Fermion algorithms

Stefan Schaefer

John von Neumann Institute for Computing, DESY

Lattice Practices 2017

HMC

Momentum Heatbath

Refresh momenta π (Gaussian random numbers)

Molecular Dynamics

Solve numerically MD equations for some MC time τ (trajectory) deriving from Hamiltonian $H = \frac{1}{2}(\pi, \pi) + S[U]$.

Acceptance Step

Correcting for inaccuracies in integration.

Model

Schwinger model

2D lattice

U(1) gauge fields

Shares many properties of QCD

- asymptotically free
- confining
- spontaneously broken chiral symmetry

Fields
two component spinor fields on sites $\psi(x,d)$ with $d=0,1$

gauge fields on the links $U(x,\mu)=e^{-\mathrm{i}\theta(x,\mu)}$.

Action

Gauge action

$$S_g[U] = -\beta \sum_{\mathrm{plaq}} \mathrm{Re} U_{\mathrm{plaq}}$$
$$= -\beta \sum_{x} \cos(\theta(x, 0) + \theta(x + \hat{0}, 1) - \theta(x + \hat{1}, 0) - \theta(x, 1))$$

Fermion action
We consider the two flavor model.

$$S_f[U, \psi, \bar{\psi}] = \sum_f \bar{\psi}_f M(\kappa) \psi_f$$

With $M = 1 - \kappa H$ and

$$H(x, y) = \sum_{\mu} \delta_{x-\hat{\mu}}(1 + \gamma_\mu) U_{x-\hat{m}u,\mu} + \delta_{x+\hat{\mu},y}(1 - \gamma_\mu) U^*_{x,\mu}$$

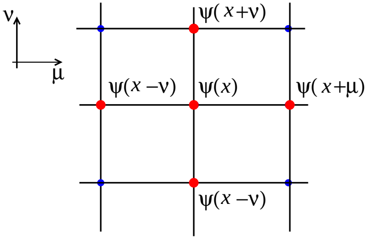
γ matrices are the Pauli matrices.

Wilson Dirac operator

With $M = 1 - \kappa H$ and

$$H(x, y) = \sum_{\mu} \delta_{x-\hat{\mu}}(1 + \gamma_{\mu})U_{x-\hat{m}u,\mu} + \delta_{x+\hat{\mu},y}(1 - \gamma_{\mu})U^{*}_{x,\mu}$$

γ matrices are the Pauli matrices.
Spinor fields $\psi(x, d) \quad \rightarrow$ `complex psi[i], i=1...2V`



$$\begin{pmatrix} \times & \times & 0 & 0 & 0 & \times \\ \times & \times & \times & 0 & 0 & 0 \\ 0 & \times & \times & \times & 0 & 0 \\ 0 & 0 & \times & \times & \times & 0 \\ 0 & 0 & 0 & \times & \times & \times \\ \times & 0 & 0 & 0 & \times & \times \end{pmatrix}$$

The Dirac operator

Dirac operator D can be viewed as a matrix acting in $\mathbb{C}^{2V}$

For **Wilson, staggered** and **domain wall** fermions this matrix is sparse.

Application of D on vector scales $\propto V$.

Methods to solve Dirac equation based on

Matrix $\times$ vector

Itererative methods, huge applied math literature.

Integrating out the fermions

ψ are Grassmann fields. Not suitable for computer simulation.

$$\int [d\psi][d\bar{\psi}] e^{\sum_f \bar{\psi}_f M \psi_f} = (\det M[U])^2 = \det M^\dagger[U] M[U]$$

because for the Wilson action

$\gamma_5 M \gamma_5 = M$

Can again be cast into the form of an action

$$\det M^\dagger M = e^{\log \det M^\dagger M}$$

Have given up locality, important ingredient of QFT.

First step away from a natural formulation

Fermion determinant

$$S_f[U] = -\log \det M^\dagger M$$

Virtually impossible to compute.

Computation of determinant takes $\propto V^3$ steps.

Numerically highly problematic.

Would lose sparcity of fermion matrix in the process.

How to proceed?

For all MCMC algorithms, ratios of probabilities have to be computed.

$$e^{-(S_g[U]-S_g[U'])} \prod_f \frac{\det M_f[U]}{\det M_f[U']}$$

even an infinitesimal change needs a matrix inversion

$$\det(M[U] + \delta M) = \det M[U] \left(1 + \mathrm{tr}\{M[U]^{-1}\delta M\} + \dots\right)$$

Makes link update algorithms virtually impossible.

