Tuning HMC using Poisson Brackets

M. A. Clark[1], A. D. Kennedy[2], and P. J. Silva[2]

[1]Centre for Computational Sciences
Boston University
mikec@bu.edu

[2]SUPA & School of Physics, The University of Edinburgh
psilva@ph.ed.ac.uk, adk@ph.ed.ac.uk

We demonstrate how measurement of the average values of the contributions to the Poisson brackets $\{S,\{S,T\}\}$ and $\{T,\{S,T\}\}$ allow us to optimize the integrators used for generating dynamical fermion configurations.

Theory

Symplectic Integrators

We are interested in finding the classical trajectory in phase space of a system described by the Hamiltonian $H(q,p) = T(p) + S(q) = \frac{1}{2}p^2 + S(q)$. The idea of a *symplectic integrator* is to write the evolution operator as $\exp\left(\tau\frac{d}{dt}\right) = \exp\left(\tau\left\{\frac{dp}{dt}\frac{\partial}{\partial p} + \frac{dq}{dt}\frac{\partial}{\partial q}\right\}\right) \equiv e^{\tau\hat{H}}$ where the *vector field*

$$\hat{H} = -\frac{\partial H}{\partial q}\frac{\partial}{\partial p} + \frac{\partial H}{\partial p}\frac{\partial}{\partial q} = -S'(q)\frac{\partial}{\partial p} + T'(p)\frac{\partial}{\partial q} \equiv \hat{S} + \hat{T}.$$

Since the kinetic energy T is a function only of p and the potential energy S is a function only of q it follows that the action of $e^{\tau\hat{S}} : f(q,p) \mapsto f(q, p - \tau S'(q))$ and $e^{\tau\hat{T}} : f(q,p) \mapsto f(q + \tau T'(p), p)$ are just translations of the appropriate variable.

We now make use of the Baker–Campbell–Hausdorff (BCH) formula, which tells us that the product of exponentials in any associative algebra can be written as $\ln(e^{A/2}e^{B}e^{A/2}) - (A+B) = \frac{1}{24}\left\{[A,[A,B]] - 2[B,[A,B]]\right\} + \cdots$ where all the terms on the right hand side are constructed out of commutators of A and B with known coefficients. We find that for a simple PQP symmetric integrator with step size $\delta\tau$ the evolution operator for a trajectory of length τ may be written as

$$\begin{aligned} U_{\text{PQP}}(\delta\tau)^{\tau/\delta\tau} &= \left(e^{\frac{1}{2}\delta\tau\hat{S}}e^{\delta\tau\hat{T}}e^{\frac{1}{2}\delta\tau\hat{S}}\right)^{\tau/\delta\tau} \\ &= \left(\exp\left[(\hat{T}+\hat{S})\delta\tau - \tfrac{1}{24}\left([\hat{S},[\hat{S},\hat{T}]] + 2[\hat{T},[\hat{S},\hat{T}]]\right)\delta\tau^3 + \mathcal{O}(\delta\tau^5)\right]\right)^{\tau/\delta\tau} \\ &= \exp\left[\tau\left(\hat{T}+\hat{S} - \tfrac{1}{24}\left([\hat{S},[\hat{S},\hat{T}]] + 2[\hat{T},[\hat{S},\hat{T}]]\right)\delta\tau^2 + \mathcal{O}(\delta\tau^4)\right)\right]. \end{aligned}$$

