Hamiltonian MD in LQCD

Bálint Joó, Jefferson Lab

FASTMath-USQCD-SUPER Workshop

LLNL

Nov 10, 2011

Hybrid MDMC methods in LQCD

- MC Simulations are the only realistic way to 'solve' QCD.
- Need to update every degree of freedom (every lattice link)
- Most basic: update a link, and then Accept/Reject

$$P_{\rm acc}(U' \leftarrow U) = \min\left(1, e^{-\Delta S}\right) \;\; \Delta S = S(U') - S(U)$$

- Fermion determinant is a 'global' weight and involves solving a linear system:

$$\int \mathcal{D}U \; \det(\mathcal{M}) e^{-S_g(U)} = \int \mathcal{D}U \; d\phi \; d\phi^\dagger \; e^{-S_g(U) - \phi^\dagger \mathcal{M}^{-1} \phi}$$

- With O(100M) degrees of freedom updating links one by one is too costly.
- MD allows 'global updating' (all links at once). Energy conservation keeps acceptance rate high.

Hybrid Monte Carlo

- For each link $U_{\mu,x}$ pick conjugate momenta: $\pi_{\mu,x}$ from heatbath
- Hamiltonian

$$H = \frac{1}{2}\pi^2 + S(U)$$

- MD evolution for some 'time': propose U'
- Accept/Reject

$$P_{\text{acc}}(U' \leftarrow U) = \min\left(1, e^{-\Delta H}\right)$$

- If U' is rejected, U is next configuration
- For detailed balance: MD must be ***reversible and area preserving***
- Symplecticity is sufficient for area preservation.

S. Duane, A. D. Kennedy, B. J. Pendleton, D. Roweth, Phys. Lett. B. vol 195, No 2, 216-222

Typical MD Integrators in LQCD

- Time symmetric combinations of symplectic pieces:
  - Leap Frog:

$$U_{PQP}(\tau) = \left(e^{(1/2)\hat{S}\delta\tau} \; e^{\hat{T}\delta\tau} \; e^{(1/2)\hat{S}\delta\tau} \right)^{\tau/\delta\tau}$$

  - 2nd order Minimum Norm:

$$U_{PQP}(\tau) = \left(e^{\lambda\hat{S}\delta\tau} \; e^{(1/2)\hat{T}\delta\tau} \; e^{(1-2\lambda)\hat{S}\delta\tau} \; e^{(1/2)\hat{T}\delta\tau} \; e^{\lambda\hat{S}\delta\tau} \right)^{\tau/\delta\tau}$$

  - 'Truncation error' given by BCH expansion.

Higher Order Integrators

- Creutz, Gocksch, Campostrini: M. Creutz, A. Gocksch, Phys. Rev. Lett, 63, p 9-12, 1989
  - complicated series of forward/backward steps to cancel errors at 2nd order: M. Campostrini, P. Rossi, Nucl. Phys. B, 329:753, 1990

$$U_{2k+2}(\delta\tau) = U_{2k}(b_1\delta\tau)\ U_{2k}(b_2\delta\tau)\ U_{2k}(b_1\delta\tau)$$

$$b_1 = \frac{1}{2 - 2^{1/(2k+1)}}$$

$$b_2 = 1 - 2b_1$$

- Minimum Norm:
  - 4 & 5 Force evaluation variants
  - Tunable parameters: λ,ϱ,θ,μ

$$U(\delta\tau) = e^{\theta\delta\tau\hat{S}}\ e^{\rho\delta\tau\hat{T}} e^{\lambda\delta\tau\hat{S}}\ e^{\mu\delta\tau\hat{T}}\ e^{(1/2)(1-2(\lambda+\theta))\delta\tau\hat{S}}\ e^{(1-2(\mu+\rho)\delta\tau\hat{T}}\ e^{(1/2)(1-2(\lambda+\theta))\delta\tau\hat{S}}\ e^{\mu\delta\tau\hat{T}}\ e^{\lambda\delta\tau\hat{S}}\ e^{\rho\delta\tau\hat{T}}\ e^{\theta\delta\tau\hat{S}}$$

