Better HMC integrators for dynamical simulations

M. A. Clark
Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, U.S.A.
E-mail: `mikec@seas.harvard.edu`

Bálint Joó
Jefferson Lab, 12000 Jefferson Avenue, Newport News, VA 23606, U.S.A.
E-mail: `bjoo@jlab.org`

A. D. Kennedy
SUPA, School of Physics & Astronomy, The University of Edinburgh, Edinburgh EH9 3JZ,
Scotland, U.K.
E-mail: `adk@ph.ed.ac.uk`

P. J. Silva
Centro de Física Computacional, Universidade de Coimbra, Portugal
E-mail: `psilva@teor.fis.uc.pt`

We show how to improve the molecular dynamics step of Hybrid Monte Carlo, both by tuning the integrator using Poisson brackets measurements and by the use of force gradient integrators. We present results for moderate lattice sizes.

arXiv:1011.0230v1 [hep-lat] 1 Nov 2010

1. Introduction and Motivation

Hybrid Monte Carlo (HMC) [1] is the algorithm of choice to generate lattice configurations including the effect of dynamical fermions. Nevertheless, the generation of gauge field configurations at large volumes and light quark masses is still very expensive computationally.

One principal ingredient of HMC is the molecular dynamics (MD) step, which consists of a reversible volume-preserving approximate MD trajectory of $\tau/\delta\tau$ steps (with τ being the length of the trajectory in suitable units, and $\delta\tau$ the step size) followed by a Metropolis accept/reject test with acceptance probability $\min(1, e^{-\delta H})$ where δH is the change in the Hamiltonian H over the trajectory.

A molecular dynamics trajectory is not only an approximate integral curve of the Hamiltonian vector field $\hat{H}$ corresponding to H, but is also an exact integral curve of the Hamiltonian vector field $\widehat{\tilde{H}}$ of an exactly conserved Shadow Hamiltonian $\tilde{H}$. The asymptotic expansion of this Shadow Hamiltonian in the step size $\delta\tau$ may be computed using the Baker–Campbell–Hausdorff (BCH) formula and expressed in terms of Poisson brackets [2].

As a simple example consider a single level Omelyan (PQPQP) integrator [3]

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

whose shadow Hamiltonian is

$$\tilde{H}_{\text{PQPQPQ}} = H_{\text{PQPQPQ}} + \left(\frac{6\lambda^2 - 6\lambda + 1}{12}\{S,\{S,T\}\} + \frac{1-6\lambda}{24}\{T,\{S,T\}\} \right) \delta\tau^2 + O(\delta\tau^4). \quad (1.1)$$

Note that we have one free tunable parameter, λ, which is often set to some *ad hoc* value not taking Poisson brackets into account [4].

2. Integrator tuning

We define the difference between the shadow ($\tilde{H}$) and actual (H) Hamiltonians as $\Delta H = \tilde{H} - H$. We have previously suggested [5] that $\frac{1}{2}\langle\delta H^2\rangle \approx \text{Var}(\Delta H)$, where the right hand side is the variance of the distribution of values of ΔH over phase space. This formula assumes that the trajectories are long enough that the end points are more-or-less independent of the starting points, and accurate enough that their distribution is still close to e^{-H}. We can therefore estimate the acceptance rate from $\text{Var}(\Delta H)$

$$P_{\text{acc}} = \text{erfc}\left(\sqrt{\frac{1}{8}\langle\delta H^2\rangle} \right) = \text{erfc}\left(\sqrt{\frac{1}{4}\text{Var}(\Delta H)} \right) \quad (2.1)$$

The advantage of using $\text{Var}(\Delta H)$ is that one only needs to measure the Poisson brackets from equilibrated configurations. We can thus express P_{acc} as a function of the integrator parameters and find their optimal values that maximize the acceptance rate.