Shadow Hamiltonians

For every symplectic integrator there is a *shadow Hamiltonian* $\tilde{H}$ that is exactly conserved; this may be obtained by replacing the commutators $[\hat{S},\hat{T}]$ in the BCH expansion with the *Poisson bracket* $\{S,T\} \equiv \frac{\partial S}{\partial p}\frac{\partial T}{\partial q} - \frac{\partial S}{\partial q}\frac{\partial T}{\partial p}$ [1]. For example, the integrator above exactly conserves the shadow Hamiltonian $\tilde{H} \equiv T + S - \frac{1}{24}\left(\{S,\{S,T\}\} + 2\{T,\{S,T\}\}\right)\delta\tau^2 + \mathcal{O}(\delta\tau^4)$. We now make the simple observation that all symplectic integrators are constructed from the same Poisson brackets, and that these Poisson brackets are extensive quantities. We therefore measure the average values of the Poisson brackets $\langle\{S,\{S,T\}\}\rangle$ and $\langle\{T,\{S,T\}\}\rangle$ over a few equilibrated trajectories at the parameters of interest and then optimize the integrator (by adjusting the step sizes, order of the integration scheme, integrator parameters, number of pseudofermion fields, etc. [2, 3]) offline so as to minimize the cost. This is possible because the acceptance rate and instabilities are completely determined by $\Delta H = \tilde{H} - H$.

As a very simple example consider the minimum norm PQPQP integrator $U_{\text{PQPQP}}(\delta\tau)^{\tau/dt} = \left(e^{\alpha\hat{S}\delta\tau}e^{\frac{1}{2}\hat{T}\delta\tau}e^{(1-2\alpha)\hat{S}\delta\tau}e^{\frac{1}{2}\hat{T}\delta\tau}e^{\alpha\hat{S}\delta\tau}\right)^{\tau/dt}$ whose shadow Hamiltonian is

$$\tilde{H} = H + \left(\frac{6\alpha^2 - 6\alpha + 1}{12}\{S,\{S,T\}\} + \frac{1-6\alpha}{24}\{T,\{S,T\}\}\right)\delta\tau^2 + \mathcal{O}(\delta\tau^4).$$

One way of optimizing this integrator is to minimize $|\langle\Delta H\rangle|$ with respect to α, the average being taken with respect to the equilibrium distribution e^{-H}; this makes H as close to the conserved $\tilde{H}$ as possible. Another option is to minimize the quantity $|\langle\delta H\rangle| = |\langle\Delta H' - \Delta H\rangle|$ where the $\Delta H'$ is the value of ΔH at the end of a trajectory (before the Metropolis test); this is probably closer to minimizing the computational cost. In practice we make use of the fact that in equilibrium $\langle\delta H\rangle = \frac{1}{2}\langle\delta H^2\rangle + \cdots$ and choose to minimize $\langle\delta H^2\rangle$, as this is more stable numerically.

Gauge Fields

We must construct the Poisson brackets for gauge fields, where the field variables are constrained to live on a group manifold: to do this we need to use some differential geometry. The following table summarizes the difference between the formulation on flat space that we have discussed up to this point and that on general manifolds.

	Flat Manifold	General
Symplectic 2-form	$dp \wedge dq$	$\omega : d\omega = 0$
Hamiltonian vector field	$\hat{H} = \frac{\partial H}{\partial p}\frac{\partial}{\partial q} - \frac{\partial H}{\partial q}\frac{\partial}{\partial p}$	$dH = i_{\hat{H}}\omega$
Equations of motion	$\dot{q} = \frac{\partial H}{\partial p}, \dot{p} = -\frac{\partial H}{\partial q}$	$\left.\frac{d}{dt}\right\vert_{\sigma} = \hat{H}$
Poisson bracket	$\{A,B\} = \frac{\partial A}{\partial p}\frac{\partial B}{\partial q} - \frac{\partial A}{\partial q}\frac{\partial B}{\partial p}$	$\{A,B\} = -\omega(\hat{A},\hat{B})$

