Algorithmic Challenges in Lattice QCD

Carsten Urbach[1,2]

[2]University of Liverpool

[1]John von Neumann-Institut für Computing NIC

ILFTN Workshop, JLAB, 2005

χLF Collaboration

T. Chiarappa, K. Jansen, K.-I. Nagai, M. Papinutto, L. Scorzato, A. Shindler, U. Wenger, I. Wetzorke

Outline

Motivation

- Maximally twisted mass formulation shares many properties with the overlap formulation, but not exact chiral symmetry:
  - What is the price for exact chiral symmetry on the lattice?
- Algorithms used to simulate full QCD suffer from a substantial slowing down when
  - quark masses get light
  - lattice spacing gets small
- Are stable simulations with Wilson fermions at light masses and small lattice spacings possible and realistic?

Notation

- Wilson-Dirac operator:

$$D_\mathrm{W} = \sum_\mu \frac{1}{2}\left\{\gamma_\mu(\nabla_\mu + \nabla^*_\mu) - a\nabla_\mu\nabla^*_\mu\right\}$$

- Twisted mass operator at maximal twist:

$$D_\mathrm{tm} = D_\mathrm{W} + m_\mathrm{crit} + i\mu\gamma_5\tau_3\,.$$

- Overlap operator ($\rho = 1.6$):

$$D_\mathrm{ov} = \left(1 - m_\mathrm{ov}\frac{a}{2\rho}\right) D_\mathrm{ov}^{(0)} + m_\mathrm{ov}\,,$$

with

$$D_\mathrm{ov}^{(0)} = \frac{\rho}{a}\left[1 - A(A^\dagger A)^{-1/2}\right], \qquad A = \rho - aD_\mathrm{W}\,.$$

Setup for the Cost Comparison

- Wilson plaquette gauge action, $\beta = 5.85$, quenched.
- Two volumes: 12^4 and 16^4 with 20 gauges each.
- sign function approximated by means of Chebyshev polynomials.
  20 (40) lowest EV of $A^\dagger A$ projected out.
- Mass values for cost comparison:

m_{PS} [MeV]	$a\mu$	am_{ov}
230	0.004	0.01
390	0.0125	0.03
555	0.025	0.06
720	0.042	0.10

Iterative solvers

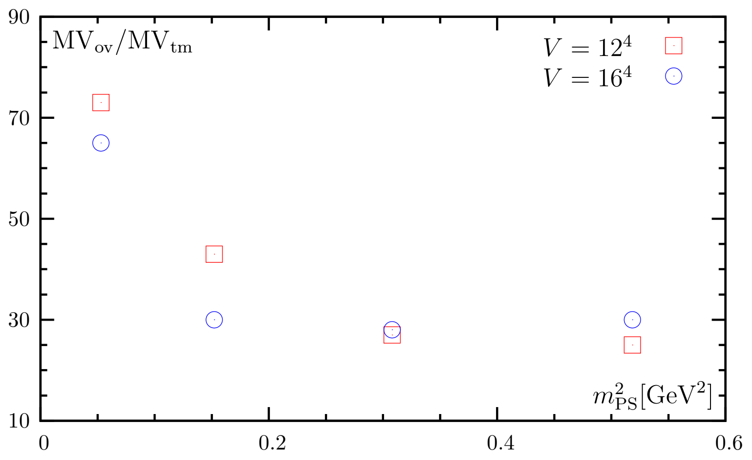
Solver	overlap	tmQCD
BiCGstab	X	X
CG(NE)	X	X
CGS	X	X
SUMR	X	
GMRES(20)	X	X
MR	X	

Best solver:
overlap: $\mathrm{GMRES}_{\mathrm{ap}}$
tm: $\mathrm{CG}_{\mathrm{EO}}$

Improvements:

- overlap: adaptive precision, chiral separation
- tm: even/odd preconditioning

Costs in matrix-vector (MV) multiplications



HMC Algorithm

(Duane, Kennedy, Pendleton, Roweth, 1987)

- Introduce traceless Hermitian momenta $P_{x,\mu}$ conjugate to fundamental fields $U_{x,\mu}$ and Hamiltonian

$$H = \frac{1}{2}\sum_{x,\mu} P_{x,\mu}^2 + S[U]\ .$$

- Molecular dynamics evolution of P and U by numerical integration of the corresponding equations of motion.
- Metropolis accept/reject step to correct for discretization errors of the numerical integration.

