi)(x) = (m_0+4)\psi(x) - \frac{1}{2}\sum_{\mu=0}^{3}(I-\gamma_\mu)U_\mu(x)\psi(x+\hat{\mu}) - (I+\gamma_\mu)U_\mu^H(x-\hat{\mu})\psi(x-\hat{\mu})$$

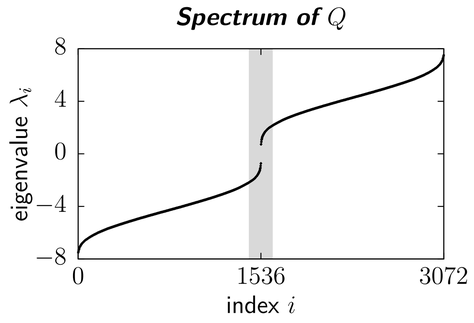
- Eta-meson correlator with $n_f = 2$ flavors and two-point function

$$C_\eta(x,y) = \langle O_x^\eta \bar{O}_y^\eta \rangle \propto \underbrace{\operatorname{tr}\left(\gamma_5 D^{-1}_{x,y}\gamma_5 D^{-1}_{y,x}\right)}_{\text{trace of } 12\times 12 \text{ matrix}} - n_f \underbrace{\operatorname{tr}\left(\gamma_5 D^{-1}_{x,x}\right)\operatorname{tr}\left(\gamma_5 D^{-1}_{y,y}\right)}_{\text{full matrix, infeasible to compute!}}$$

→ Use Monte Carlo simulation to estimate trace via $v^H Q^{-1} v$, $v \in \mathbb{Z}_2 + i\mathbb{Z}_2$, $Q = \Gamma_5 D$

- ⊕ Converges *almost surely* (with probability 1)
- ⊕ Only needs one inversion + inner product → cheap iterations
- ⊖ Error decreases with $\mathcal{O}(1/\sqrt{N_{stoch}})$ → many iterations needed

- Noise reduction technique → *low-mode averaging*
  - Split Operator $Q^{-1} = Q_{low}^{-1} + Q_{high}^{-1}$
  - $Q_{low}^{-1} = \sum_{i=1}^{k} \frac{1}{\lambda_i} v_i v_i^H$ smallest eigenpairs
  - Treat Q_{high}^{-1} with Monte Carlo
- Eta correlator dominated by low modes
  - ⇝ Few estimates for Q_{high}^{-1} are sufficient
- **Computation of small eigenpairs dominates overall cost**

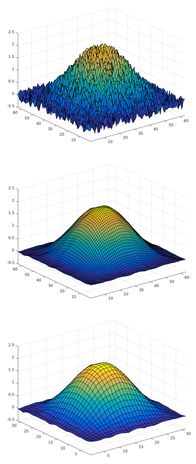


The Basic Ingredients

(Generalized) Davidson Framework

- Two-phase algorithm
  - Subspace extraction
  - Subspace expansion
- Broad spectrum of customization
- Supports preconditioning (line 13)
- Connection to other eigensolvers:
  - Solve $(I-uu^H)(Q-\theta I)(I-uu^H)t = -r$ → *Jacobi*-Davidson method
  - Choosing $t = r$ produces same subspace as Arnoldi's method

```
1   Input: initial guess t = v_0, accuracy ε
2   Output: eigenpair (λ, x)
3   for m = 1, 2, ...
4     t = (I − VV^H)t
5     v_m = t/||t||_2
6     V = [V|v_m]
7     H = V^H QV
8     get target eigenpair (θ, s) of H
9     u = Vs
10    r = Qu − θu
11    if ||r||_2 ≤ ε
12      λ = θ, x = u finished
13    t = M_m(r)
```



Domain Decomposition Adaptive Algebraic Multigrid: DD-αAMG

- Initial setup procedure generates interpolation/restriction operators
- Error gets *smoothed* → Gauss-Seidel, SAP, GMRES
- Represent smooth error on coarse grid
- Coarse grid treats part of spectrum not efficiently reduced by smoother
  → Can be approximated by small eigenvalues
- Solve coarse system and update error
- ⊕ Insensitive to condition number
- ⊕ DD-αAMG orders of magnitude faster than Krylov methods
- ⊖ Can not treat larger shifts efficiently

Subspace Extraction Phase

Finding Interior Eigenvalues

Rayleigh-Ritz extraction

Construct $H = V^H Q V$ → standard eigenvalue problem

- ⊕ Computationally cheap
- ⊖ Approximates exterior eigenvalues first ↔ interior needed!

vs.

Harmonic Ritz extraction

Solve $(QV)^H(QV)s_i = \theta_i (QV)^H V s_i$ → general eigenvalue problem

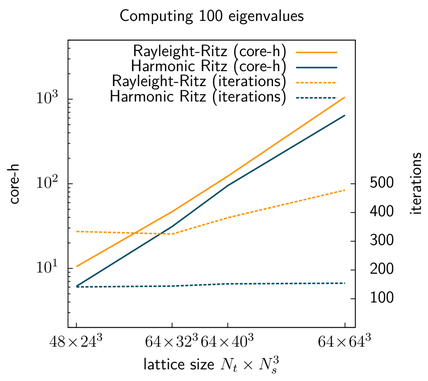
- ⊖ 2x IPs, GEP more expensive
- ⊕ Mimics EVs of Q^{-1} ↔ finds interior EVs

→ Harmonic Ritz reduces iterations by up to a factor 3 and time by 2

Locally Minimal Residuals[1]

- Extremal Ritz pairs ↮ extremal eigenpairs
- Either Ritz value or Ritz vector might be off (or both)
- **Here:** *Trust* vector and recompute value via Rayleigh quotient