Change due to one link requires (expensive) operation on full lattice.

Naturally led to algorithms updating many links at a time.

What makes fermions special?

Local, natural formulation not suitable

Many choices, which seem logical, but are not unique.

Can have a large influence on algorithm performance.

Fermions

Formulation of the theory

Pseudofermions

Pseudofermions — PETCHER, WEINGARTEN'81

$$\det Q^2 \propto \int [\mathrm{d}\phi][\mathrm{d}\phi^\dagger]\, e^{-(\phi,\, Q^{-2}\phi)} \;, \qquad\qquad Q = \gamma_5 M$$

Gaussian integral $\rightarrow$ apply transformation $\phi = Q\,\eta$

$$\int [\mathrm{d}\eta][\mathrm{d}\eta^\dagger]\, e^{-(\eta,\,\eta)} = \int [\mathrm{d}\phi][\mathrm{d}\phi^\dagger]\, \det{}^{-2} Q\, e^{-(\phi,\, Q^{-2}\phi)}$$

Determinant is the **Jacobian** of this transformation.

Q is Hermitian

Comments

$$\det Q[U]^2 = \frac{1}{Z_\phi} e^{-(\phi,\, Q[U]^{-2}\phi)}$$

- Essential for making fermions amenable to computer simulations

- Pseudofermion ϕ can be generated with simple heatbath

  Generate Gaussian complex-valued quark field η

  $$P[\eta] \propto e^{-(\eta,\eta)}$$

  Multiply with Q

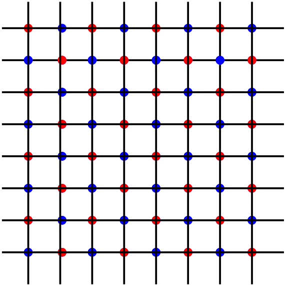
  $$\phi = Q\eta$$

- PF action nonlocal.

Even–odd preconditioning

The Wilson Dirac operator connects only neighboring sites.

Label them “even” and “odd”.



$$M = \begin{pmatrix} M_{ee} & M_{eo} \\ M_{oe} & M_{oo} \end{pmatrix}$$

M_{oo} and M_{ee} are site-diagonal matrices.

Even–odd preconditioning

Matrix identity

$$\begin{pmatrix} M_{ee} & M_{eo} \\ M_{oe} & M_{oo} \end{pmatrix} =$$
$$\begin{pmatrix} 1 & M_{eo}M_{oo}^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} (M_{ee} - M_{eo}M_{oo}^{-1}M_{oe}) & 0 \\ 0 & M_{oo} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ M_{oo}^{-1}M_{oe} & 1 \end{pmatrix}$$

For the determinant this means

$$\det M = \det M_{oo}\ \det(M_{ee} - M_{eo}M_{oo}^{-1}M_{oe}) \equiv \det M_{oo} \det \hat{M}$$

with $\hat{M}$ the Schur complement.

In the following, I will mostly write M or $Q = \gamma_5 M$.
In practice, this frequently means $\hat{M}$ or $\hat{Q}$.

Partition function

Include pseudofermions in path integral.

$$Z = \int [dU][d\pi][d\phi][d\phi^\dagger]\, e^{-\frac{1}{2}(\pi,\pi) - S_g[U] - (\phi, \frac{1}{Q^2}\phi) + 2\log\det Q_{oo}}$$

S_g: gauge action

effective fermion action for $N_f = 2$.

$$S_{f,eff} = (\phi, \frac{1}{\hat{Q}^2}\phi) - 2\log\det Q_{oo}$$

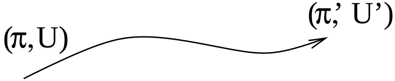
HMC

Momentum and pseudofermion Heatbath

Refresh momenta π
Refresh pseudofermions $\phi \to$ kept fixed during trajectory

Molecular Dynamics

Solve numerically MD equations for some MC time τ (trajectory) deriving from Hamiltonian $H = \frac{1}{2}(\pi, \pi) + S[U]$.



Acceptance Step

Correcting for inaccuracies in integration.

