Multigrid Methods for Chiral Fermions

Peter Boyle

- Lattice QCD and MCMC
- HMC
- New multigrid methods for chiral fermions

Monte Carlo Integration

- Integration

$$\int_U f(U)dU = Vol \times \langle f \rangle \qquad ; \qquad Vol = \int_U dU$$

- Monte Carlo Integration (x_i uniform over compact domain)

$$\langle f \rangle = \frac{1}{N} \sum_i f(x_i)$$

- Importance sampling: draw x_i with positive normalised probability density $P(x_i)$

$$\langle f \rangle = \frac{1}{N} \sum_i \frac{f(x_i)}{P(x_i)}$$

- If $|f(x_i)| \propto P(x_i)$ this may converge better.
  NB: analogy to removing bias from fixing machine learned MC integrators
- Variance reduction: if $\tilde{f}$ is a good, cheap approximation for f

$$\langle f \rangle = \frac{1}{N} \sum_i \frac{\tilde{f}(x_i)}{P(x_i)} + \frac{1}{N \times M} \sum_j \frac{f(x_j) - \tilde{f}(x_j)}{P(x_j)}$$

Euclidean Path Integral

- Pure gauge path integral
$$\frac{1}{Z}\int_U e^{-S_G[U]}\mathscr{O}(U)dU$$
- Importance sample: seek to distribute gluon configuratoins according to
$$P(U) = \frac{e^{-S_G[U]}}{\int_U e^{-S_G[U]}dU}$$
- Calculate observables on each *configuration*
$$\langle\mathscr{O}\rangle = \frac{1}{N}\sum_i \mathscr{O}(U_i)$$
- $\Rightarrow$ Markov chain monte carlo:
  - 10^{10} degrees of freedom(!)
  - Sharply probability weight
  - Variance of $\mathscr{O}(U)$ determines how many samples are required.
  - 100-2000 samples typically good for 1% scale statistical errors
- This is observable dependent...

Markov chains

- Sequence of states generated by *transition probability* $M(X \to X')$ from X to X'
  Rule depends only on X.
- Usually composed of proposal and acceptance probablities

$$M(X \to X') = P_p(X \to X')P_{acc}(X \to X')$$

- Design rule $M(X \to X')$ to yield *desired* equilibrium probability distribution after many transitions$P_{eq}(X)$
- P_{eq} must map to itself under of the transition rule:

$$P_{eq}(X') = \sum_X P_{eq}(X)M(X \to X')$$

- An ergodic update satisfying this *and* is a contration mapping and Markov transitions converge on the desired equilibrium.

(Clear pedagogical review: Anthony Kennedy Nara lectures 2006)

Metropolis algorithms

- Detailed balance property:

$$P_{eq}(X)M(X \to X') = P_{eq}(X')M(X' \to X)$$

- Sum over X to obtain

$$\sum_X P_{eq}(X)M(X \to X') = \sum_X P_{eq}(X')M(X' \to X) = P_{eq}(X')$$

- So P_{eq} is a *fixed point* of the Markov process!
- We can sample any probability distribution we desire with such an update.

Metropolis algorithms

- Make the update combine proposal and acceptance probablities

$$M(X \to X') = P_p(X \to X') P_{acc}(X \to X')$$

- detailed balance

$$P_{eq}(X) P_p(X \to X') P_{acc}(X \to X') = P_{eq}(X') P_p(X' \to X) P_{acc}(X' \to X),$$

- is satisfied with the Metropolis acceptance probability,

$$P_{acc}(X \to X') = \mathrm{Min}(1, \frac{P_{eq}(X') P_p(X' \to X)}{P_{eq}(X) P_p(X \to X')})$$

- either $P_{acc}(X \to X') = 1$, or $P_{acc}(X' \to X)$; considering cases leads to trivial proof.
- Simplifies if $P_p(X' \to X) = P_p(X \to X')$ (reversible, area preserving constraint)

Basis of most Markov Chain Monte Carlo

- Aspects of this might be of interest to the numerical integration / ML / journal club (?)

QCD path integral

- Partition function becomes a real, statistical mechanical probability weight

$$Z = \int d\bar{\psi} d\psi dU e^{-S_G[U] - S_F[\bar{\psi}, \psi, U]}$$

- Dirac differential operator represented via discrete derivative approximations: sparse matrix
- (Until Taku & Gumaro sort QIS)
  use pseudofermion approach to replace with Gaussian integral $\sqrt{\pi\lambda} = \int dt e^{-t^2/\lambda}$

$$\int \mathscr{D}\bar{\psi} \mathscr{D}\psi e^{-\bar{\psi}(x) A_{xy} \psi(y)} = \det A$$

$$\pi\lambda = \int d\phi_r e^{-\phi_r \frac{1}{\lambda} \phi_r} \int d\phi_i e^{-\phi_i \frac{1}{\lambda} \phi_i} = \int d\phi^* d\phi e^{-\phi^* \frac{1}{\lambda} \phi}$$

- replace two flavour determinant with a two flavour *pseudofermion* integral

$$(\det M)^2 = (\det \gamma_5 M)^2 = \det M^\dagger M = \int \mathscr{D}\phi^* \mathscr{D}\phi e^{-\phi^*(x)(M^\dagger M)^{-1} \phi(y)}$$

Hybrid Monte Carlo

- Auxiliary Gaussian integral over conjugate momentum field $\int d\pi e^{\frac{-\pi^2}{2}}$
  Lives in Lie algbra; serves only to move U round the group Manifold

$$\int d\pi \int d\phi \int dU \quad e^{-\frac{\pi^2}{2}} e^{-S_G[U]} e^{-\phi^*(M^\dagger M)^{-1}\phi}$$

- Outer Metropolis Monte Carlo algorithm
  - Draw momenta
  - Draw pseudofermion as gaussian $\eta = M^{-1}\phi$
  - Metropolis acceptance step
- Metropolis proposal includes inner molecular dynamics at constant Hamiltonian:

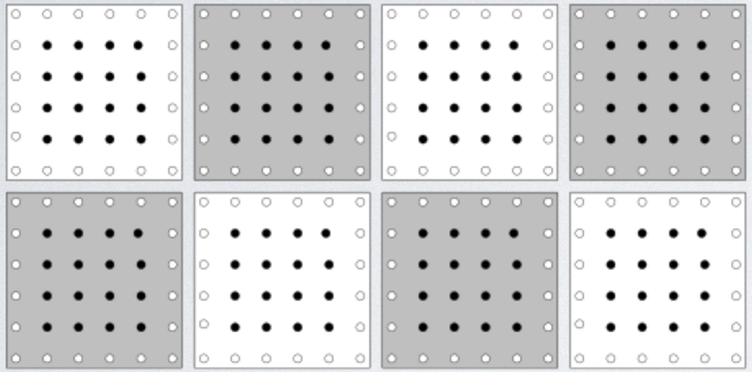
$$H = \frac{\pi^2}{2} + S_G[U] + \phi^*(M^\dagger M)^{-1}\phi$$

$$\dot{U} = i\pi U \qquad ; \qquad i\dot{\pi} = (U\nabla_U S)_{TA}$$

- Must invert $M^\dagger M$ at each timestep of evolution in MD force

$$\delta(M^\dagger M)^{-1} = -(M^\dagger M)^{-1}[(\delta M^\dagger)M + M(\delta M)](M^\dagger M)^{-1}$$

- ★ Domain Decompose the Fermionic path integral at the node level
- ★ Large domains: algorithm tailored to GPU computation
- ★ Local domain solvers decouple from interconnect, realise full potential



node local - runs well on an island!

