Numerical methods for lattice field theory

Mike Peardon

Trinity College Dublin

August 11, 2007

Numerical methods - references

- Good introduction to the concepts: "Simulation", Sheldon M. Ross, Academic Press ISBN 0-12-598053-1
- Detail on the theory of Markov chains: "Markov chains, Gibbs fields, Monte Carlo simulations and queues", Pierre Brémaud, Springer ISBN 0-387-98509-3
- A classic: "The Art of Computer Programming, Volume 2" Donald E. Knuth, Addison-Wesley ISBN 0-201-48541-9.
- And another: "Numerical Recipes: The Art of Scientific Computing (3rd Edition)", Press, Teukolsky, Vetterling and Flannery, CUP ISBN 0-521-88068-8
- Applications to field theory: "The Monte Carlo method in quantum field theory", Colin Morningstar `arXiv:hep-lat/0702020`

Monte Carlo integration of the path integral of QFTs

- The ergodic theorem tells us Markov chains can be used to perform importance sampling Monte Carlo. We have seen a few methods for building suitable Markov processes with a stationary state of our choice.
- Making predictions from the (Euclidean) path integral of a lattice quantum field theory is now possible.

$$\langle F \rangle = \frac{1}{Z} \int \mathcal{D}\phi \;\; F(\phi) e^{-S(\phi)}$$

- The (finite) lattice has made the path integral a high-dimensional "ordinary" integral, which we can estimate using Monte Carlo.
- The Boltzmann weight of configurations means only miniscule fractions of the configuration space contribute to the integral - it is crucial to use importance sampling.

Monte Carlo integration of the path integral of QFTs

- Start with bosonic fields. For a field ϕ representing a scalar boson on the lattice, we have

$$S(\phi) = \sum_{x,y} \phi_x M_{xy} \phi_y + \sum_x S_{\text{int}}(\phi_x)$$

- $M_{x,y}$ describes the Klein-Gordon operator, so we could discretise on a lattice with spacing a using

$$M_{x,y} = (am)^2 \delta_{x,y} - \sum_\mu (\delta_{x+\hat{\mu},y} + \delta_{x-\hat{\mu},y} - 2\delta_{x,y})$$

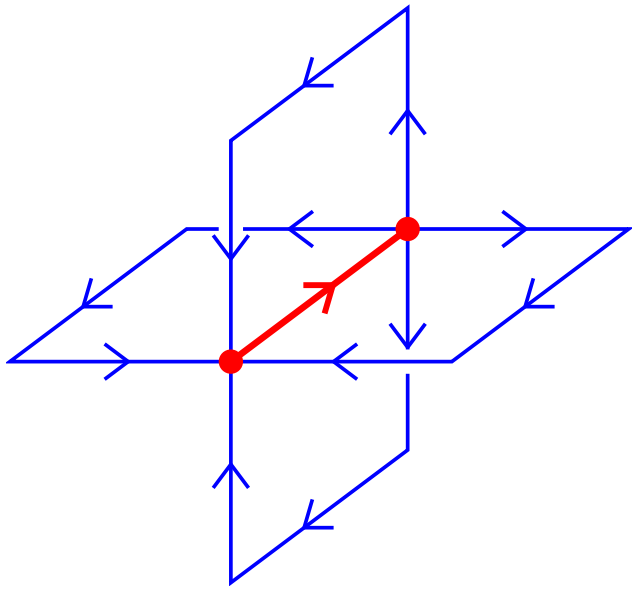
- Since (usually) S_{int} is local, Metropolis-Hasting with a change at a single site will often work well. A suitable proposal step might be

$$\phi'_x = \phi_x + \delta$$

with δ drawn from a normal distribution with mean 0. The variance of this distribution can be used to control the acceptance rate, and so the efficiency of the simulation.

Gauge bosons

- For the pure Yang-Mills theory, the Wilson gauge action is localised.
- A link appears in the action in just $2d$ plaquette terms.
- A local update (Gibbs sampler) is computationally efficient.



- For a Gibbs sampler, the action break-up ($S_G = S_{\mathrm{Gibbs}}(U_\mu(x) + \tilde{S}_\mu(x))$ becomes

$$S_{\mathrm{Gibbs}}[U_\mu(x)] = \mathrm{ReTr}\ (U_\mu(x)\Sigma_\mu(x))$$

with

$$\Sigma_\mu(x) = \sum_{\nu \neq \mu} U_\nu(x+\hat{\mu})U_\mu^\dagger(x+\hat{\nu})U_\nu^\dagger(x) + U_\nu^\dagger(x+\hat{\mu}-\hat{\nu})U_\mu^\dagger(x-\hat{\nu})U_\nu(x-\hat{\nu})$$

Gauge bosons

- Updating one link would require generating a group element from

$$\mu(U) = \frac{\exp(-\beta \text{ ReTr } U\Sigma)}{\int dU \exp(-\beta \text{ReTr } U\Sigma)}$$

