Algorithms for DWF

Peter Boyle (BNL),
Christopher Kelly (BNL),
Azusa Yamaguchi (University of Edinburgh)
RBC-UKQCD collaboration

- Domain Wall Multigrid
- Large domain DDHMC for GPUs
  - Goal: enable 2+1+1f simulations, 4GeV cut off

The RBC & UKQCD collaborations

UC Berkeley/LBNL
Aaron Meyer

BNL and BNL/RBRC
Yasumichi Aoki (KEK)
Peter Boyle (Edinburgh)
Taku Izubuchi
Yong-Chull Jang
Chulwoo Jung
Christopher Kelly
Meifeng Lin
Hiroshi Ohki
Shigemi Ohta (KEK)
Amarjit Soni

CERN
Andreas Jüttner (Southampton)

Columbia University
Norman Christ
Duo Guo
Yikai Huo
Yong-Chull Jang
Joseph Karpie
Bob Mawhinney
Ahmed Sheta
Bigeng Wang
Tianle Wang
Yidi Zhao

University of Connecticut
Tom Blum
Luchang Jin (RBRC)
Michael Riberdy
Masaaki Tomii

Edinburgh University
Matteo Di Carlo
Luigi Del Debbio
Felix Erben
Vera Gülpers
Tim Harris
Raoul Hodgson
Nelson Lachini
Michael Marshall
Fionn Ó hÓgáin
Antonin Portelli
James Richings
Azusa Yamaguchi
Andrew Z.N. Yong

KEK
Julien Frison

University of Liverpool
Nicolas Garron

Michigan State University
Dan Hoying

Milano Bicocca
Mattia Bruno

Peking University
Xu Feng

University of Regensburg
Davide Giusti
Christoph Lehner (BNL)

University of Siegen
Matthew Black
Oliver Witzel

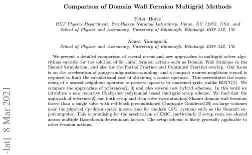
University of Southampton
Nils Asmussen
Alessandro Barone
Jonathan Flynn
Ryan Hill
Rajnandini Mukherjee
Chris Sachrajda

University of Southern Denmark
Tobias Tsang

Stony Brook University
Jun-Sik Yoo
Sergey Syritsyn (RBRC)

Domain Wall Multigrid

- Preprint: https://arxiv.org/pdf/2103.05034.pdf
- Spectrum of DWF makes coarsening nearest neighbour operator *hard*
    - Polynomial approximation to $\frac{1}{z}$ in region of complex plane enclosing origin
    - Typically solve normal equations on positive definite $M^\dagger M$
    - Nearest neighbour coarsenings of $\gamma_5 R_5 D_{dwf}$ (Herm, indefinite)
- Novel chebyshev polynomial setup of multigrid
- Result:
  Set up and solve twice D_{dwf} faster than red-black CG
- HMC focus; use compressed Lanczos for valence analysis



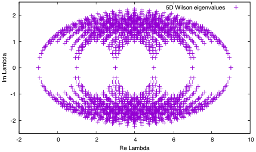
-lat] 8 Mar 2021

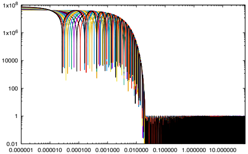
Comparison of Domain Wall Fermion Multigrid Methods

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

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 multigrid setup scheme. We find that the approach of reference[2], can both setup and then solve twice standard Shamir domain wall fermions faster than a single solve with red-black preconditioned Conjugate Gradients[29] on large volumes near the physical up/down quark masses 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.