Accelerating the HMC algorithm

- Most expensive part: fermion determinant
- Precondition by factorization (with suitable C and E):

$$\det Q^2 = \det(C) \cdot \det(E)$$

  with C and E better "behaved" than Q^2.
- mass preconditioning (`Hasenbusch`), polynomial filtering (`Peardon, Sexton`), domain decomposition (`Lüscher`), n^{th}-root (`Clark, Kennedy`)
- whereas often:
  - C is cheap and
  - E is expensive to invert

Multiple Time Scale Integration

(Sexton, Weingarten, 1992)

- Assume: $H = \frac{1}{2}\sum_{x,\mu} P_{x,\mu}^2 + S_0 + S_1$
- Define ($j = 0, 1$):

$$T_{\mathrm{U}}(\Delta\tau): \quad U \quad \rightarrow \quad U' = \exp\left(i\Delta\tau P\right) U\,,$$
$$T_{\mathrm{S}_j}(\Delta\tau): \quad P \quad \rightarrow \quad P' = P - i\Delta\tau\delta S_j$$

- and recursively:

$$T_0 = T_{\mathrm{S}_0}(\Delta\tau_0/2)\; T_{\mathrm{U}}(\Delta\tau_0)\; T_{\mathrm{S}_0}(\Delta\tau_0/2)\,,$$
$$T_1 = T_{\mathrm{S}_1}(\Delta\tau_1/2)\; [T_0]^{N_0}\; T_{\mathrm{S}_1}(\Delta\tau_1/2)$$

- trajectory of length τ: $\left[T_1\right]^{N_1}$
- time steps must fulfill: $N_1 = \tau/\Delta\tau_1\,, \qquad N_0 = \Delta\tau_1/\Delta\tau_0$

Multiple Time Scale Integration

- Also Sexton-Weingarten (SW) improved scheme can be generalized to multiple time scales.
- SW impr. scheme is one particular version of so-called second order minimal (2MN) norm integration schemes. (`de Forcrand, Takaishi, 2005`)
  However, SW impr. scheme is close to optimal.
- Interchange of the order of momentum and gauge field updates reveals a speedup.

Mass Preconditioning

- Precondition the fermion determinant ($Q = \gamma_5 D_\mathrm{W}$, $n_f = 2$):

$$\det Q^2 = \det\left[Q^2 + \mu^2\right] \cdot \det\left[\frac{Q^2}{Q^2 + \mu^2}\right] .$$

(Hasenbusch, 2001)

- Corresponding effective action:

$$S_\mathrm{eff} = S_\mathrm{G} + \phi_1^\dagger \frac{1}{Q^2 + \mu^2} \phi_1 + \phi_2^\dagger \frac{Q^2 + \mu^2}{Q^2} \phi_2 = S_\mathrm{G} + S_{\mathrm{PF}_1} + S_{\mathrm{PF}_2} .$$

- Can be extended to $N_\mathrm{PF} > 2$ pseudo-fermion fields.
- Can be combined with even/odd preconditioning.

Mass Preconditioning

- Precondition the fermion determinant ($Q = \gamma_5 D_{\mathrm{W}}$, $n_f = 2$):

$$\det Q^2 = \det\left[Q^2 + \mu^2\right] \cdot \det\left[\frac{Q^2}{Q^2 + \mu^2}\right] .$$

- Original idea: Choose μ such that the condition numbers of $Q^2 + \mu^2$ and $Q^2/(Q^2 + \mu^2)$ are equal
- condition number: $K \quad \to \quad \sqrt{K}$ (Hasenbusch, Jansen, 2002, ALPHA, 2003)
- Pseudo-fermion forces are reduced $\quad \to \quad$ larger HMC step sizes (factor two (at large mass))
- Caveat: Q^2 must still be inverted.

Strategy

If possible, tune μ such that:

- the more expensive the computation of δS_{PF_j} is, the less it contributes to the total force
- different parts can be integrated on *different* time scales chosen according to their force magnitude.

$$\Delta\tau_j \|F_j\| = \mathrm{const}$$

- Remark: also variance and time dependence of F_j is of importance

Similar approaches: `Peardon, Sexton, 2003, QCDSF, 2003`

Simulation Set-up

- Wilson-Dirac operator, $r = 1$, bare mass m_0, $\kappa = (2m_0 + 8)^{-1}$
- Wilson plaquette gauge action with $\beta = 5.6$, $a \approx 0.08\ \mathrm{fm}$. on $24^3 \times 32$ lattices (cf. `Lüscher, 2004`)
- HMC with even/odd preconditioning and up to three pseudo fermion fields, trajectory length $\tau = 0.5$
- Used $r_0 = 0.5\ \mathrm{fm}$ to set the scale
- Reversibility violation in H: $\mathcal{O}(10^{-5})$
- Multiple time scale integration schemes:
  - Plain Leap Frog (LF) integration scheme
  - Sexton-Weingarten (SW) improved scheme

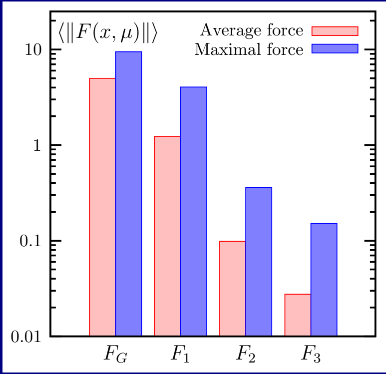
Simulation Points

	Int.	κ	m_{PS} [MeV]	a [fm]	$\tau_{\mathrm{int}}(P)$
A	SW	0.1575	665(17)	0.083(1)	6(2)
B	SW	0.1580	485(13)	0.081(1)	7(2)
C	LF	0.15825	380(17)	0.078(2) -	10(4)
D	SW	0.15835	294	–	15