Tetsuya Takaishi[1] and Philippe de Forcrand

Phys. Rev. E 73, 036706 (2006)

Multiple Time Steps

- Consider action: S = S0 + S1

$$U_0(\delta\tau) = e^{(1/2)\delta\tau_0 \hat{S}_0} \; e^{\delta\tau_0 \hat{T}} \; e^{(1/2)\delta\tau_0 \hat{S}_0}$$

- Then to integrate S, define recursively:

$$U_1(\delta\tau_1) = e^{(1/2)\delta\tau_1 \hat{S}_1} \; U_0\left(\frac{\delta\tau_1}{N}\right)^N \; e^{(1/2)\delta\tau_1 \hat{S}_1}$$

- Two time scales: $\delta\tau_1$ and $\delta\tau_1/N$
- Generalize to multiple time-scales/levels
- Can nest more sophisticated integrators than just leapfrog

Shadow Hamiltonian

- These integrators actually conserve exactly a Shadow Hamiltonian.
- Defined as an asymptotic (Baker-Campbell-Haussdorff) expansion

$$\tilde{H}_{\mathrm{PQPQP}} = H + \left(\frac{6\lambda^2 - 6\lambda + 1}{12}\{S,\{S,T\}\} + \frac{1-6\lambda}{24}\{T,\{S,T\}\}\right)\delta\tau^2 + \left(\frac{-1 + 30\lambda^2 - 60\lambda^3 + 30\lambda^4}{720}\right.$$
$$\times\{S,\{S,\{S,\{S,T\}\}\}\} + \frac{-4 + 15\lambda + 15\lambda^2 - 30\lambda^3}{720}\{T,\{S,\{S,\{S,T\}\}\}\} + \frac{-7 + 30\lambda}{1440}$$
$$\times\{T,\{T,\{S,\{S,T\}\}\}\} + \frac{-7 + 30\lambda}{5760}\{T,\{T,\{T,\{S,T\}\}\}\} + \frac{-2 + 15\lambda - 35\lambda^2 + 30\lambda^3}{240}$$
$$\left.\times\{\{S,T\},\{S,\{S,T\}\}\} + \frac{-2 + 15\lambda - 30\lambda^2}{720}\{\{S,T\},\{T,\{S,T\}\}\}\right)\delta\tau^4 + \mathcal{O}(\delta\tau^6).$$

- Can measure 'Poisson Bracket' terms: {S,{S,T}}, {S, {S, {S, {S,T}}}} etc.
- BCH expansion generalizes to multi-level & higher order integrators (but gets very complicated)

Integrator Tuning

- Measure Poisson Bracket commutators
- Minimize Distance between H and its shadow, ie adjust step sizes, and λ-s so as to minimise