Large domain DDHMC for GPUs

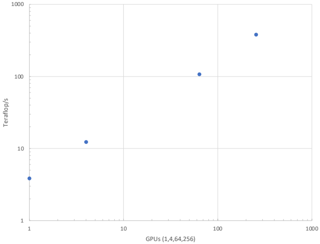
Motivation:

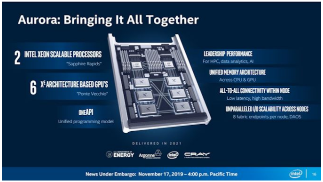
- GPU speed is increasing rapidly over time
- Interconnect speeds are *not* keeping pace.
- Expense spent on interconnect is significant
- If we get nearer commodity pricing, this will be even worse

- Project 30% efficiency on future systems
- GPU cache sizes are growing
  - DWF reuses gauge links L_s times
  - DWF reuses spinors $2N_d$ times
  - $\Rightarrow$ cache bound performance on a single node
- Interconnect will rapidly become bottleneck

System balance

System	GPUs	Node peak FP32	Node interconnect (GB/s Snd+Rcv)
Booster/Jülich	4 x A100	78TF/s	200GB/s
Tursa/Edinburgh	4 x A100	78TF/s	200GB/s
Summit	6 x V100	94TF/s	50GB/s
Aurora	6 x Intel Xe	$\geq$ 130TF/s fp64 $\geq$ 260TF/s fp32 (conjectured 2x)	300-400GB/s 300GB/s to GPU's?





Future machines will not scale with current algorithms

Scaling on Edinburgh Tursa / Jülich Booster

Nodes	GPUs	Measured Perf / GPU	Measured TF/s	Ideal GPU scaling	Ideal Node scaling
1	1	3.85	3.85	3.85	
1	4	3.075	12.3	15.4	12.3
16	64	1.671875	107	246.4	196.8
64	256	1.484375	380	985.6	787.2

DDHMC refresher

- hep-lat/0409106
- Partition the lattice into hypercuboids
- Colour them black and white according to parity
- Call "white" domain Ω and complement $\bar{\Omega}$
- Schur factoring the Fermion determinant leaves local and non-local terms that can be integrated on different timescales.

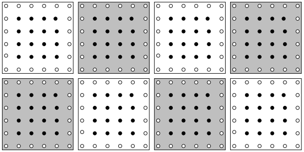
$$D = \begin{pmatrix} D_\Omega & D_\partial \\ D_{\bar{\partial}} & D_{\bar{\Omega}} \end{pmatrix}$$

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

Schwarz-preconditioned HMC algorithm
for two-flavour lattice QCD

Martin Lüscher

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



- *small* domains 4^4 to 6^4
- HMC MD integrate gauge action and local determinants for each domain without communication
- Fits within L2 cache of a CPU core
- Small cell provides IR regulator for Dirichlet Dirac solves
- Exterior boundary gauge links are frozen (cross domain and in surface plane)

Boundary determinant

- Handling the Schur complement "boundary" determinant requires care

$$\chi = 1 - D_{\Omega}^{-1} D_{\partial} D_{\bar{\Omega}}^{-1} D_{\bar{\partial}}$$

- Can restrict to exterior boundary of Ω

$$R = \mathbb{P}_{\bar{\partial}} - \mathbb{P}_{\bar{\partial}} D_{\Omega}^{-1} D_{\partial} D_{\bar{\Omega}}^{-1} D_{\bar{\partial}}$$

- because in the right basis χ takes the form

$$\chi = \begin{pmatrix} 1 - X & 0 \\ Y & 1 \end{pmatrix}$$

  so $\det \chi = \det R = \det(1 - X)$

- For pseudofermion action $\phi_{\bar{\partial}}^{\dagger} (RR^{\dagger})^{-1} \phi_{\bar{\partial}}$,

$$R^{-1} = \hat{\mathbb{P}}_{\bar{\partial}} - \hat{\mathbb{P}}_{\bar{\partial}} D^{-1} \hat{D}_{\bar{\partial}}$$