In order to construct a Hamiltonian system on a manifold we need not only a Hamiltonian function but also a fundamental closed 2-form ω. On a Lie group manifold this is most easily found using the globally defined *Maurer–Cartan* forms θ^i that are dual to the generators and satisfy the relation $d\theta^i = -\frac{1}{2}c^i_{jk}\theta^j \wedge \theta^k$, where c^i_{jk} are the structure constants of the group. We choose to define $\omega \equiv -d\sum_i \theta^i p^i = \sum_i(\theta^i \wedge dp^i - p^i d\theta^i) = \sum_i(\theta^i \wedge dp^i + \frac{1}{2}p^i c^i_{jk}\theta^j \wedge \theta^k)$. Using this fundamental 2-form we can define a Hamiltonian vector field $\hat{A}$ corresponding to any 0-form A through the relation $dA = i_{\hat{A}}\omega$ or equivalently $dA(\boldsymbol{x}) = \omega(\hat{A},\boldsymbol{x})\ \forall\boldsymbol{x}$. The classical trajectories $\sigma_t = (Q_t, P_t)$ are then the integral curves of this vector field, $\dot{\sigma}_t = \hat{A}(\sigma_t)$.

Poisson Brackets

For a Hamiltonian of the form $H = S + T$ we find that the leading Poisson brackets that appear in the shadow Hamiltonian for a symmetric symplectic integrator are $\{S,\{S,T\}\} = e_i(S)e_i(S)$ and $\{T,\{S,T\}\} = -p^ip^je_ie_j(S)$ where the p^i are the momentum coordinates and the e_i are linear differential operators satisfying $e_i(U) = T_iU$ for gauge fields $U \in \mathrm{SU}(n)$ with generators T_i.

Consider the Wilson pseudofermionic action $S = \phi^\dagger\mathcal{M}^{-1}\phi$, thus $e_i(S) = -\phi^\dagger\mathcal{M}^{-1}e_i(\mathcal{M})\mathcal{M}^{-1}\phi$, and $p^ip^je_ie_j(S) = p^ip^j\phi^\dagger\mathcal{M}^{-1}\left(2\,e_i(\mathcal{M})\mathcal{M}^{-1}e_j(\mathcal{M}) - e_ie_j(\mathcal{M})\right)\mathcal{M}^{-1}\phi$. $e_i(\mathcal{M})$ is straightforward to evaluate given the linearity of the Wilson–Dirac operator in the gauge field: we just use Leibniz rule and then replace the gauge field U by PU.

Nested Integrators

If it is much cheaper to evaluate the force for one part of the action, such as the pure gauge part, we can use a nested integrator with a very small step size for the “inner” cheap part. One might expect that one could then tune the “outer” part without reference to the cheap part, but this is not the case.

Let the Hamiltonian be $H = \frac{\pi^2}{2} + S_1 + S_2$ with $\|S_2\| \ll \|S_1\|$ and consider a nested integrator with a composite step of the form $U(\delta\tau) = \exp\frac{\hat{S}_2\delta\tau}{2}\left(\exp\frac{\hat{S}_1\delta\tau}{2m}\exp\frac{\hat{T}\delta\tau}{m}\exp\frac{\hat{S}_1\delta\tau}{2m}\right)^m\exp\frac{\hat{S}_2\delta\tau}{2}$. For the inner integrator the BCH formula tell us that $\left(\exp\frac{\hat{S}_1\delta\tau}{2m}\exp\frac{\hat{T}\delta\tau}{m}\exp\frac{\hat{S}_1\delta\tau}{2m}\right)^m$ may be written as

$$\exp\left[(\hat{S}_1 + \hat{T})\delta\tau + \left(\alpha[\hat{S}_1,[\hat{S}_1,\hat{T}]] + \beta[\hat{T},[\hat{S}_1,\hat{T}]]\right)\frac{\delta\tau^3}{m^2} + \mathcal{O}(\delta\tau^5)\right]$$