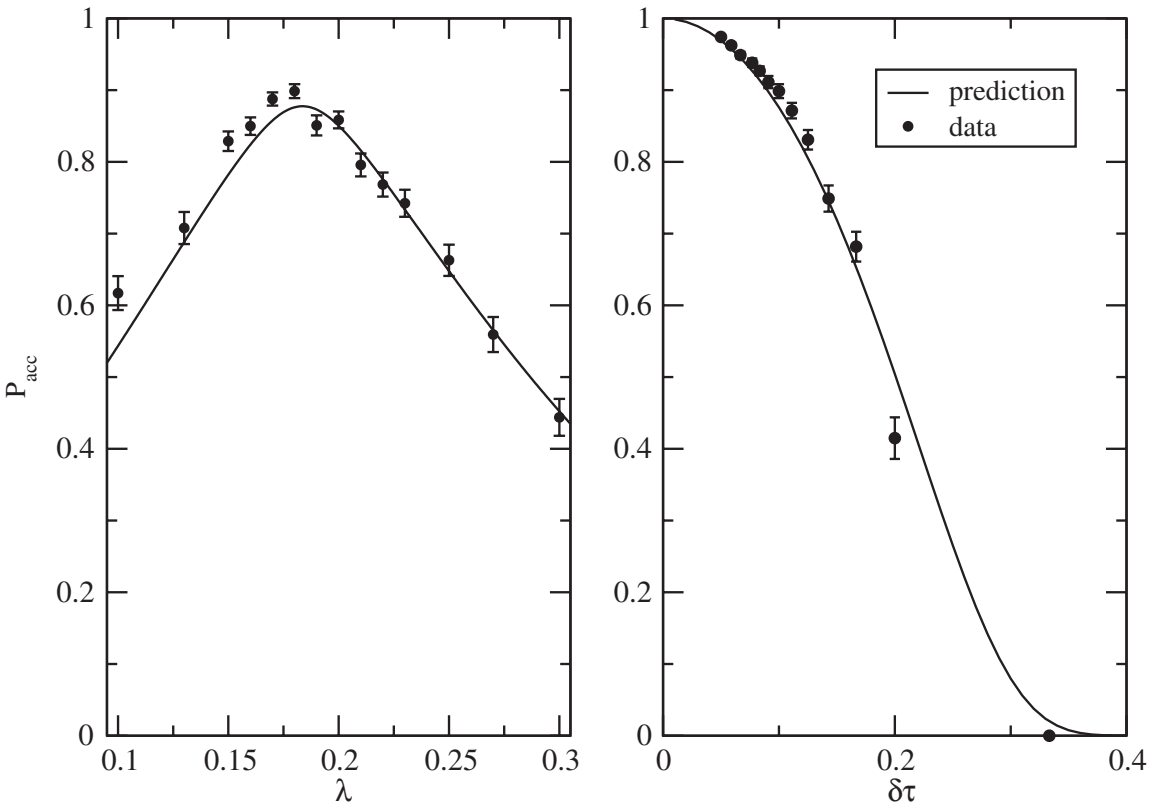
$$\sigma^2(\Delta H),\ \Delta H = \tilde{H} - H$$

- On average, this should be the closely related to (same as?)

$$\sigma^2(\delta H),\ \delta H = H(U', \pi') - H(U, \pi)$$

- PBs in LQCD are extensive, and should be fairly stable over a simulation (not like MD where they can fluctuate a lot)

Recent Results



- PB's allow prediction of acceptance rate as a function of step-size and tuning parameters
- A more systematic way of tuning than 'balancing forces'
- PBs allowed to improve an already well tuned set of parameters...

M. A. Clark[1], Bálint Joó[2], A. D. Kennedy[3], and P. J. Silva[4,]

Phys. Rev. D 84, 071502(R) (2011)

Force Gradient Integrator

- In MN Integrator choose λ=1/6, kill {T,{S,T}} term at O($\delta\tau^2$)
- Leaves {S,{S,T}} term at O($\delta\tau^2$) - so called Force Gradient term
- Define Force Gradient Integrator

$$U_{PQP}(\delta\tau) = e^{(1/6)\hat{S}\delta\tau}\, e^{(1/2)\hat{T}\delta\tau}\, e^{48\hat{S}\delta\tau - \{S,\widehat{\{S,T\}}\}\delta\tau^3/72}\, e^{(1/2)\hat{T}\delta\tau}\, e^{(1/6)\hat{S}\delta\tau}$$

- This is a bona fide 4th order integrator
  - Campostrini: ($\delta\tau^4$ / 34560) x (commutator terms)
  - Force Gradient: ($\delta\tau^4$ / 155520) x (commutator terms)
- FG term ~ 2 force calculations (force + force of force), 5 in total
- Not much back tracking :-)

Hantao's Trick...