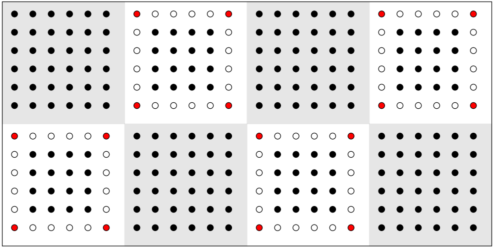
- $\delta R^{-1} = \mathbb{P}_{\bar{\partial}} D^{-1} \delta D D^{-1} D_{\bar{\partial}}$.
- Pauli-Villars (or Hasenbusch) requires

$$\phi_{\bar{\partial}}^{\dagger} P^{\dagger} R^{-\dagger} R^{-1} P \phi_{\bar{\partial}}.$$

  and $\delta R = \mathbb{P}_{\bar{\partial}} D_{\Omega}^{-1} (\delta D_{\Omega}) D_{\Omega}^{-1} D_{\partial} D_{\bar{\Omega}}^{-1} D_{\bar{\partial}} + \mathbb{P}_{\bar{\partial}} D_{\Omega}^{-1} D_{\partial} D_{\bar{\Omega}}^{-1} (\delta D_{\bar{\Omega}}) D_{\bar{\Omega}}^{-1} D_{\bar{\partial}}$.

Symmetric domain shapes

- Luscher's domain structure



- Boundary pseudofermion lives on the interior boundary of Ω
- Spin structured: sites on only one face are spin projected
  - Red dots are four component pseudofermion
  - Open dots are two component pseudofermion
- Schur Factoring around Ω or $\bar{\Omega}$ dictates where pseudofermion "lives"
  $\Rightarrow$ choice creates asymmetry in the forces on gauge links Ω or $\bar{\Omega}$
  - Easiest to see in 1D where all boundary sites are spin projected

Large domain DDHMC

- GPU's offer large parallelism within the node
  - 32^4 or greater subvolume per domain
  - Local solves can outstrip the network.
  - Node 10x(?) faster than Booster, network 1.5x faster (?)
  - Domain decompose HMC on *large* domains
- Cell local Dirichlet determinants are "obvious"
  - Create an adaptor for any Grid Fermion operator that zeroes gauge links, removes communication
  - Standard two flavour pseudofermion action otherwise.
  - Local determinant equally ill conditioned as light solve
    $\longrightarrow$ this is exactly what GPU's are good at!
- *Will also change subdomain shapes*

Domain Wall force

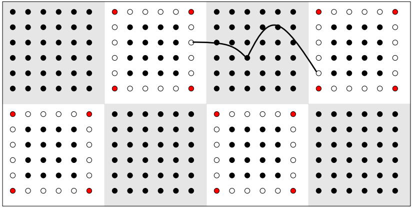
- Normal equations on 5D system uses single solve in force

$$\phi^\dagger (M^\dagger M)^{-1}\phi$$

  - Can also be used for local determinant
- Boundary projector means number of solves is doubled (normal equations twice)

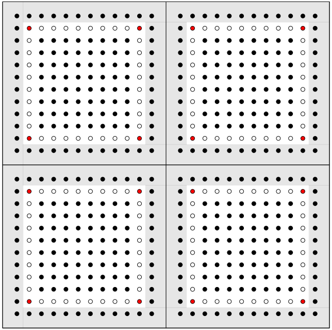
$$\delta\left(\phi^\dagger_{\bar\partial} R^{-\dagger} R^{-1}\phi_{\bar\partial}\right) = 2\mathrm{Re}\langle R^{-1}\phi_{\bar\partial}|\mathbb{P}_{\bar\partial} D^{-1}\delta D D^{-1} D_{\bar\partial}\phi_{\bar\partial}\rangle$$

- Must have a good integrator timestep ratio between local and boundary determinants
- Force is suppressed by two light quark propogators
  - Can suppress force arbitrarily by using a broader band of inactive links
  - Short distance propagator is not dictated by pion mass