- Creutz developed an algorithm for $SU(2)$ (M. Creutz, Phys.Rev.D21 (1980) 2308-2315.)
- $SU(2)$ has the useful property that $\Sigma = k\bar{U}, \bar{U} \in SU(2)$. so re-write the problem in terms of $U' = U\bar{U}$ and draw from

$$\frac{1}{Z} \exp(\frac{1}{2}\beta k \text{Tr } U')dU'$$

- For $SU(2)$, we can parameterise U' in terms of 4 real numbers

$$U = a_0 + i\sum_{k=1}^{3} a_k \sigma_k$$

Gauge bosons

- The group-invariant Haar measure (for this parameterisation) is $dU' = \frac{1}{2\pi^2}\delta(|a|^2 - 1)d^4a$ and changing to (three-dimensional) spherical polar co-ordinates deals with the constraint to give a distribution

$$d\Omega da_0 \;\; \frac{1}{2}\sqrt{1 - a_0^2}e^{\beta k a_0}$$

  which means a_0 and the direction of the three-vector $\underline{a}$ are independent (it has length $\sqrt{1 - a_0^2}$).

- a_0 is generated with the rejection method, first using $e^{\beta k a_0}$, then accepting with probability $\sqrt{1 - a_0^2}$.
- "We leave it to the interested reader to design his own scheme for randomly selecting the direction for $\underline{a}$" (Creutz).
- Cabibbo and Marinari used this to make a "quasi-heatbath" method for general $SU(N)$ by updating $SU(2)$ subgroups. $SU(2)$

Fermions

- The lattice physicist's complaint: "why are there no grassmann-valued variables in my computer?"
- While the algebra is simple to formulate, it is tricky to handle directly in a computer (it *can* be represented as a matrix algebra)

$$\int \eta d\eta = 1, \quad \int 1 \, d\eta = 0$$

- Direct representation of the algebra is avoided; integrate out a fermion bilinear to give an effective action.

$$\int \mathcal{D}\bar{\psi}\mathcal{D}\psi e^{-\bar{\psi}M\psi} = \det M$$

- "γ_5-hermiticity" ($M^\dagger = \gamma_5 M \gamma_5$) implies $\det M$ is real, but not that it is positive. Odd numbers of (mass-degenerate) fermions are potentially a problem: importance sampling need a +ve-definite weight.
- Use $|\det M|$ with $\operatorname{sgn}\det M$ in the observables. This introduces a "sign problem": variances of estimators can be prohibitively large.

Fermions (2)

- Local Four-fermi (and more) interactions can be handled by introducing an auxiliary field. A simple example illustrates the idea:

$$e^{\sigma\bar{\psi}\Gamma\psi-\sigma^2/2} = 1 + \sigma\bar{\psi}\Gamma\psi + \frac{\sigma^2}{2}\left(\bar{\psi}\Gamma\psi\right)\left(\bar{\psi}\Gamma\psi\right) \quad \text{if} \quad \left(\bar{\psi}\Gamma\psi\right)^3 = 0$$

so

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} d\sigma \;\; e^{\sigma\bar{\psi}\Gamma\psi-\sigma^2/2} = 1 + \frac{1}{2}\left(\bar{\psi}\Gamma\psi\right)\left(\bar{\psi}\Gamma\psi\right) = e^{\frac{1}{2}\left(\bar{\psi}\Gamma\psi\right)\left(\bar{\psi}\Gamma\psi\right)}$$

- The four-fermi interaction at one site is then replaced by an interaction with a new (non-propagating!) field and a fermion bilinear
- We can now integrate out the fermion and put the auxiliary field on the computer.

Fermions (3)

- We have just postponed the problem of simulating fermions.
- Now we have a **non-local** function of the remaining (bosonic/gauge/auxiliary) fields that coupled to the fermions.
- Manipulating large determinants is computationally expensive. A "bosonisation" trick is commonly used: the pseudofermion field.
- For two flavours of fermions with a γ_5-hermian lattice representation of the Dirac matrix, we have

$$\det M^2 = \det M^\dagger M = \frac{1}{\det\left[M^\dagger M\right]^{-1}}$$

and this is represented as a gaussian integral;

$$\frac{1}{\det\left[M^\dagger M\right]^{-1}} = \int \mathcal{D}\phi\mathcal{D}\phi^* \quad e^{-\phi^*\left[M^\dagger M\right]^{-1}\phi}$$

- The pseudofermions inherit the spin structure and colour charge.
- For (eg) quarks coupled to a gluon field, M depends on the links.

QCD

- With this machinery, Euclidean lattice QCD can be manipulated in a computer. An expression for an observable is:

$$\langle F \rangle = \int \mathcal{D}U\mathcal{D}\phi^*\mathcal{D}\phi \quad F[U] \quad \mu(U, \phi^*, \phi)$$

where μ is the (importance sampling) weight

$$\mu(U, \phi^*, \phi) = \frac{1}{\mathcal{Z}} \exp\left\{-\beta S_G[U] - \phi^* \left(M^\dagger[U]M[U]\right)^{-1} \phi\right\}$$

with $\mathcal{Z}$ the path integral

$$\mathcal{Z} = \int \mathcal{D}U\mathcal{D}\phi^*\mathcal{D}\phi \quad \exp\left\{-\beta S_G[U] - \phi^* \left(M^\dagger[U]M[U]\right)^{-1} \phi\right\}$$