Pauli antisymmetry

$$\det D = \det D_\Omega \det D_{\bar{\Omega}} \det \left\{ 1 - D_\Omega^{-1} D_\partial D_{\bar{\Omega}}^{-1} D_{\bar{\partial}} \right\}$$

Integrate non-local term on larger timestep

$$\begin{pmatrix} D_\Omega & D_\partial \\ D_{\bar{\partial}} & D_{\bar{\Omega}} \end{pmatrix} = \begin{pmatrix} 1 & D_\partial D_{\bar{\Omega}}^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} D_\Omega - D_\partial D_{\bar{\Omega}}^{-1} D_{\bar{\partial}} & 0 \\ 0 & D_{\bar{\Omega}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ D_{\bar{\Omega}}^{-1} D_{\bar{\partial}} & 1 \end{pmatrix}$$

Schur decomposition

Estimate 5x to 10x gain

Observables

Importance sampling has reduced:

$$\langle \mathscr{O} \rangle = \frac{1}{Z} \int_U e^{-S_G[U]} \mathscr{O}(U) dU \rightarrow \frac{1}{N} \sum_i \mathscr{O}(U_i)$$

- Zero momentum pion, kaon or B meson two point function

$$\sum_x \langle \bar{u} \gamma_0 \gamma_5 d(x,t) \bar{d} \gamma_0 \gamma_5 u(0,0) \rangle = \frac{1}{N} \sum_i \mathrm{trace}\{ \gamma_0 \gamma_5 M_d^{-1}(x,t;0,0) \gamma_0 \gamma_5 M_u^{-1}(0,0;x,t) \}$$

- Euclidean space $\propto Ae^{-mt}$
- Tune bare mass until interacting meson mass is correct, prefactor gives pion, kaon, B meson decay constant
- etc..

Scalar field

- $S = \phi^*(\Box + m^2)\phi$

$$M(x,x') = \delta_{x,x'} 2(N_d + m^2) - \sum_\mu \delta_{x+\mu,x'} + \delta_{x-\mu,x'}$$

- **Free case:** M is hermitian diagonalised by a unitary transformation

$$M = V^\dagger D V$$

- Call it the "discrete fourier transform" in the free case, and the eigenvalues are

$$D(p) = (2\sin p/2)^2 + m^2$$

  - Propagator is the inverse of this

- **Interacting case:** M is hermitian diagonalised by a unitary transformation

$$M = V^\dagger D V$$

- covariant derivative couples to gauge fields, numerical solution of propagator
- Still diagonalisable, eigenvectors no longer plane waves etc...

$$M^{-1} = V \mathrm{Diag}\{\frac{1}{\lambda_i}\} V^\dagger \simeq V \mathrm{Diag}\{P(\lambda_i)\} V^\dagger$$

- If P is polynomial approximating $\frac{1}{x}$ over the whole spectral range $\Rightarrow$ Krylov solvers

Greens functions in a lattice calculation

Each gauge *configuration* individually breaks translation invariance; restored under gauge average

$$S_F(x,y) \neq S_F(x-y)$$

Quark propagator is the Greens function of the Dirac operator D on each gauge sample

$$G(y|x) = D^{-1}_{y,z}\delta_{z,x}$$

Never (exactly) compute all elements of $G(y,z)$
Must *solve* the Dirac equation for each position independently.

$$D_{xy}\psi(y) = \eta(x)$$

Algorithm minimising

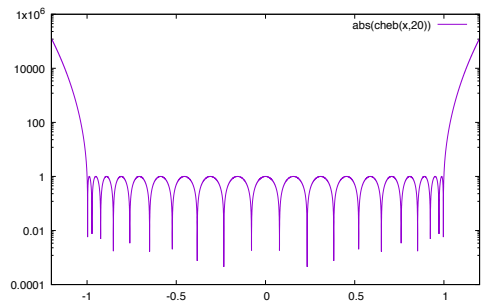
$$|r| = |D_{xy}\psi(y) - \eta(x)|$$

will find $\psi = D^{-1}\eta$
Standard algorithms (Krylov solvers) are just (matrix) polynomial approximations to $\frac{1}{x}$
Conjugate Gradients, Hestene and Steiffel (1952) has 8733 citations!

Chebyshev polynomials

$$T_n(x) = \cos(n \cos^{-1} x)$$

- n half periods over $[-1, 1]$
  - Uniform amplitude ripples
  - x^n growth outside range
  - School boy induction using double angle formula $\Rightarrow$ recursion relation
- Orthogonal under weight $\rho(x) = (1 - x^2)^{\frac{1}{2}}$
- Expansion in orthogonal polynomials $\Rightarrow$ Chebyshev approximation



Critical slowing down

Chebyshev worst case convergence bound for Conjugate Gradients

$$\frac{|r|_n}{|r|_{n-1}} = \sigma \leq \frac{\sqrt{k}-1}{\sqrt{k}+1}$$

Condition number

$$k = \frac{\lambda_{\max}}{\lambda_{\min}}$$

Limiting cases:

- $k = 1 \rightarrow \sigma = 0$ Converges in one iteration if given a matrix with all eigenvalues equation (basis change of identity).
- $k = 1 \rightarrow \sigma = 1$ Never converges if the matrix is singular $\lambda_{min} = 0$
- $\sigma(k)^n = \text{tolerance}$ estimates iteration count n

Spectrum of Dirac operator is important

Dirac Spectrum

- Atiyah Singer index theorem
  - Density of zero modes of Dirac operator $\rightarrow$ eigenvalues separated from zero only by quark mass
- Lowest (dimensionless) eigenvalue $am_u \rightarrow 0$ in continuum limit
- Domain Wall / chiral lattice fermions reproduce the chiral anomaly:
  - exact zero modes in classical topological fields
  - integer topological index
  - satisfy Atiyah-Singer index theorem *in a discrete system*
- Probe non-perturbative dynamics of Yang-Mills theory

UnKrylov solvers

Possible resolutions to critical slowing down

1. Determine and accurately project a few thousand low eigenmodes[1]
   - Cannot help HMC gauge field sampling
2. Multigrid methods *approximately* treating this low mode space
   - Introduced for Wilson (Luscher 2007, Brower et al 2007)