- Implementation: follow Horner scheme $a + bx + cx^2 = a + x(b + cx)$
- The force gradient step is similar

$$p_i \longleftarrow p_i - \frac{2}{3}\tau e_i(S) + \frac{1}{36}\tau^3 e^j(S) e_j e_i(S)$$

- rewrite as $(F^j = e^j(S))$

$$p_i \longleftarrow p_i - \frac{2}{3}\tau \left(1 - \frac{1}{24}\tau^2 F^j e_j\right) e_i(S)$$

- approximate it by (using Taylor expansion)

$$-\frac{2}{3}\tau \left(1 - \frac{1}{24}\tau^2 F^j e_j\right) e_i(S) = -\frac{2}{3}\tau \exp\left(-\frac{\tau^2}{24} F^j e_j\right) e_i(S) + \mathcal{O}(\tau^5)$$

- $\exp\left(-\frac{\tau^2}{24} F^j e_j\right) S[U] = S\left[e^{-\frac{\tau^2}{24} F^j T_j} U\right]$

reproduced from from Hantao Yin's, Lattice 2011 contribution

Scaling Behaviour of FG for DWF

Tested with the following integration scheme on a $16^3 \times 32 \times 16$ lattice with 2+1 flavor DWF simulations, $m_\pi = 420$MeV.

$$\underbrace{\det\left(\frac{m_l}{m_s}\right)}_{\text{top level}} \underbrace{\det\left(\frac{m_s}{1}\right)^{1/2} \det\left(\frac{m_s}{1}\right)^{1/2} \det\left(\frac{m_s}{1}\right)^{1/2}}_{\text{2nd level}} + \underbrace{\text{gauge field}}_{\text{3rd level}}$$

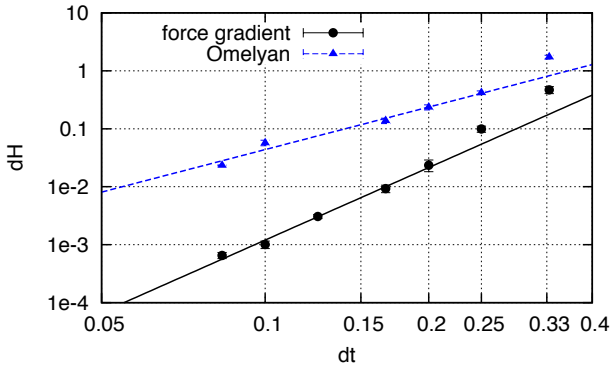


Figure: Scaling behavior
Omelyan 2.44 ± 0.21
Force Gradient 4.16 ± 0.21

- With force gradient, the top level step size can be increased to 1/3 compared with 1/4 in Omelyan.
- With the first and last updaing step combined, force gradient PQPQP requires 3 CG inversion while Omelyan requires 2 each step.

reproduced from from Hantao Yin's, Lattice 2011 contribution

Costs: MD times

We are using this new integration scheme on our production job.
The following shows a $16^3 \times 8 \times 48$ lattice with 160MeV pion

integrator	λ	step size	acc ratio	MD time(s)
Original	0.22	1/5	76/112=0.70	2.92e3
FG QPQPQ	N/A	1/7	75/88=0.85	1.95e3
Ome QPQPQ(top)	0.22	1/7	81/104=0.78	1.65e3
Ome QPQPQ(all)	0.22	1/7	88/122=0.72	1.60e3

- Except the original scheme, which has CG stopping condition 1e-8, all the rest have CG stopping condition 1e-6 for the quotient actions.