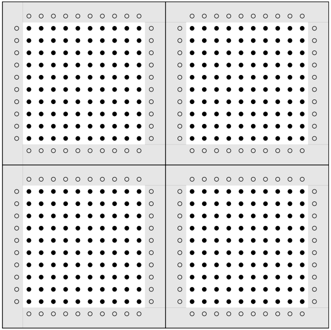
Non-symmetric domain shapes

$$S_{\text{Pseudofermion}} = \phi_\Omega^\dagger (D_\Omega^\dagger D_\Omega)^{-1} \phi_\Omega + \phi_{\bar\partial}^\dagger (R^\dagger R)^{-1} \phi_{\bar\partial}$$



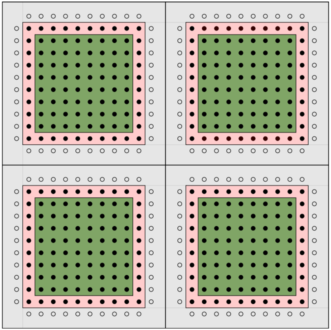
Large Domain DDHMC domain structure

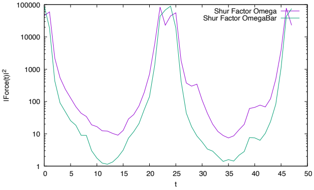
- Boundary pseudofermion lives on the interior boundary of Ω (or $\bar\Omega$)
- Spinor structure: sites on only one face are spin projected



- $\det D_\Omega$ is local to a node and maximally large
- Freeze all links in $\bar\Omega$, *do not need to compute* $\det D_{\bar\Omega}$

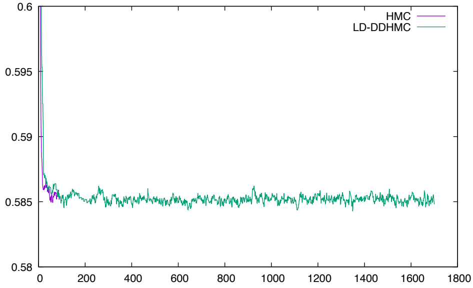
- HMC slow zone close to boundary
  - scale the HMC time evolution in a coordinate dependent way
  - power law "slow down" in red zone
  - Counterbalance rise in propagator





$16^3 \times 48$ 2f test

- DWF+Iwasaki 2 flavor: $\beta = 2.13, 16^3 \times 48, m_f = 0.01, L_s = 16$
- Produced on 2 GPUs
- 3:1 ratio of boundary determinant to local determinant timesteps
  - Omelyan integrator (2 force evaluations in nesting)
  - Adequate hierarchy in integration
- Wall clock gain depends on interconnect performance; believe this work is a substantial factor on Aurora
- Strange quark / odd flavours are a work in progress. May just use EOFA.



Clearly need more statistics on reference, but looks OK
Evolution is solid and plaquette in low stats agreement with

- 2f Grid run of same ensemble
- Plausibly close to historic 2+1f u/d/s plaquette

Summary and outlook

- Large domain DDHMC is ideal to decouple islands of high performance in future GPU systems
  - Conjecture up 10x acceleration of local domain solves (?)
  - Precise, algorithmically efficient determinant factorisation:
    SAP local multigrid smoother increases fine matrix multiplies vs. polynomial smoothers
- Can be combined with multigrid on the large domain, and on the full lattice solve
- Can also consider multilevel integration
  - no in-principle barrier for DWF
  - N^2 valence measurements
    $\Rightarrow$ For DWF need a better valence solver scheme as we have not yet achieved the Wilson multigrid speed-up
    N^2 Lanczos deflation prohibitive
- Fall of propagator with distance and computer architecture trends make this a guaranteed win in long run
- Expect it to win on Aurora
- Master field simulation proposals increase the need as we go to bigger volumes, fewer configurations.
  - Caution: must still thermalise and decorrelate, so running for many times the thermalisation time will always be sensible.
    Pre-thermalising on smaller volume; periodic replication; rethermalise may help address.