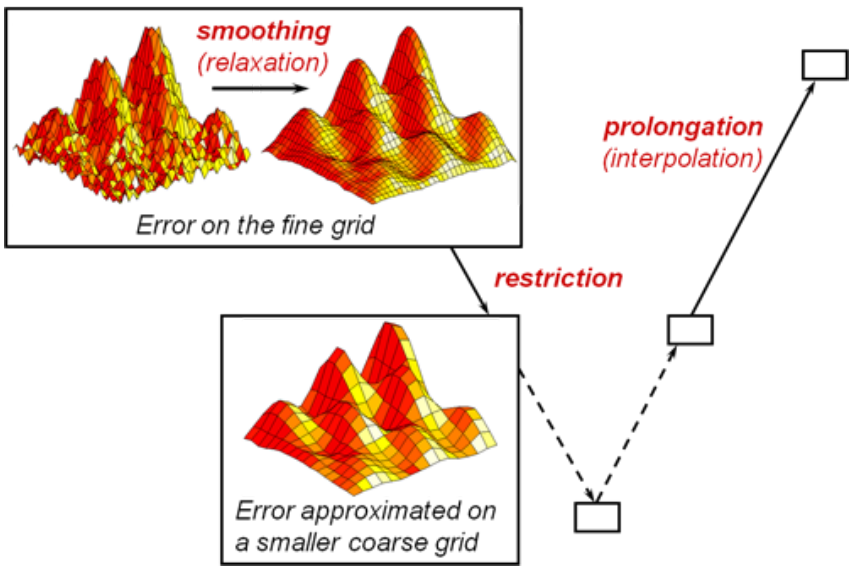
Aim:

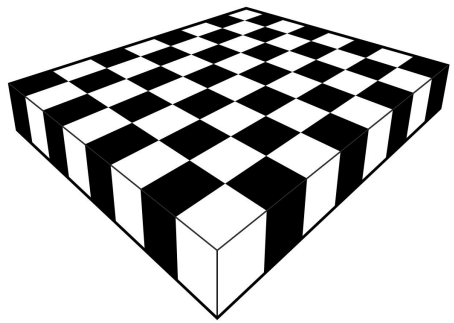
- Produce a deflation algorithm with a real gain *inside* HMC

[1] not all topological in nature

Multigrid : how it works

- Project to low dimensional basis that captures the low mode space
- Represent the original matrix in this truncated basis
- Inverse of this truncated representation corrects the current solution





	Fine	Coarse	CoarseCoarse	Evecs
λ_{min}	1.0e-6	1.0e-6	1.0e-6	1.0e-6
λ_{max}	60	11	5.0	4.0e-3

- Improve the condition number by lowering the cut-off as you go coarser
- Arguably a surface to volume suppression of the high modes as you block
- Smoother step helps cheaply wipe out the effects while preserving the low mode element of coarse correction

Domain wall multigrid

PROCEEDINGS OF SCIENCE

Hierarchically deflated conjugate residual

Azusa Yamaguchi[*†]
University of Edinburgh
E-mail: ayamaguc@staffmail.ed.ac.uk

Peter A Boyle
University of Edinburgh
E-mail: paboyle@ph.ed.ac.uk

We present a progress report on a new class of multigrid solver algorithm suitable for the solution of 5d chiral fermions such as Domain Wall fermions and the Continued Fraction overlap. Unlike HDCG [1], the algorithm works directly on a nearest neighbour fine operator. The fine operator used is Hermitian indefinite, for example $\Gamma_5 D_{dwf}$, and convergence is achieved with an indefinite matrix solver such as outer iteration based on conjugate residual. As a result coarse space representations of the operator remain nearest neighbour, giving an 8 point stencil rather than the 81 point stencil used in HDCG. It is hoped this may make it viable to recalculate the matrix elements of the little Dirac operator in an HMC evolution.

Multigrid for Chiral Lattice Fermions: Domain Wall

Richard C. Brower[*], M. A. Clark[†], Dean Howarth[*], and Evan S. Weinberg[†]

[*]Boston University, Boston, MA 02215, USA
[†]NVIDIA Corporation, Santa Clara, CA 95050, USA

April 17, 2020

Abstract

Critical slowing down for the Krylov Dirac solver presents a major obstacle to further advances in lattice field theory as it approaches the continuum solution. We propose a new multi-grid approach for chiral fermions, applicable to both the 5-d domain wall or 4-d Overlap operator. The central idea is to directly coarsen the 4-d Wilson kernel, giving an effective domain wall or overlap operator on each level. We provide here an explicit construction for the Shamir domain wall formulation with numerical tests for the 2-d Schwinger prototype, demonstrating near ideal multi-grid scaling. The framework is designed for a natural extension to 4-d lattice QCD chiral fermions, such as the Möbius, Zolotarev or Borici domain wall discretizations or directly to a rational expansion of the 4-d Overlap operator. For the Shamir operator, the effective overlap operator is isolated by the use of a Pauli-Villars preconditioner in the spirit of the Kähler-Dirac spectral map used in a recent staggered MG algorithm [1].

Comparison of Domain Wall Fermion Multigrid Methods

Peter Boyle
HET Physics Department, Brookhaven National Laboratory, Upton, NY 11973, USA. and School of Physics and Astronomy, University of Edinburgh, Edinburgh EH9 3JZ, UK

Azusa Yamaguchi
School of Physics and Astronomy, University of Edinburgh, Edinburgh EH9 3JZ, UK

Abstract

We present a detailed comparison of several recent and new approaches to multigrid solver algorithms suitable for the solution of 5d chiral fermion actions such as Domain Wall fermions in the Shamir formulation, and also for the Partial Fraction and Continued Fraction overlap. Our focus is on the acceleration of gauge configuration sampling, and a compact nearest neighbour stencil is required to limit the calculational cost of obtaining a coarse operator. This necessitates the coarsening of a nearest neighbour operator to preserve sparsity in coarsened grids, unlike HDCG[1]. We compare the approaches of references[2, 3] and also several new hybrid schemes. In this work we introduce a new recursive Chebyshev polynomial based setup scheme. We find that the approach of reference[2], can both setup, and solve standard Shamir Domain Wall Fermions faster than a single solve with red-black preconditioned Conjugate Gradients[30] on large volumes and for modern GPU systems such as the Summit supercomputer. This is promising for the acceleration of HMC, particularly if setup costs are shared across multiple Hasenbusch determinant factors. The setup scheme is likely generally applicable to other Fermion actions.

2016, QCD, D=4

2020, D=2, Schwinger model

D=4, QCD both methods
Almost done

And… 2014, PAB, HDCG

Domain wall Fermion action

$$S^5 = \int d^4x \bar{\psi} D^5_{DW} \psi,$$

$$D_W(M) = M + 4 - \frac{1}{2} D_{\text{hop}},$$

$$D_{\text{hop}} = (1-\gamma_\mu) U_\mu(x) \delta_{x+\mu,y} + (1+\gamma_\mu) U^\dagger_\mu(y) \delta_{x-\mu,y}.$$

$$D^5_{DW} = \begin{pmatrix} D_\parallel & -P_- & 0 & \dots & 0 & mP_+ \\ -P_+ & \ddots & \ddots & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \ddots & \ddots & -P_- \\ mP_- & 0 & \dots & 0 & -P_+ & D_\parallel \end{pmatrix},$$