Problems

Pseudofermions Petcher, Weingarten'81

$$\det Q^2 \propto \int \mathrm{d}\phi^\dagger \mathrm{d}\phi \, e^{-(\phi,\, Q^{-2}\phi)}$$

- Works only for pairs of degenerate flavors
  Solution: take square root $\rightarrow$ PHMC, RHMC
- Force evaluation expensive: 2 solutions of Dirac eq.

$$F_{\mathrm{pf}} = -(\phi,\, Q^{-2}\, \delta Q\, Q^{-1}\, \phi) + \text{h.c.}$$

- Seems somewhat unnatural
  Start with manifestly local action
  $\rightarrow$ quite non-local expression

Berlin Wall

Status 2000 Quarks 16× heavier than in nature.
No perspective even with 2010 computers.

Coarse lattices $a \approx 0.1$fm
(the typical length scale is 1fm)

Cost of a simulation (Ukawa Lattice 2001)

$$\text{Cost} = C \left[\frac{\#conf}{1000}\right] \cdot \left[\frac{m_q}{16 m_{\text{phys}}}\right]^{-3} \cdot \left[\frac{L}{3\text{fm}}\right]^{5} \cdot \left[\frac{a}{0.1\text{fm}}\right]^{-7}$$

$C \approx 2.8$ Tflops year

Fermions

Pseudofermions PETCHER, WEINGARTEN'81

$$\det Q^2 \propto \int [\mathrm{d}\phi][\mathrm{d}\phi^\dagger]\, e^{-(\phi,\, Q^{-2}\phi)}$$

- **HMC + single pseudofermion action not successful**
- Compare

$$F_{\mathrm{pf}} = \delta(\phi,\, Q^{-2}\phi) \qquad \text{and} \qquad F_{\mathrm{ex}} = -\delta \mathrm{tr}\, \log\, Q^2$$

- F_{pf} is "stochastic estimate" of F_{ex}
  At beginning of the trajectory $\langle F_{\mathrm{pf}} \rangle_\phi = F_{\mathrm{ex}}$
- Very large fluctuations in F_{pf}

$$|F_{\mathrm{pf}}| \gg |F_{\mathrm{ex}}|$$

Fermions
Modifications

Determinant Splitting

Insight

- Need better estimate of determinant.
- Frequency splitting.

Mass preconditioning

Hasenbusch'01, Hasenbusch,Jansen'03

$$\det Q^2 = \det \frac{Q^2}{Q^2 + \mu^2} \det(Q^2 + \mu^2)$$

- Each determinant represented by pseudo-fermion
- "Pauli-Villars" for fermion force
- more intermediate $\mu \to$ Noise reduction in force.
- success depends on choice of μ. Urbach et al'04

Numerical examples

Action

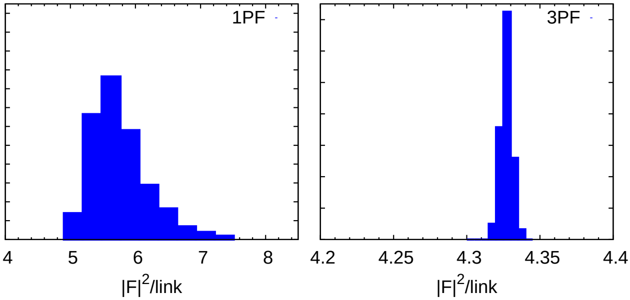
- $N_{\mathrm{f}} = 2 + 1$ NP improved Wilson fermions
- Iwasaki gauge action
- 64×32^3 lattice with $a = 0.09$fm
- studied extensively by PACS-CS AOKI ET AL'09,'10
- $m_\pi = 200$MeV
- $m_\pi L = 3$

Algorithm M. LÜSCHER, S.S.'12

- Reweighting to avoid stability problems.
- Generated with public `openQCD` code.
  `http://cern.ch/luscher/openQCD`

Effect of determinant factorization

Forces for light quark, 20 configurations. $\mu_1 = 0.05$, $\mu_2 = 0.5$



- Fluctuations of force not much reduced.
- Fluctuations in **norm** squared of force:
  Spread reduced by more than factor 100.
  (Different scale!)

Understanding the improvement

Framework — CLARK, JOO, KENNEDY, SILVA'11

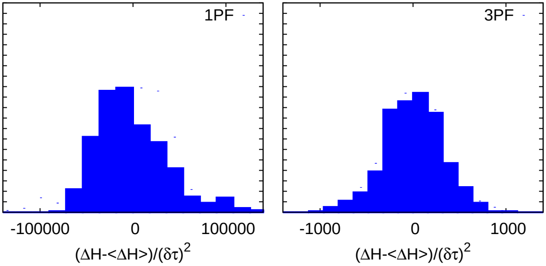
- Shadow Hamiltonian of symplectic integrators

$$\tilde{H} = H + (c_1\, \partial_x S \partial_x S - c_2\, \pi_x \pi_y \partial_x \partial_y S)\epsilon^2 + \ldots$$

- Large cancellation between the two terms
  $\rightarrow$ **potential for optimization.**

- 2nd order minimum norm integrators:
  minimum of $c_1^2 + c_2^2$ — Omelyan, Mrygold, Folk'03
- **Symplectic integrators profit from reduced fluctuations in norm of force.**

Numerical examples



- $\Delta H = \tilde{H} - H$, fermions only.
- Second order min. norm Omelyan integrator.
- Much larger step-size possible.