→ Saves up to 20% outer iterations



[1] G. Sleijpen, H. van der Vorst, E. Meijerink, *Efficient expansion of subspaces in the Jacobi-Davidson method for standard and generalized eigenproblems*, Electronic Transactions on Numerical Analysis, 7:75-89, 1998

Subspace Expansion Phase

Efficient Evaluation of Preconditionier

- Use *Jacobi-Davidson* for more robust expansion
  - Avoids stagnation if correction equation is solved exactly
  - Enables controlling of inner iterations → Hochstenbach, Notay[2]
- Modified DD-αAMG[3] solver for inverting shifted systems with Q
  - SAP replaced by GMRES as smoother
  - Γ_5-preconditioning: $(I-uu^H)(D-\theta\Gamma_5)(I-uu^H)t = -\Gamma_5 r$

correction equation	iterations outer	iterations inner	Time core-h.
no Γ_5-prec	565	10,349	83.0
with Γ_5-prec	511	3,045	41.3

Handling Many Eigenvalues

- *Explicit* locking → keep converged eigenvectors in basis V
  - ⊕ Robust & easy to implement
  - ⊖ Size of V (and H) scales with no. of eigenvalues → worse eigenvalue scaling
- Thick restarting scheme
  - $V^H Q V = H = S^{-1}\Theta S$ eigenvalue decomposition
  - Keep smallest $n_{conv} + m_{min}$ Ritz values of Θ and according vectors in S
  - Recompute basis $V \leftarrow VS$, $H \leftarrow S^H H S$

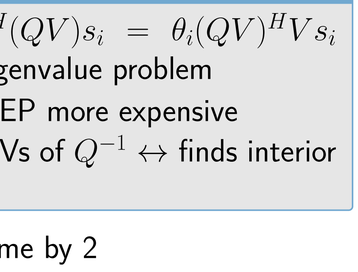
[2] M. Hochstenbach, Y. Notay, *Controlling inner iterations in the Jacobi-Davidson method*, SIAM J. Matrix Anal. Appl., 31(2):460-477, 2009

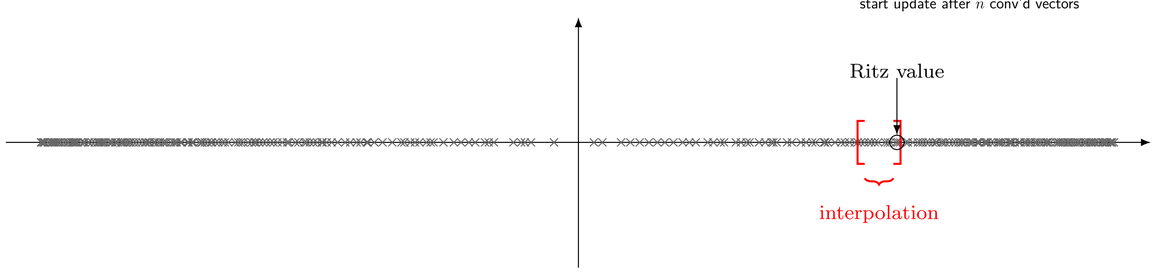
[3] A. Frommer, K. Kahl, S. Krieg, B. Leder, M. Rottmann, *Adaptive aggregation based domain decomposition multigrid for the lattice Wilson Dirac operator*, SIAM J. Sci. Comp., 36:A1581-A1608, 2014

Optimizing Solver Performance

- Shifted solves needed, once enough eigenvalues are found
- ⊖ Loss of *local coherence* if shift is too large
  → Coarse grid correction ineffective
  → Do not start shifting right away
- Update interpolation dynamically throughout eigenvalue computation
  1. Check sign of current Ritz value θ
  2. Replace interpolation vectors by nearest eigenvectors to θ and rebuild coarse grid operator

  ⇝ *One-sided deflation*
- Interpolation updating improves eigenvectors scaling significantly



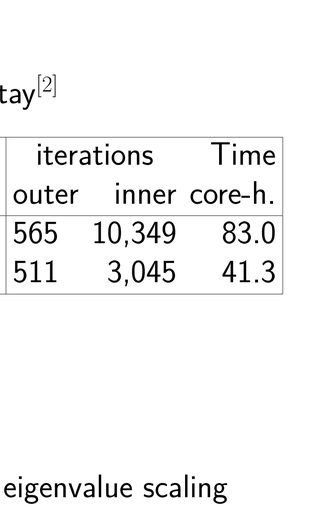


Comparison With Other Solvers

- PARPACK with (near optimal) Chebychev filter on Q^2
  - ⊖ Squaring of condition number
  - Lacks in lattice scaling
  - PARPACK with Q^{-1} also not competitive

  → Not considered in further tests
- PRIMME with mostly default parameters + Γ_5-preconditioning
  - Based on same framework ⇝ similar results expected
  - Implements state-of-the-art methods and techniques
  - Uses Rayleigh-Ritz extraction

  → Slightly better eigenvalue scaling
  → Benefits greatly from Γ_5-preconditioning



Future Work

(1) Algorithmic improvements
  - *Implicit* locking: Keep $V \perp X$
  - Multilevel Solver → lattice scaling

(2) Estimation for needed eigenvalues

(3) Use method as multigrid setup
  - First results show reduction of solver iterations by more than 50%
  - **Tradeoff:** Setup vs. Solve time

Acknowledgments

Numerical results were obtained on Jureca at Jülich Supercomputing Centre (JSC) through NIC grant HWU29

JÜLICH FORSCHUNGSZENTRUM