Simulations in Lattice QCD

Lattice Practice 2024 at The Cyprus Institute - Simulations

Jacob Finkenrath

- Basic
  - Markov Chain Monte Carlo
- Intermediate
  - Hybrid Monte Carlo algorithm
- Advance
  - Fermions in simulations and more

this talk is based on Mattia Della Brida's constribution from 2021

Goal: Evaluation of pathintegral

Compute

$$< \square >= \frac{1}{Z} \int D\phi e^{-S(\phi)} \square(\phi) \qquad D\phi = \prod_{i=1}^{M} d\phi_i \quad e.g. \quad \phi = U, \overline{\psi}, \psi$$

- Deterministic integration methods not feasible!

  Current lattice QCD simulations can have M = O(10^9)

- Monte Carlo: evaluates integral by sampling the integrand at points selected with probability under the

  integration measure

Basic idea

1. Generate sequence of field configurations with probability

$$P(\phi^{(t)}) = \frac{1}{Z} e^{-S(\phi^t)}$$

1. Evaluate

$$\overline{O} = \frac{1}{N} \sum_{t=1}^{N} \square(\phi^{(t)})$$

Monte Carlo integration

Consider

$$\langle f \rangle = \int_D dx\ f(x) \qquad x = (x_1, \dots, x_d) \quad D = [0,1]^d$$

Compute

$$\overline{f_N} = \frac{1}{N} \sum_{k=1}^{N} f(x^{(k)}) \qquad x^{(k)} = (x_1^{(k)}, \dots x_d^{(k)})$$

where *x* are random numbers uniformly distributed within [0,1] . This requires a solid random number generator.

Central limit theorem

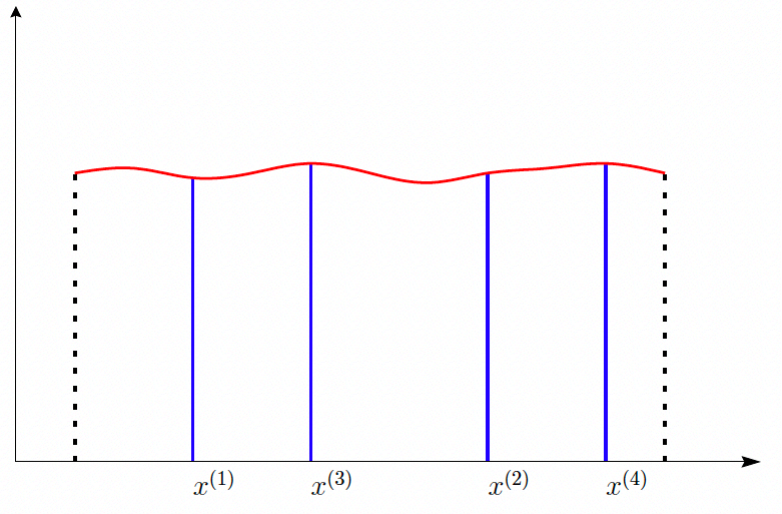
$$P(\overline{f_N}) \overset{N\to\infty}{\propto} \exp\left[-\frac{1}{2}\frac{(\overline{f_N} - \langle f \rangle)^2}{\mathrm{var}(f)/N}\right] \quad \text{with} \quad \mathrm{var}(f) = \langle (f - \langle f \rangle)^2 \rangle$$

with

$$\overline{f_N} - \langle f \rangle = (O)(1/\sqrt{N})$$

- Uncertainty of results are of statistical rather than systematic nature
- Error scales as $1/\sqrt{N}$ independently of the dimension d
- Rate of convergence depends on var(f)