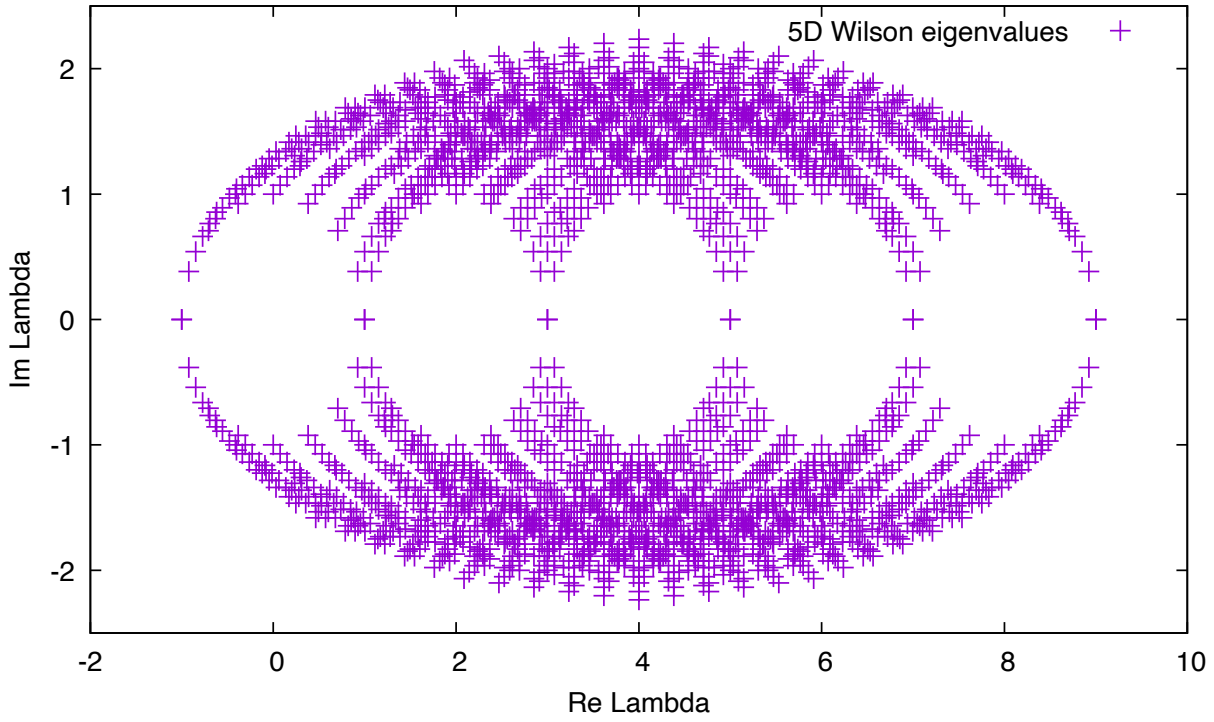
$$D_\parallel = 5 - M_5 - \frac{1}{2} D_{\text{hop}} = D_W(-M_5) + 1.$$

Gamma-5 Hermiticity:

$\Gamma_5 = \gamma_5 R_5$, where R_5 denote reflection in the fifth dimension.

$H_{DW} = \Gamma_5 D_{DW}$.

Chiral modes bound to surfaces in a fictitious fifth dimension
Exponentially accurate chiral symmetry
Good for weak matrix elements, V-A current etc…



Free field spectrum:

$$\mathrm{Re}\,\lambda \;=\; m + 5 - \sum \cos p_\mu \;\text{ and }\; |\,\mathrm{Im}\,\lambda|^2 \;=\; \sum \sin^2 p_\mu .$$

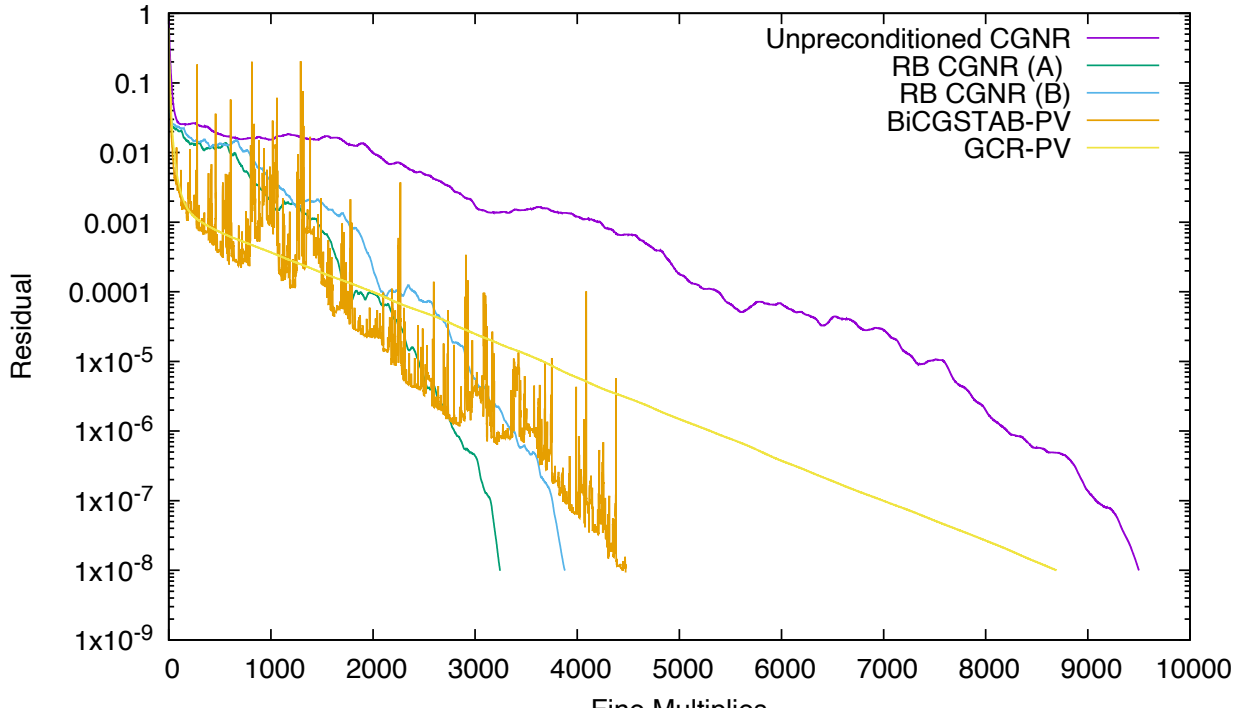
Problematic for polynomial approximation of 1/z

CGNR $\quad M_{pc}^\dagger M_{pc}\psi = \eta$

It is impossible to reproduce the phase winding of $\frac{1}{z} = \frac{1}{r}e^{-i\theta}$ around zero with an analytic function. Indeed, perhaps belabouring the point, the orthogonality of the set of functions $e^{im\theta}$ over $[0, 2\pi]$ makes it easy to show that minimising the uniformly weighted mean square error over a fixed radius circle gives precisely zero for all polynomial coefficients. In the

Krylov solvers on 16^3 configuration, m=0.001

Algorithm	Operator	Iterations	Full Matmuls	Time (s)
CGNR	$M^\dagger M$	9541	19082	183s
BiCGSTAB	$M_{PV}^\dagger M$	4140	8280	79s
prec-CGNR	$(M_{ee} - M_{eo}M_{oo}^{-1}M_{oe})^\dagger(M_{ee} - M_{eo}M_{oo}^{-1}M_{oe})$	3224	6448	62s
prec-CGNR	$(1 - M_{ee}^{-1}M_{eo}M_{oo}^{-1}M_{oe})^\dagger(1 - M_{ee}^{-1}M_{eo}M_{oo}^{-1}M_{oe})$	3880	7760	77s
GCR(32,32)	$M_{PV}^\dagger M$	8693	17386	474s