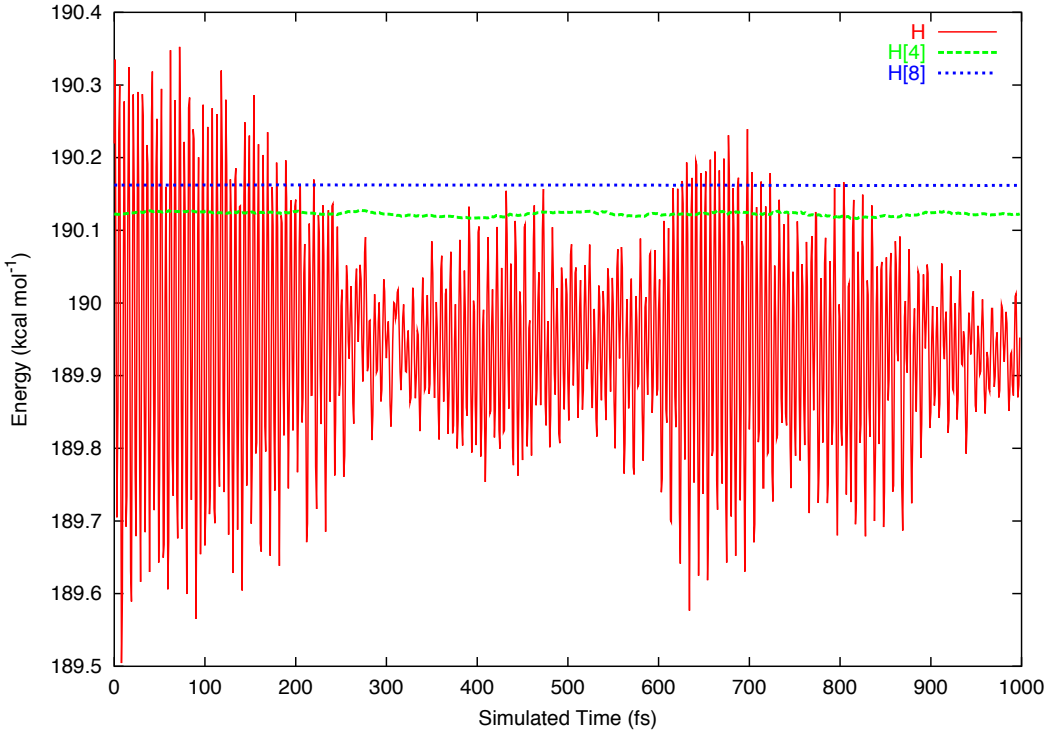
reproduced from from Hantao Yin's, Lattice 2011 contribution

FG Current Status

- In current tests, extra inversions cost in FG term could not be ameliorated by large enough step size
  - Well tuned 2nd order integrators were more efficient
  - but lattice sizes were quite small
  - Scaling:
    - $O(V^{5/4})$ - for 2nd order scheme
    - $O(V^{9/8})$ - for 4th order scheme
    - For sufficiently large lattice FG should win
    - In current study in the PRD paper, this volume is 56^4 sites
  - May already work for our large $40^3 \times 256$ lattices (?)

Shadow Hybrid Monte Carlo (SHMC)

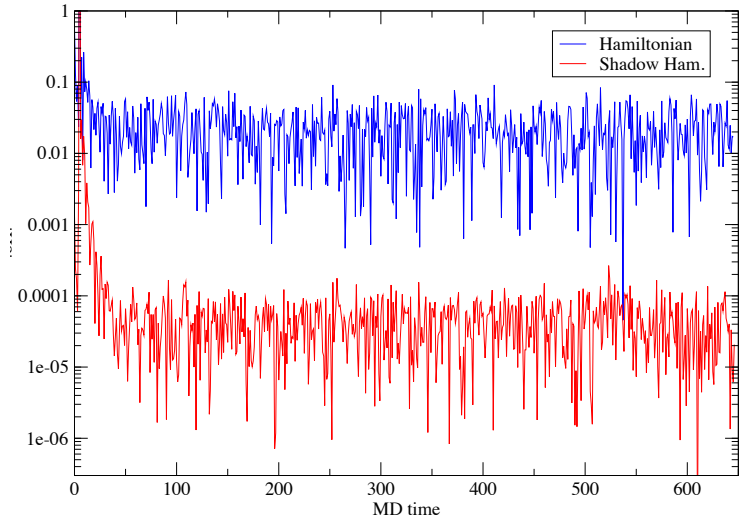
- Proposed in Material Science
- In HMC: accept/reject with approximation to H_s
  - Now need to accept/reject momenta
  - Do MD using H
    - preserves H_s exactly
- Reweight with exp(-(H-H_s))
- Motivation in Material Science:
  - Real Hamiltonian
  - Lumpy phase space
  - H_s much more stable than H, easier to tune