As a simple test, we consider a HMC simulation of two flavors of Wilson fermions at $\kappa = 0.158$ and Wilson gauge action at $\beta = 5.6$ on an 8^4 lattice. We use a single level PQPQP integrator and a unit trajectory length, therefore we have two parameters to tune, namely λ and the step size $\delta\tau$. In Figure 1 we compare the acceptance rates predicted by the formula above (red curve) with
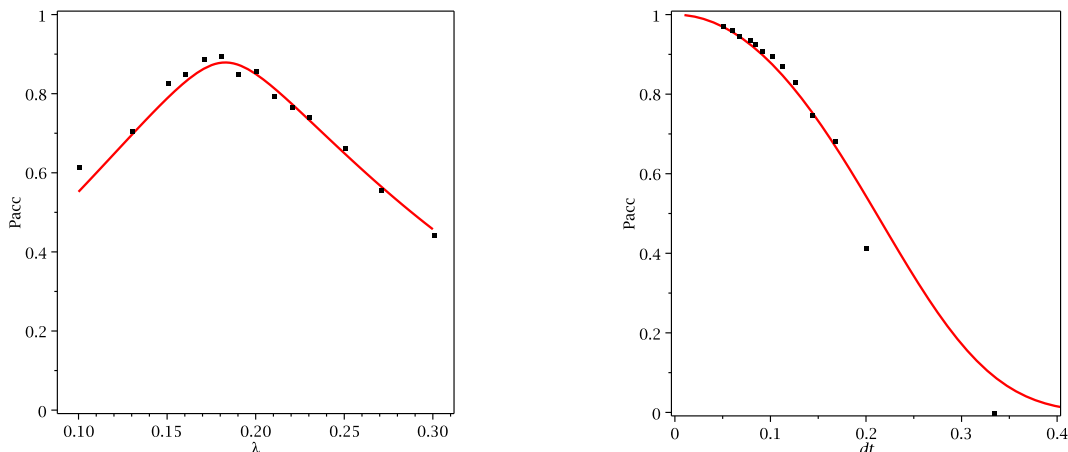
(a) Acceptance rate as a function of λ, with $\delta\tau = 0.1$.

(b) Acceptance rate as a function of $\delta\tau$, with $\lambda = 0.18$.

Figure 1: Comparison of measured acceptance rates and their predictions from average Poisson brackets.

numerical data taken from simulations at various values of λ and $\delta\tau$ (black dots). The Poisson Bracket values used for the predictions were measured at $\lambda = 0.18$ and $\delta\tau = 0.1$. In Figure 1(a) we have fixed $\delta\tau = 0.1$ and we leave λ as a free parameter; whereas in Figure 1(b) we take $\lambda = 0.18$ and we plot P_{acc} as a function of the step size.

The figures show good agreement between predicted and measured acceptance rates. We now use eq. (2.1) to tune the MD integrator on a larger volume. Ultimately, we are interested in reducing the computational cost, which depends on the wall-clock time spent computing the force terms as well as the acceptance rate, and the autocorrelation time for the observables. We neglect the autocorrelation time in this discussion as they are not sensitive to the choice of integrator parameters as long as the acceptance rate is reasonable, and define our cost metric as

$$\text{cost} = \frac{\text{trajectory CPU time}}{P_{\text{acc}}} .$$

For a nested integrator the numerator of this cost function is a function of the number of steps at each level times the CPU time required to compute the forces at that level.

3. Tuning a real simulation

As an application of our tuning technology, we are going to consider a HMC simulation of a $24^3 \times 32$ lattice [6], with two flavours of Wilson fermions with $\kappa = 0.1580$ and $\beta = 5.6$. The authors of [6] include two Hasenbusch fields with twisted mass fermions as "preconditioners", and they use a nested PQPQP integrator scheme with $\lambda = 1/6$ and one force term at each level, as shown in Table 1, where 0 is the outermost level. For each level i, we show the corresponding number of steps m_i, the type of force and its parameters, and typical times spent on force and force gradient computation. All times refer to runs on 128 nodes on the BlueGene/L at the University of Edinburgh.

Level i	Steps m_i	Force	F time	FG time
0	3	Hasenbusch ($\mu = 0$ / $\mu = 0.057$)	37 s	43 s
1	1	Hasenbusch ($\mu = 0.057$ / $\mu = 0.25$)	9 s	12 s
2	2	Wilson ($\mu = 0.25$)	2.5 s	3.3 s
3	3	Gauge	0.25 s	0.34 s

Table 1: Set-up used in the HMC simulation described in [6], together with typical times spent on force computation. For convenience times for the force gradient computation used in §3.2 are also shown here.