Factorizations

Hasenbusch Hasenbusch'03

$$\det Q^2 = \det \frac{Q^2}{Q^2+\mu_1^2} \; \det \frac{Q^2+\mu_1^2}{Q^2+\mu_2^2} \cdots \det(Q^2+\mu_N^2)$$

RHMC Kennedy, Horvath, Sint'99, Clark, Kennedy'07

$$\det Q^2 = \prod_{i=1}^{N} \det \sqrt[N]{Q^2}$$

Domain decomposition Lüscher '04

$$\det Q = \det Q_{\text{block}} \det R$$

RHMC

For realistic simulations need single (non-degenerate) quark flavors

$$\det Q^2 = W \det R^{-1}$$

with R a rational approximation to $(Q^2)^{-1/2}$

Zolotarev approximation

The Zolotarev rational function

$$R_{n,\epsilon}(x) = A \frac{(x+a_1)(x+a_3)\cdots(x+a_{2n-1)})}{(x+a_2)(x+a_4)\cdots(x+a_{2n})}$$

with degree (n,n) approximates the function $f(x) = 1/\sqrt{x}$ with the smallest possible deviation

$$\delta = \max_{\epsilon \leq x \leq 1} |1 - \sqrt{x} R_{n,\epsilon}(x)|$$

RHMC

For realistic simulations need single (non-degenerate) quark flavors

$$\det Q^2 = W \det R_{n,\epsilon}^{-1}(r_b^{-2} Q^2)$$

Need stable upper and lower bounds r_a and r_b of spectrum of Q^2.

Can be a problem for light Wilson fermions.

Two choices for W

- Acceptance step of the metropolis
- Reweighting factor included in the measurement

PHMC

Frezzotti,Jansen '98

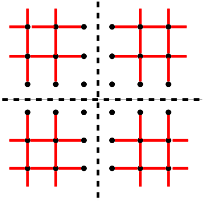
Similar to RHMC

Use polynomial approximation to $1/\sqrt{x}$.

No need to solve Dirac equation.

Might have advantage due to simplicity for well-conditioned Q^2.

Domain decomposition



- Domain decomposition
  $\rightarrow$ Divide the lattice in blocks

$$\det D = \det D_{\text{block}} \cdot \det D_{\text{R}}$$

- Do not update links connecting blocks
  $\rightarrow$ longer autocorrelations
- Good for slow communication.
- Brings back some locality to the theory

Reweighting

No need to simulate action S, can generate ensembles with a different action.

$$
\begin{aligned}
\langle A \rangle &= \frac{1}{Z} \int [dU] e^{-S[U]} A[U] \\
&= \frac{Z'}{Z} \frac{1}{Z'} \int [dU] e^{-S'[U]} (e^{-(S[U]-S'[U]} A[U]) \\
&= \frac{\langle AW \rangle'}{\langle W \rangle'}
\end{aligned}
$$

With $W = e^{-(S[U]-S'[U]}$ and

$$
Z = \int [dU] e^{-S[U]} \qquad \text{and} \qquad Z' = \int [dU] e^{-S'[U]} \tag{1}
$$

Can be used to improve simulations.

Has its limitations: large fluctuations in W will lead to large errors.

Twisted mass reweighting

Lüscher, Palombi'09

Wilson Dirac operator can have zero eigenvalues.

Action $-\log\det D$ becomes infinite.

Configuration space separated in disconnected sectors.

Ergodicity with continuous algorithms compromized.

Simulate with action that does not allow zero eigenvalues

$$D \to D + i\gamma_5\mu$$

μ needs to be small enough in order for reweighting to work.

Summary

Fermion action $-\mathrm{tr}\log D$ cannot be simulated directly.

Use pseudofermions *together* with matrix factorization

Several factorizations lead to working setups.

Need of solving the Dirac equation in each force evaluation.