Lattice multigrid

Generate near null vectors – new with Chebyshev:

$$deflation\ \ basis\ \ \text{of vectors}\ \ \phi_k$$

Restrict to cuboidal blocks

$$\phi_k^b(x) = \begin{cases} \phi_k(x) & ;\ x \in b \\ 0 & ;\ x \notin b \end{cases}$$

Use as a cheap, compressed short-hand basis for the low modes

Efficient because linear space spanned by blocks is O(volume) bigger

$$\text{span}\{\phi_k\} \subset \text{span}\{\phi_k^b\}.$$

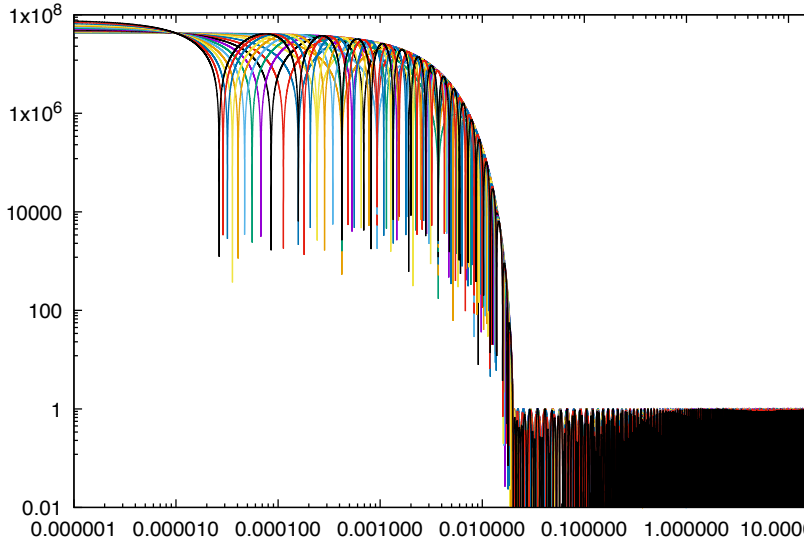


FIG. 2. Overlay of Chebyshev low-pass filter functions used to create the subspace.

Local coherence and deflation of the low quark modes in lattice QCD

Martin Lüscher

CERN, Physics Department, TH Division
CH-1211 Geneva 23, Switzerland

If the 64×32^3 lattice is divided into blocks of size 4^4, for example, and if 12 eigenmodes out of 48 are selected for the construction of the domain-decomposed subspace, the remaining 36 modes turn out to lie in the subspace up to deficits ϵ ranging from 0.03 to 0.06. The deficits increase with the block size, but become smaller if more modes are used for the subspace construction. On the 48×24^3 lattice the situation is practically the same, i.e. similar deficits are obtained for a given block size and subspace dimension.

We introduce projectors to the subspace S and its complement $\bar{S}$,

$$P_S = \sum_{k,b} |\phi_k^b\rangle\langle\phi_k^b| \qquad ; \qquad P_{\bar{S}} = 1 - P_S,$$

$$M = \begin{pmatrix} M_{\bar{S}\bar{S}} & M_{S\bar{S}} \\ M_{\bar{S}S} & M_{SS} \end{pmatrix}$$

Define coarse grid operator:

$$A_{jk}^{ab} = \langle\phi_j^a|M|\phi_k^b\rangle \qquad ; \qquad (M_{SS}) = A_{ij}^{ab}|\phi_i^a\rangle\langle\phi_j^b|.$$

Chebyshev smoothers – cj are the Chebyshev expansion of 1/x:

$$S_{\text{chebyshev([a,b],N)}} = \left[\frac{1}{2}c_0 + \sum_{j=1}^{N} c_j T_j\left(\frac{2(M^\dagger M - a)}{b-a} - 1\right)\right] M^\dagger$$

V(11)-cycle multigrid preconditioner

$$x_1 = x_0 + Sr_0$$

$$x_2 = x_1 + Qr_1$$

$$x_3 = x_2 + Sr_2.$$

$$[SP_L + Q + P_R S + SP_L MS]$$

Two ways to fix inappropriate spectrum and retain nearest neighbour coarsening

BNL / Edinburgh

Mimic how CGNR works
Hermitian operator has real indefinite spectrum

$$\lambda \in [-\lambda_{\max}, -\lambda_{\min}] \cup [\lambda_{\min}, \lambda_{\max}]$$

Coarsen

$$H_{DW} = \Gamma_5 D_{DW}.$$

Gamma5 compatible

$$\mathbb{P}^\dagger \Gamma_5 D_{DW} \mathbb{P} \mathbb{P}^\dagger \Gamma_5 D_{DW} \mathbb{P} = \mathbb{P}^\dagger D_{DW}^\dagger \mathbb{P} \mathbb{P}^\dagger D_{DW} \mathbb{P}$$

Boston

$M_{PV}^\dagger M_l$

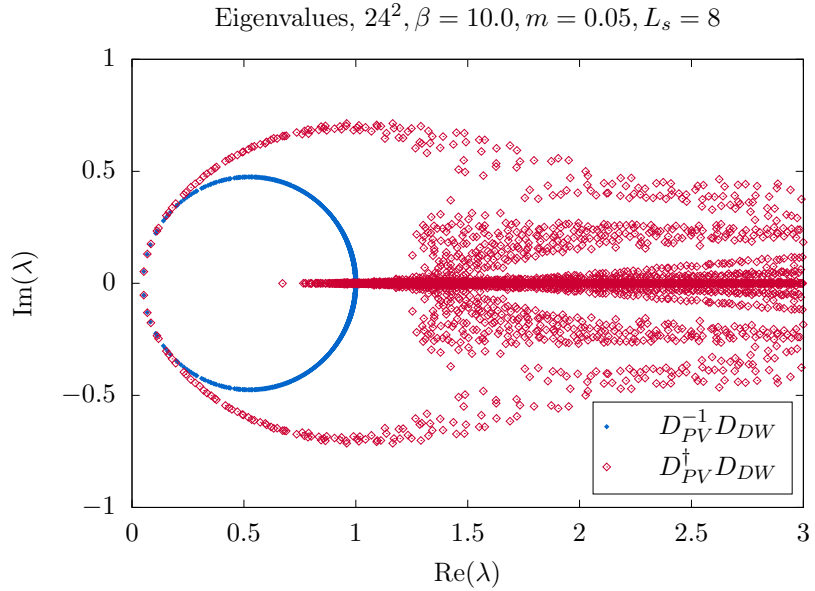


Figure 2.5: The spectrum of our target multigrid operator, $D_{PV}^\dagger D_{DW}$, compared with the effective overlap spectrum, $D_{PV}^{-1} D_{DW}$. For clarity of presentation we truncate the x-axis; the spectrum of $D_{PV}^\dagger D_{DW}$ extends out to Re(λ)$\approx$ 25.