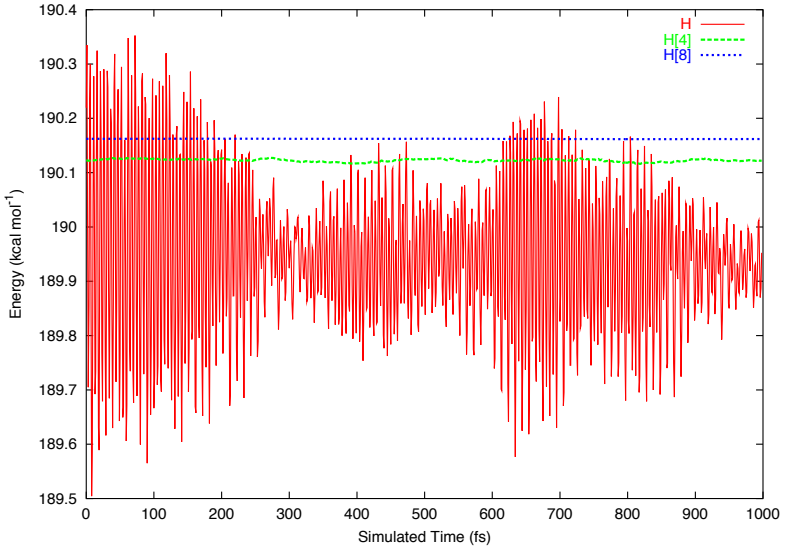


JA Izaguirre, S. Hampton
Journal of Computational Physics, vol 200 (2004), p 581-604