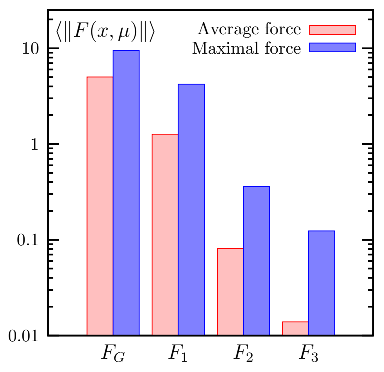
- Simulation point D still running, τ_{int} extrapolated
- Full agreement with `Lüscher, 2004` and `Orth et al., 2005`
- Acceptance around 80% for all runs.

Molecular Dynamics Forces

run *B* (485 MeV):

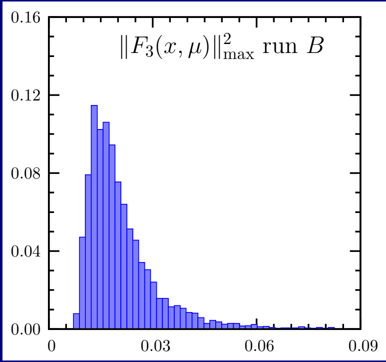


run *D* (294 MeV):

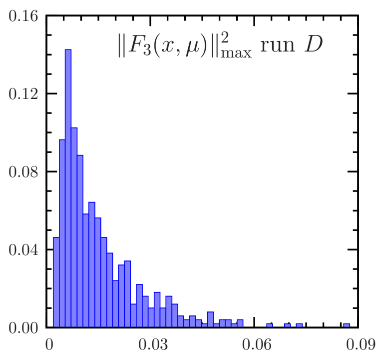


Molecular Dynamics Forces

run B (485 MeV):



run D (294 MeV):



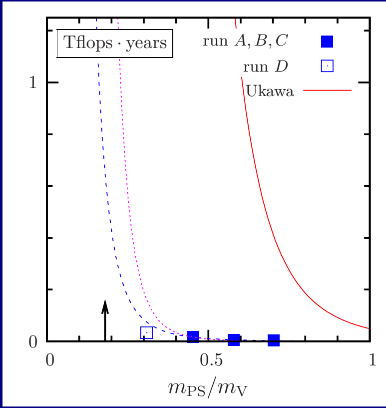
Performance

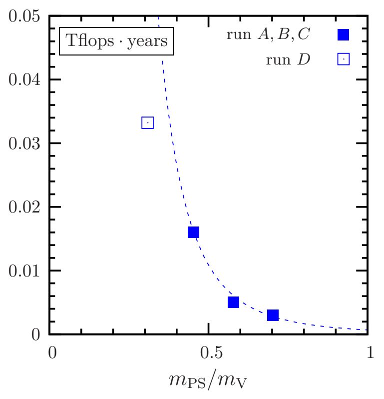
- Cost figure
$$\nu = 10^{-3}(2n+3)\tau_{\mathrm{int}}(P)$$
n: number of integration steps for physical operator
- Machine and implementation independent

	κ	ν	ν (Lüscher)	ν (Orth et al.)
A	0.15750	0.09(3)	0.69(29)	1.8(8)
B	0.15800	0.11(3)	0.50(17)	5.1(5)
C	0.15825	0.23(9)	0.62(23)	-
D	0.15835	0.35	0.74(18)	-

Simulation Cost

Cost for 1000 independent configurations,
$a \approx 0.08$ fm, $24^3 \times 40$ lattices.





How exactly does it work...?

Is the main effect

- the noise reduction due to the additional pseudo fermion fields?
- the infra-red regulation of the eigenvalue spectrum provided by the preconditioning?

A comprehensive understanding of the mechanism would possibly help us to further improve the HMC algorithm.

Conclusion

- Overlap versus twisted mass cost comparison (quenched): depending on the mass value overlap is a factor $20 - 70$ more expensive.
- HMC with a combination of mass preconditioning and multiple time scale integration:
  - Straightforward to implement and applicable to a wide variety of lattice Dirac operators.
  - Performance (in terms of the cost figure) comparable to SAP (`Lüscher, 2004`)
  - Stable simulations with $n_f = 2$ flavors of Wilson fermions are possible and affordable with $a \approx 0.08\ \mathrm{fm}$ and pseudo scalar masses as low as $300\ \mathrm{MeV}$.
  - Promising update of the Berlin wall figure (`Ukawa, 2002, Jansen, 2004`)

Outlook

- Dependence on the lattice spacing needs to be investigated
- Further improvements possible
  - Chronological Inverter (Brower et al., 1997)
  - PHMC (Frezzotti, Jansen, 1997, 1999) direction $n_f > 2$ flavours.