Coarsen Dw – make a coarse Ddwf out of coarse Dw

$$\mathbb{P}^\dagger D_{PV}^\dagger \mathbb{P} \mathbb{P}^\dagger D_{DW} \mathbb{P}$$

Share code between fine Grid Mobius and Coarse Grid Mobius

$$\hat{D}^5_{GDW} = \begin{pmatrix} \hat{D}_+ & -\hat{D}_- P_- & 0 & \dots & 0 & m\hat{D}_- P_+ \\ -\hat{D}_- P_+ & \ddots & \ddots & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \ddots & \ddots & -\hat{D}_- P_- \\ m\hat{D}_- P_- & 0 & \dots & 0 & -\hat{D}_- P_+ & \end{pmatrix}$$

Implemented Boston in D=4 QCD and run an algorithm shoot out on 16^3 x 32 Ls=16 m=0.001

BNL/Edinburgh

Name	Algorithms
HDCR	A,B
MG-PV	C,D
MG-$M^\dagger M$	E
Hybrid-$M^\dagger M$	F
HDCG	G

Boston

BNL-Boston hybrid

Edinburgh 2014

Algorithm shoot out

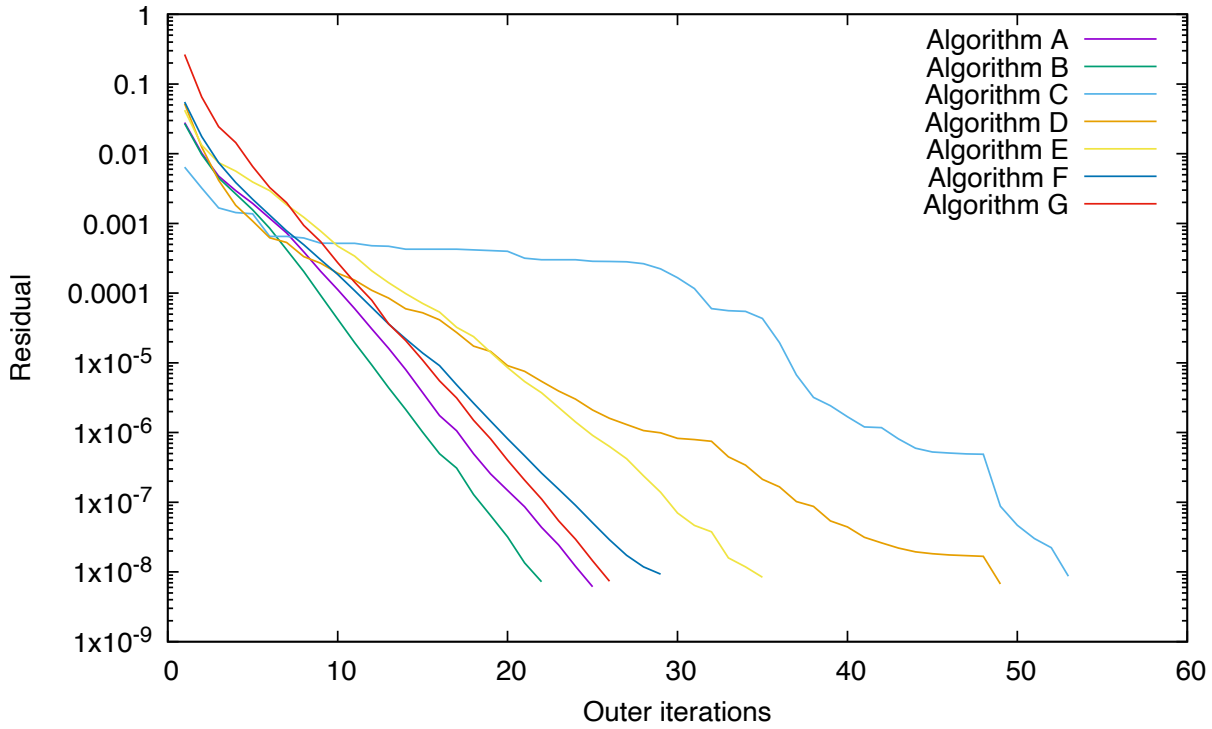
Many parameters tuned

Algorithm	A	B
Fine Grid	$16^3 \times 32 \times 16$	$16^3 \times 32 \times 16$
Block	$4 \times 2 \times 4 \times 2 \times 16$	$2^4 \times 16$
Coarse Grid	$4 \times 8 \times 4 \times 16 \times 1$	$8^3 \times 16 \times 1$
Outer Krylov	pGCR(H_{DW})	pGCR(H_{DW})
Basis vectors	40	32
Smoother	$S_{\text{chebyshev}}([0.5, 60], 12)$	$S_{\text{chebyshev}}([0.5, 60], 12)$
Coarsening	H_{DW}	H_{DW}
Coarse Solver	Deflated CGNR	Deflated CGNR
Coarse Tolerance	0.02/0.04	0.02/0.04
Coarse Eigenvectors	48/64	48/128
Subspace $\lambda_{\max}$	60.0	60.0
Subspace λ_{lo}	0.05	0.05
Subspace m	500	500
Subspace Δ	100	100