QCD and SHMC

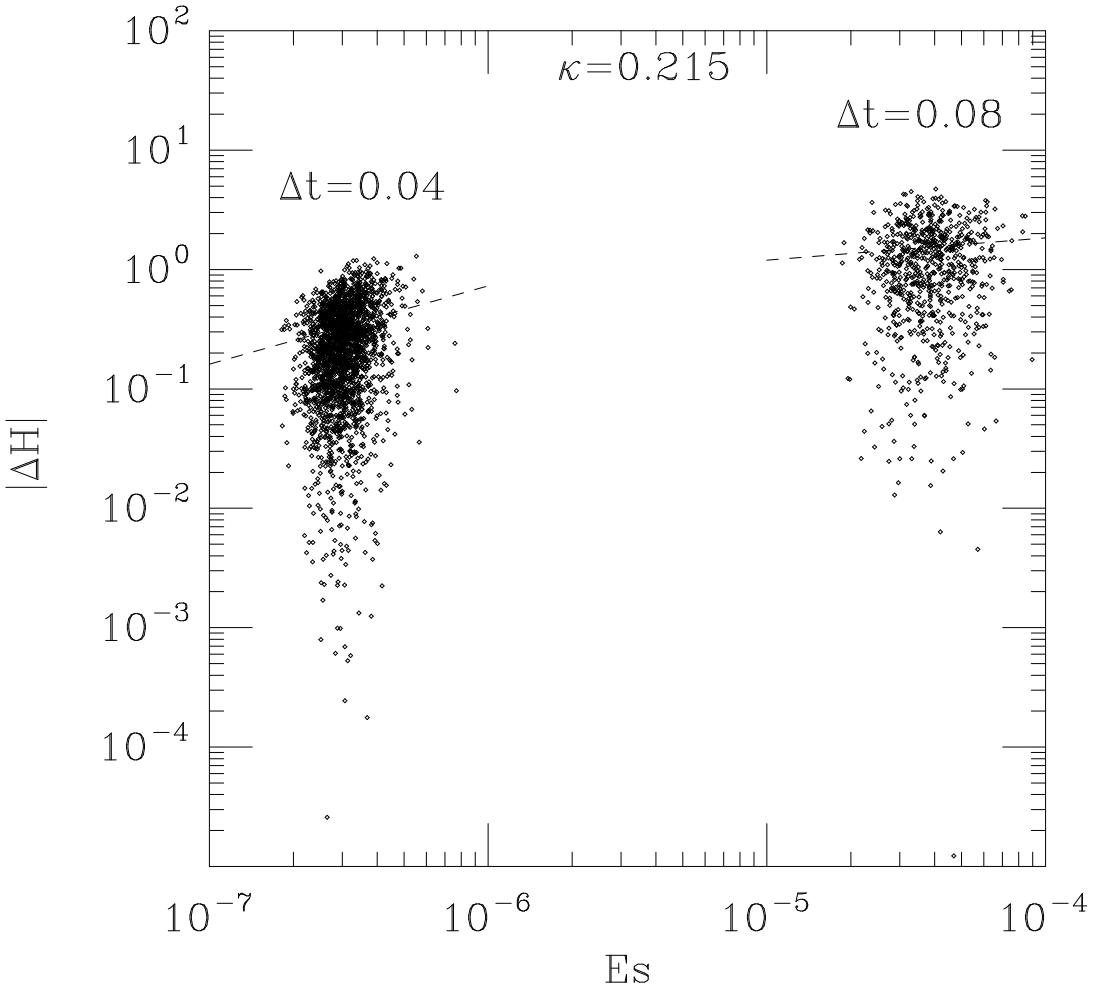
- Can construct truncated Hs using PBs
  - In material science forward/ backward timestepping was used
- In QCD δH (and ΔH) is quite stable (unlike material science) since PBs are extensive
- In terms of PBs
  - tuning MD in HMC <=> tuning reweighting in SHMC
- Would there be any gain?





Adaptive Step Size

- Adaptive step size integrators are not symplectic usually. Need to be very careful not to loose area preservation.
- One particular technique tried by deForcrand and Takaishi.
- Algorithm by Stoffer, time symmetric step-size controller
- Problem: the step-size controller Es was not strongly correlated with ΔH which controls acceptance
- Overheads for variable step size method outweighed benefits compared to a fixed step-size at the same acceptance rate



Philippe de Forcrand and Tetsuya Takaishi

Phys. Rev. E 55, 3658–3663 (1997)

Perspectives

- Domain Science perspectives
  - Currently explored in Wilson and DWF systems
  - For production use, will need PBs for Wilson-Clover quarks
  - Test FG in large volume runs
  - Use PBs for Shadow HMC?

- Applied Math perspectives (conversation with Emil)
  - Stability & Stiffness Analysis, algorithm optimization
  - Variable time stepping with time symmetric controls
    - improve on the works of deForcrand and Takaishi
    - time step changing can provide insight on when to use higher order methods

Primaries

- External (to USQCD) collaborators: (originators)
  - Mike Clark (NVIDIA),
  - Tony Kennedy (University Edinburgh)
  - Paulo Silva (University of Coimbra)
- In USQCD
  - Balint Joo (PBs & FG with Chroma, for Clover)
  - Hantao Yin (FG for DWF in CPS?)
  - Kostas Orginos (Shadow Hamiltonian algorithms)
  - James Osborn, Rich Brower
- FastMATH
  - Emil Constantinescu (ANL) - stability and stiffness analysis, improvements on variable step size techniques, tuning