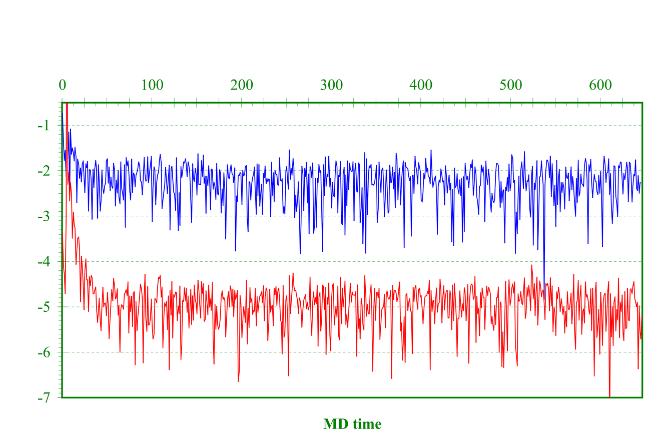
with $\alpha = -\frac{1}{24}$ and $\beta = \frac{1}{12}$. Applying the BCH formula again leads to the shadow Hamiltonian

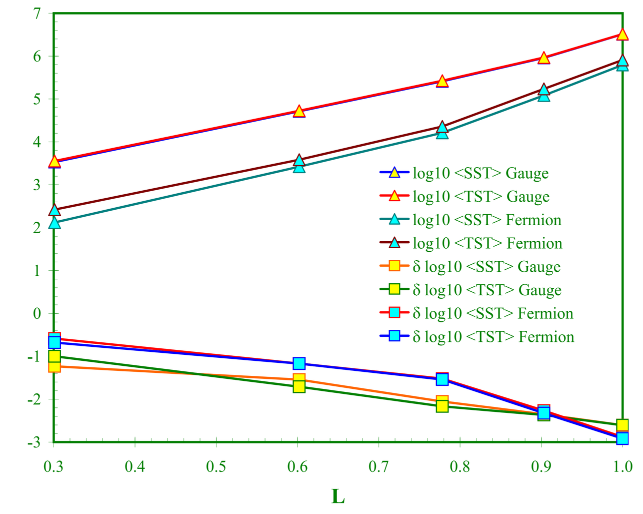
$$\begin{aligned}\tilde{H} = H\ +\ &\left(\alpha\{\hat{S}_2,\{\hat{S}_2,\hat{T}\}\} + \beta\{\hat{S}_1,\{\hat{S}_2,\hat{T}\}\} + \beta\{\hat{T},\{\hat{S}_2,\hat{T}\}\}\right. \\ &\left. + \frac{1}{m^2}\left(\alpha\{\hat{S}_1,\{\hat{S}_1,\hat{T}\}\} + \beta\{\hat{T},\{\hat{S}_1,\hat{T}\}\}\right)\right)\delta\tau^2 + \mathcal{O}(\delta\tau^4).\end{aligned}$$

Observe that the Poisson bracket $\{\hat{S}_1,\{\hat{S}_2,\hat{T}\}\}$ depends on the cheap action S_1 but is not supressed by any inverse power of m; it is therefore necessary to measure this quantity in order to optimize the integrator.

Results

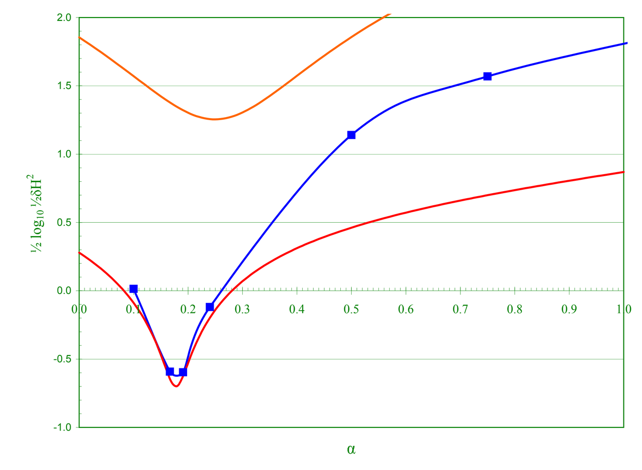
The blue curve in the first figure following shows how $\log_{10}|\delta H|$ behaves as a function of MD time, compared with the red curve $\log_{10}|\delta\tilde{H}|$ for the shadow Hamiltonian up to leading non-trivial order in $\delta\tau$. This demonstrates that once the system has reached equilibrium the shadow Hamiltonian is indeed conserved.