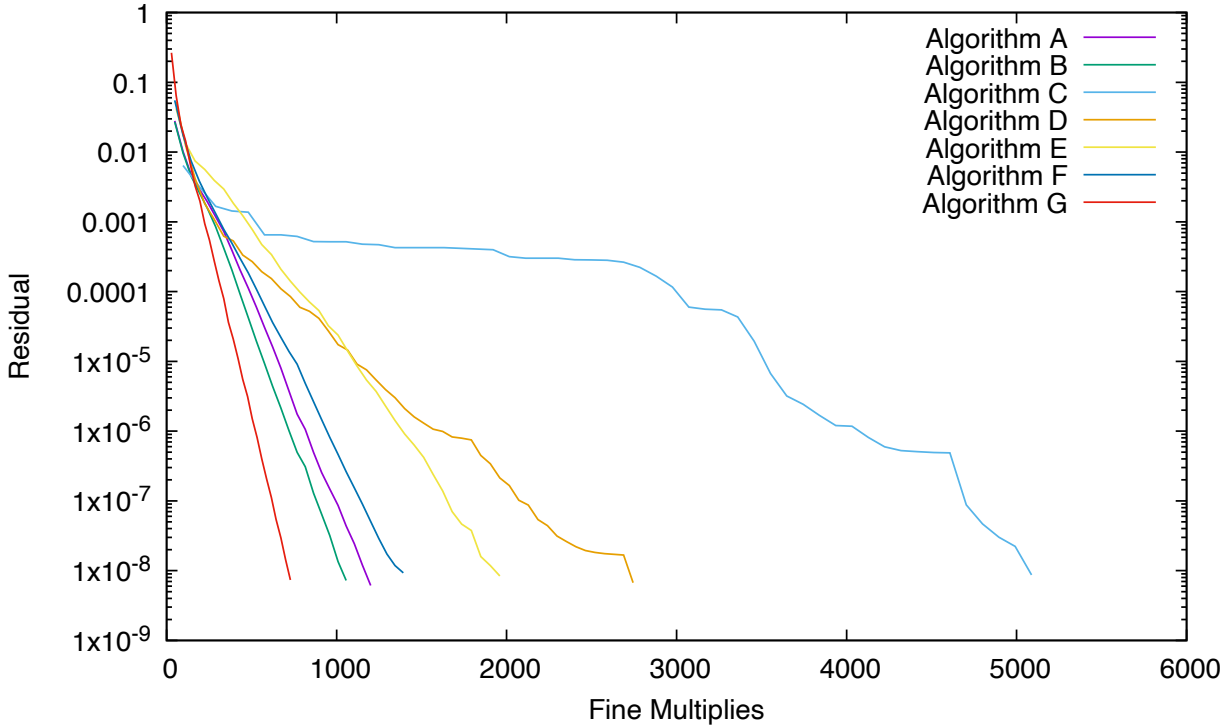
Algorithm	C	D	E
Fine Grid	$16 \times 32 \times 16$	$16^3 \times 32 \times 16$	$16^3 \times 32 \times 16$
Block	$2^4 \times 1$	$2^4 \times 1$	$2^4 \times 1$
Coarse Grid	$8^3 \times 16 \times 16$	$8^3 \times 16 \times 16$	$8^3 \times 16 \times 16$
Fine Krylov	pGCR($M_{PV}^{\dagger} M$)	pGCR($M_{PV}^{\dagger} M$)	pGCR(M)
Smoother	$S_{\text{BiCGSTAB}}(24)$	$S_{\text{GCR}}(14)$	$S_{\text{chebyshev}}([0.5, 60], 14)$
Coarsening	D_W	D_W	D_W
Coarse Solver	BiCGSTAB($M_{PV}^{\dagger} M$)	GCR($M_{PV}^{\dagger} M$)	CGNR($M^{\dagger} M$)
Coarse Tolerance	0.02	0.02	0.02/0.1
Coarse Eigenvectors	0	0	0/64
Subspace basis	24	24	24
Subspace $\lambda_{\max}$	60.0	60.0	60.0
Subspace λ_{lo}	4.0	4.0	4.0
Subspace m	600	600	600
Subspace Δ	250	250	250

Algorithm	F
Fine Grid	$16^3 \times 32 \times 16$
Fine Krylov	pGCR(M)
Smoother	$S_{chebyshev}([0.5, 60], 12)$
Coarsening	D_W
Coarse Grid	$8^3 \times 16 \times 16$
Coarse Solver	pGCR(M)
Coarse Tolerance	0.01
Subspace basis	24
Subspace $\lambda_{\max}$	60.0
Subspace λ_{lo}	4.0
Subspace m	600
Subspace Δ	250
Coarsening	H_{DW}
Coarse Grid	$8^3 \times 16 \times 1$
Coarse Solver	$CGNR(M^{\dagger} M)$
Coarse Tolerance	0.02
Coarse Eigenvectors	128
Subspace basis	32
Subspace $\lambda_{\max}$	60.0
Subspace λ_{lo}	4.0
Subspace m	600
Subspace Δ	250

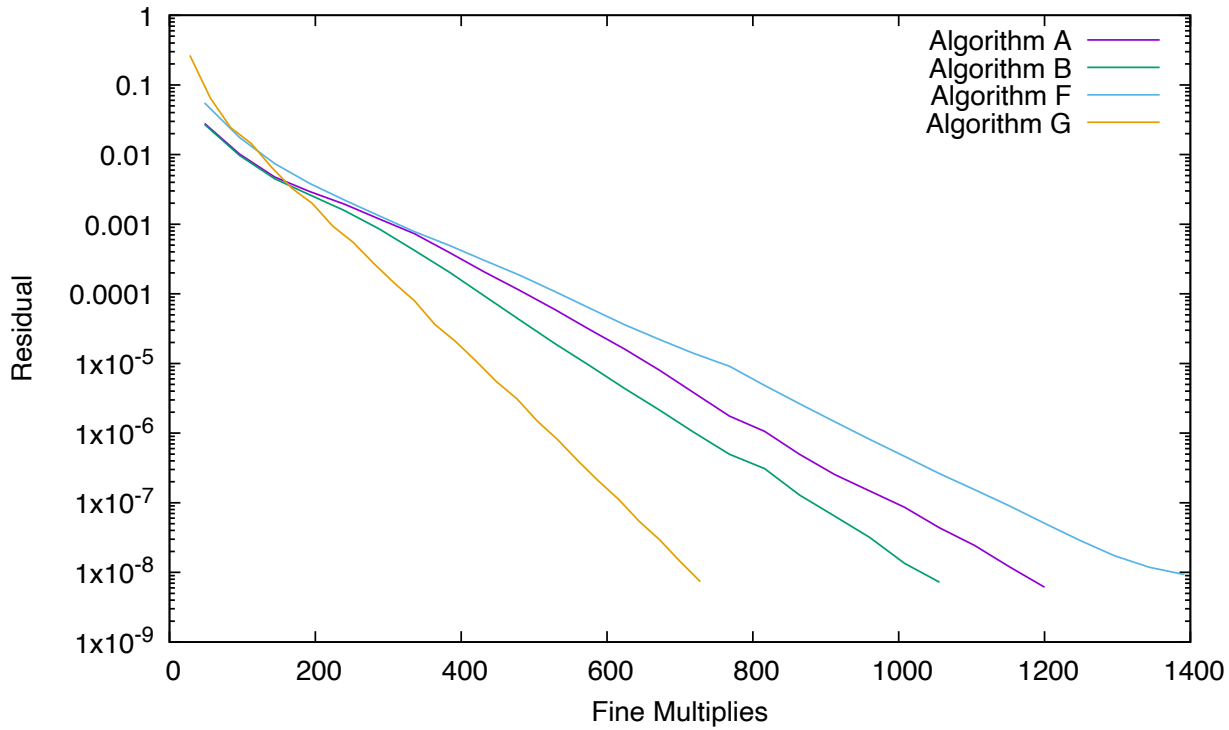
Cost of the preconditioner is NOT included – deflation effect is clear