- So if a Markov chain of configurations can be generated with stationary state μ then the ergodic theorem gives us $\langle F \rangle = E[F(U)]_\mu$

Molecular dynamics

- Consider a bosonic theory with a (possibly non-local) action S. The importance sampling measure we want to draw from is

$$\mu[\phi] = \frac{1}{Z} e^{-S[\phi]}$$

- For every degree of freedom in ϕ, add an independent partner, π. The joint probability distribution for this system is

$$\mu'[\phi, \pi] = \frac{1}{Z'} e^{-S[\phi] - T[\pi]}$$

$E[f(\phi)]$ is unchanged: the density is a separable product $e^{-S} \times e^{-T}$

- We'll choose $T[\pi] = \frac{1}{2} \sum_x \pi^2$ so π is normally distributed.
- Now the magic sleight-of-hand; write

$$\mathcal{H} = S + T$$

and pretend $\mathcal{H}$ is a (classical) hamiltonian, which defines equations of motion in a new fictitious time co-ordinate when π is thought of as the momentum conjugate to ϕ.

. . . a word from my sponsors. . .

William Rowan Hamilton (1805-1865)

Andrews Professor of Astronomy
Trinity College Dublin

Molecular dynamics (2)

- These equations of motion define an **area-preserving (Liouville's theorem) reversible mapping** on the phase space $\{\phi, \pi\}$.
- We can use this to define a Markov process suitable for importance sampling:
  1. Draw π from the normal distribution.
  2. Integrate the equations of motion for $\mathcal{H}$ for some time interval, τ.
- The equations of motion are

$$\dot{\phi}_i = \pi_i, \quad \dot{\pi}_i = -\frac{\partial S}{\partial \phi_i}$$

- but unfortunately it is usually impossible to write a simple expression that solves the equations of motion. We must use a numerical integrator and this will introduce a finite-step-size error.
- Langevin dynamics (stochastic differential equations) use a similar idea. The R-algorithm is adaptable to fractional powers of the fermion matrix, and is popular for staggered fermions.

Molecular dynamics (3)

- To avoid finite-step-size integrator errors for the equations of motion, a modified algorithm, **HMC** was proposed (S. Duane, A.D. Kennedy, B.J. Pendleton and D. Roweth, Phys.Lett.B195 (1987) 216-222.)
- Remember that the Metropolis-Hastings algorithm has two parts; a reversible, area-preserving proposal step followed by an accept-reject test and it obeys detailed balance for a given stationary state.
- Numerical integrators that exactly preserve $\mathcal{H}$ are hard to build, but ones that *almost* conserve $\mathcal{H}$ but retain the important properties (reversibility and area-preservation) can be found.
- The best-known integrators with these properties are called reversible **symplectic** integrators.

The leap-frog integrator

- Update $\phi : \phi_i(t + \frac{h}{2}) = \phi_i(t) + \frac{h}{2}\pi_i(t)$
- Update $\pi : \pi_i(t + h) = \pi_i(t) - h\frac{\partial S}{\partial \phi_i}(t + \frac{h}{2})$
- Update $\phi : \phi_i(t + h) = \phi_i(t + \frac{h}{2}) + \frac{h}{2}\pi_i(t)$

Hybrid Monte Carlo

Hybrid Monte Carlo

1. Draw new conjugate momenta, π from the normal distribution
2. Store the current field, ϕ and compute $\mathcal{H}(\pi, \phi)$
3. Integrate the equations of motion using a reversible, symplectic integrator with step-size h (such as leap-frog) so $(\phi, \pi) \overset{\text{leapfrog}}{\longrightarrow} (\phi', \pi')$
4. Compute $\mathcal{H}(\pi', \phi')$ and the change, $\Delta\mathcal{H}$
5. Accept ϕ' as the new entry in the Markov chain with probability

$$\mathcal{P}_{\text{acc}} = \min\left[1, e^{-\Delta\mathcal{H}}\right]$$

if the new configuration is rejected, then make ϕ the new entry.

- As $h \to 0, \Delta\mathcal{H} \to 0$, so $\mathcal{P}_{\text{acc}} \to 1$
- $E[e^{-\Delta\mathcal{H}}] = 1$ and $E[\Delta\mathcal{H}] = \frac{1}{2}E[(\Delta\mathcal{H})^2]$ for small $\Delta\mathcal{H}$

Molecular dynamics (4)

- For gauge theories, our degrees of freedom are constrained, so we need to define hamiltonian dynamics on curved manifolds.
- A Lie group G (all continuous gauge groups) has particularly helpful properties. At all points, there is a well-defined, tangent space (the Lie algebra at the identity element) in which conjugate momenta naturally live.
- A useful definition for a momentum variable p conjugate to a group element U is

$$p = p_a T_a \text{ so } p \in \mathcal{L}(G)$$

with T_a the (hermitian) generators of the group and define the equation of motion for U to be

$$\dot{U} = ipU$$

- The extra appearance of U shows we need to rotate the Lie algebra (where p lives) to be tangent to U. The left multiplication is a convention (right works just as well).