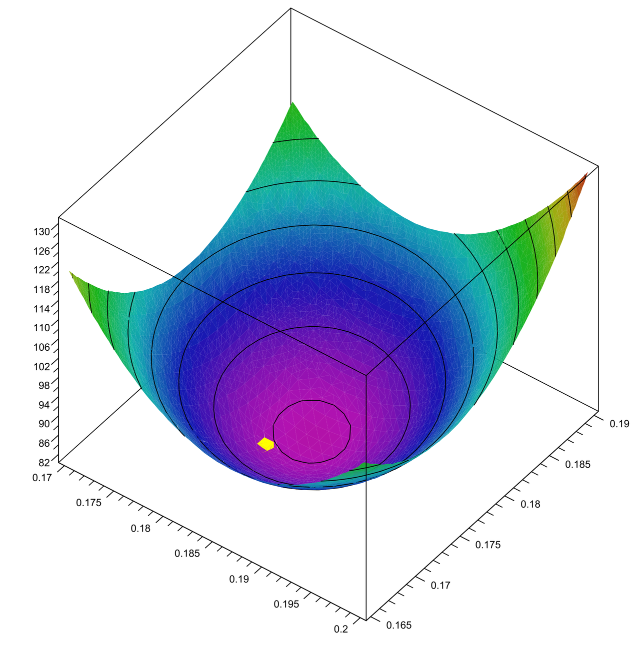




The second graph shows how several different Poisson brackets and their fluctuations depend on the lattice size. As expected the Poisson brackets are more-or-less extensive (they grow as L^4); the statistical fluctuations in the Poisson brackets are also shown, and they fall as L^{-2} as expected.

The first graph below shows how we tune the PQPQP integrator.





The orange curve shows the value of $|\langle\Delta H\rangle|$ as a function of the free parameter α on a logarithmic scale. This quantity reaches its minimum near $\alpha = 0.25$, whereas the minimum of $\langle\delta H^2\rangle$ (red curve) occurs for $\alpha = 0.179$, which agrees well with the measured values of this quantity (blue). Note that the red curve is computed from Poisson bracket values measured at the single value $\alpha = 0.24$. The fact that the phase space distribution at the end of trajectories depends on α probably accounts for the discrepancy between the blue and red curves for the values of α where the acceptance rate is essentially zero.

The second graph shows similar results for tuning the parameters for a dynamical fermion computation on a 8^4 lattice with a Wilson gauge action with $\beta = 5.6$ and Wilson fermions with $\kappa = 0.1575$. We used a two level PQPQP integrator with two gauge steps per fermion step, and a trajectory length of one. The yellow point shows values of the α parameters at which the Poisson brackets were measured.

Acknowledgements & References

P. J. Silva acknowledges support from FCT via grant SFRH/BPD/40998/2007. This work was supported in part by NSF grant PHY-0427646 and PPARC/STFC grant PP/D000238/1.

[1] M. A. Clark and A. D. Kennedy. Asymptotics of Fixed Point Distributions for Inexact Monte Carlo Algorithms. *Phys. Rev.*, D76:074508, 2007.

[2] Michael A. Clark and A. D. Kennedy. Accelerating dynamical fermion computations using the Rational Hybrid Monte Carlo (RHMC) algorithm with multiple pseudofermion fields. *Phys. Rev. Lett.*, 98:051601–051604, 2007.

[3] Tetsuya Takaishi and Philippe de Forcrand. Testing and tuning symplectic integrators for the Hybrid Monte Carlo algorithm in lattice QCD. *Phys. Rev. E*, 73:036706, 2006.