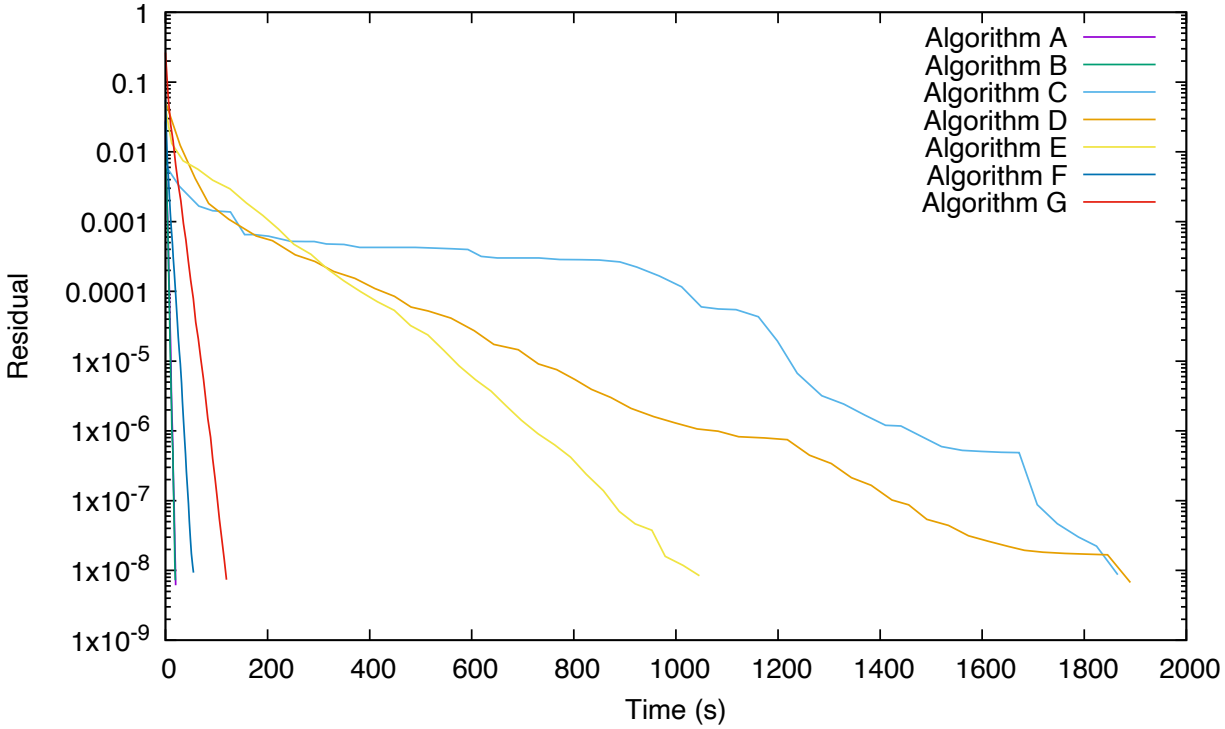
Fine matrix multiples includes cost of smoothers and outer steps.
Ignores cost of coarse space solves.
Had to use higher order smoother for Boston method

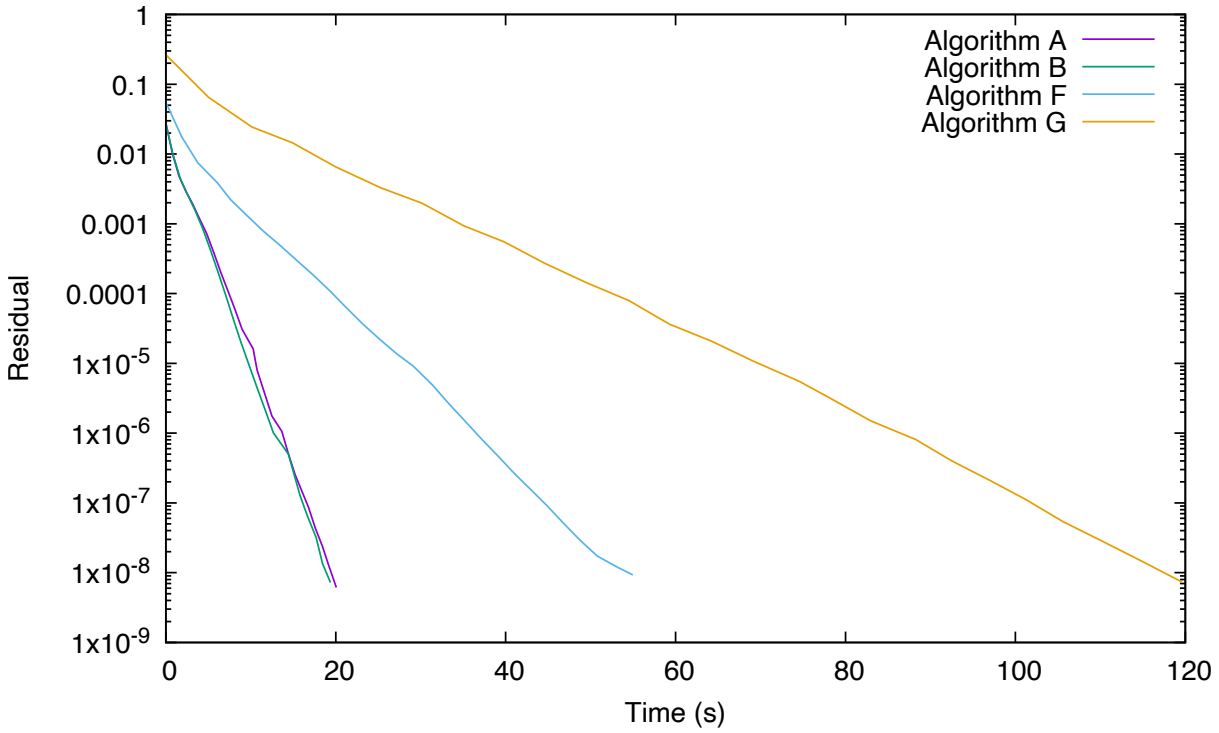


Zoom – Hybrid MdagM with Dw coarsening not crazy…



Time: so far most successful with BNL methods A,B





Physical point!

We have reoptimised the most successful of the algorithms studies on a larger system with lighter quark masses. We used a single $48^3 \times 96$ Mobius domain wall fermion configuration number 1000 from our $m_{ud} = 0.00074$ 2+1 flavour Iwasaki gauge ensemble at $\beta = 2.13$[29] with $L_s = 24$. The valence quark action was $L_s = 24$ but with the Shamir Domain Wall

Algorithm	Coarse Subspace	Coarse operator	Solve Time
rbCGNR	-	-	502s
H	110s	12.5s	195s
H (RB CG coarse)	110s	12.5s	149s
F (RB CG coarse)	50s	12.5s	280s

Can set up *and* solve twice to 10^-8 in 420s – faster than a single solve with the right baseline

Reusing subspace set up:

The Fermion determinant in domain wall Fermions is that of a ratio of the two flavour and Pauli Villars operators,

$$\det \frac{M_l^\dagger M_l}{M_{PV}^\dagger M_{PV}}$$

; normally factored as several intermediate Hasenbusch terms such as,

$$\det \frac{M(m_l)^\dagger M(m_l)}{M(m_1)^\dagger M(m_1)} . \det \frac{M(m_1)^\dagger M(m_1)}{M(m_2)^\dagger M(m_2)} . \det \frac{M(m_2)^\dagger M(m_2)}{M(m_{PV})^\dagger M(m_{PV})}$$

Alternate splitting may share subspace generation?

$$\det \frac{M(m_l)^\dagger M(m_l)}{M(m_l)^\dagger M(m_l) + \Delta_1} . \det \frac{M(m_l)^\dagger M(m_l) + \Delta_1}{M(m_l)^\dagger M(m_l) + \Delta_2} . \det \frac{M(m_l)^\dagger M(m_l) + \Delta_2}{M(m_{PV})^\dagger M(m_{PV})},$$

Conclusions

- (multiple) new DWF multigrid methods
  - Dw coarsening remains intriguing
  - Hdwf coarsening so far wins – Boston were a big distraction!

- Demonstrated cross over to modest net gain in HMC application
  - Compared to the correct baseline!

- Possible to use with Hasenbusch scheme