For simplicity, we have fixed $\tau = 1$ in our tuning exercise. We also have restricted our search space to $m_0 \geq 3$ to avoid integrator instabilities (which occur when the BCH expansion breaks down).

3.1 PQPQP tuning

We now describe how we tuned the PQPQP integrator, and what results we obtained. We considered two different nested schemes:

A1. The original scheme, but with tuned values of λ.

B1. The two Hasenbusch fields appear now at the same level (so we have only 3 different levels).

Table 2 shows the parameters which minimize the cost metric defined in the last section. For each scheme we show the number of steps at each level, the optimal λ parameters, our predictions for the acceptance rate, the estimated time spent in force computation in one trajectory, and measurements of acceptance rates and trajectory times. For comparison we also show data for the original scheme.

Scheme	m_i				λ_i				Prediction		Measurement	
	0	1	2	3	0	1	2	3	P_{acc}	F time / traj.	P_{acc}	Time / traj.
Original	3	1	2	3	1/6	1/6	1/6	1/6	0.85	655 s	0.89	709 s
A1	3	1	1	2	0.185	0.188	0.184	0.183	0.80	541 s	0.75	578 s
B1	3	3	1	–	0.177	0.183	0.176	—	0.82	454 s	0.83	554 s

Table 2: Tuning of the PQPQP integrator scheme.

We see that both the tuning of the integrator parameters and changes to the scheme provide further improvements over an already well-tuned scheme. Indeed, with the B1 scheme we get a $1.3\times$ speedup.

3.2 Force gradient integrator tuning

We have again considered the two integrator schemes above, but used a PQPQP force gradient integrator [7] at all levels. As in this case there are no tuneable parameters, we could only vary the number of steps at each level. In Table 3 we show the best parameters we found as well as the measured values of acceptance rates and trajectory times.

Scheme	m_i				Prediction		Measurement	
	0	1	2	3	P_{acc}	$F+FG$ time / traj.	P_{acc}	Time / traj.
Original	3	1	2	3	0.85	655 s	0.89	709 s
A2	3	1	1	1	0.96	778 s	0.91	814 s
B2	3	1	1	–	0.91	565 s	0.78	626 s

Table 3: Tuning of the force gradient integrator scheme.

In this case, the acceptance rates are not sufficiently high to amortize the higher wall-clock time cost coming from the computation of the force gradient term. Thus overall, we could see no measurable improvement in the cost metric as compared to the regular PQPQP case. However, further improvement could be possible by either using other force gradient integrators [3], or tuning the Hasenbusch masses. We are currently working on these issues.

4. Conclusions

We have presented a novel way of tuning an integrator, together with a practical example using a moderate lattice size. This tuning procedure can be used for all lattice gauge and fermionic actions. We are working towards a general implementation of the calculation of Poisson brackets and force gradient terms in Chroma [8]. In the near future we will consider the tuning of HMC simulations on larger lattices and smaller quark masses, and we will also consider other widely used lattice actions.

Acknowledgments

The numerical results have been obtained using Chroma library [8].

References

[1] S. Duane, A. D. Kennedy, B. J. Pendleton and D. Roweth, *Phys. Lett. B* **195** (1987) 216.

[2] A. D. Kennedy and M. A. Clark, PoS(LATTICE 2007)038, arXiv:0710.3611 [hep-lat].

[3] I. P. Omelyan, I. M. Mryglod and R. Folk, *Comp. Phys. Comm.* **151** (2003) 272.

[4] Tetsuya Takaishi and Philippe de Forcrand, *Phys. Rev.* **E73** (2006) 036706.

[5] M. A. Clark, A. D. Kennedy and P. J. Silva, `PoS(LATTICE 2008)041`, `arXiv:0810.1315 [hep-lat]`.

[6] C. Urbach, K. Jansen, A. Shindler and U. Wenger. *Comp. Phys. Comm.* **174** (2005) 87.

[7] A. D. Kennedy, M. A. Clark and P. J. Silva, `PoS(LAT2009)021`, `arXiv:0910.2950 [hep-lat]`.

[8] R. G. Edwards and B. Joo (SciDAC Collaboration and LHPC Collaboration and UKQCD Collaboration), *Nucl. Phys. Proc. Suppl.* **140** (2005) 832.