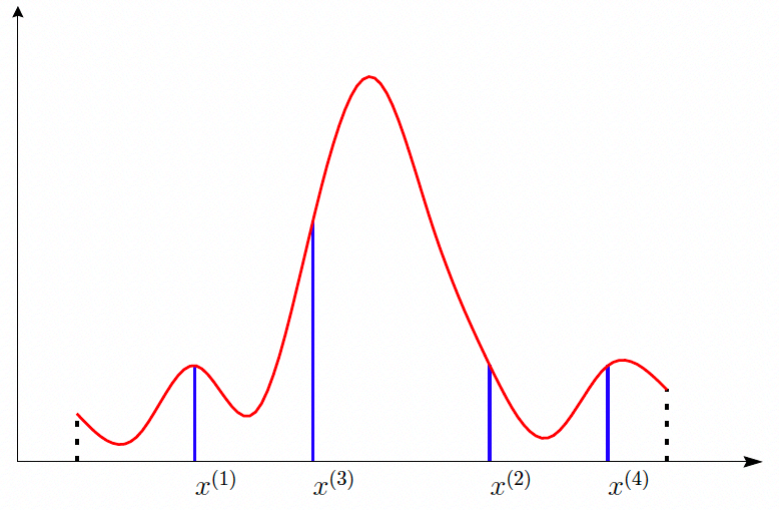
Monte Carlo integration



Uniformly sampling is effective for approximately constant functions;

here:

var(f) is small



in contrast:

Sampling of more complicated functions is more difficult

here

Using uniformly distributed random numbers can easily require large values of N to reach a good precision.

Importance sampling

Consider

$$\langle f \rangle = \int_D dx \; p(x) \left[\frac{f(x)}{p(x)} \right] = \langle g \rangle_p \qquad \text{with} \quad p(x) > 1 \quad \text{and} \quad \int_D dx \; p(x) = 1$$

with

$$g(x) = \frac{f(x)}{p(x)}$$

Compute

$$\overline{g_N} = \frac{1}{N} \sum_{k=1}^{N} g(x(k)) \qquad x^{(k)} = (x_1^{(k)}, \ldots x_d^{(k)})$$

with x random vectors distributed according to p(x)

Importance sampling

Central limit theorem

$$P(\bar{g}_N) \overset{N\to\infty}{\propto} \exp\left[-\frac{1}{2}\frac{(\overline{g_N} - \langle f\rangle)^2}{\mathrm{var}(g)/N}\right] \quad \text{with} \quad \mathrm{var}(f) = \langle (g - \langle f\rangle)^2\rangle_p$$

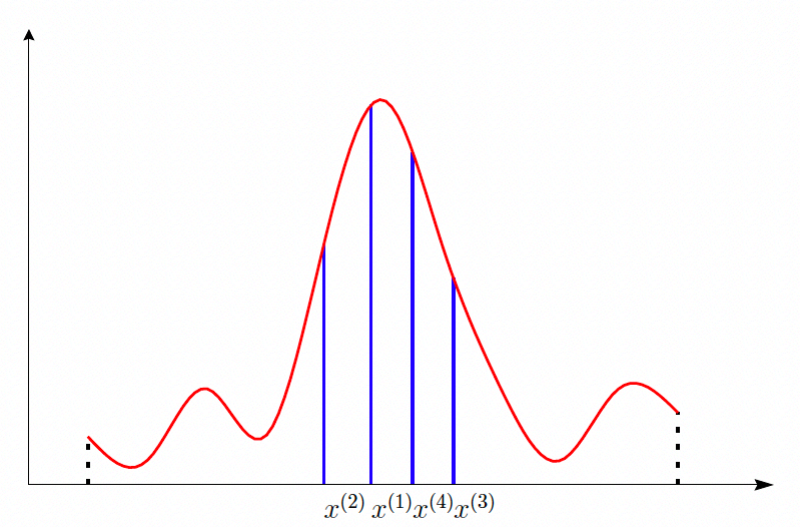
with

$$\langle f\rangle = \overline{g_N} \pm \sigma(\overline{g_N}) \quad \sigma(\overline{g_N}) = \sqrt{\mathrm{var}(g)/N}$$

Choice of p(x) can signicantly affect convergence

- Optimal p(x) but this requires <|f|>
- Only relatively simple distributions p(x) can be directly sampled, e.g. via inverse transform, hit-and-miss, etc.

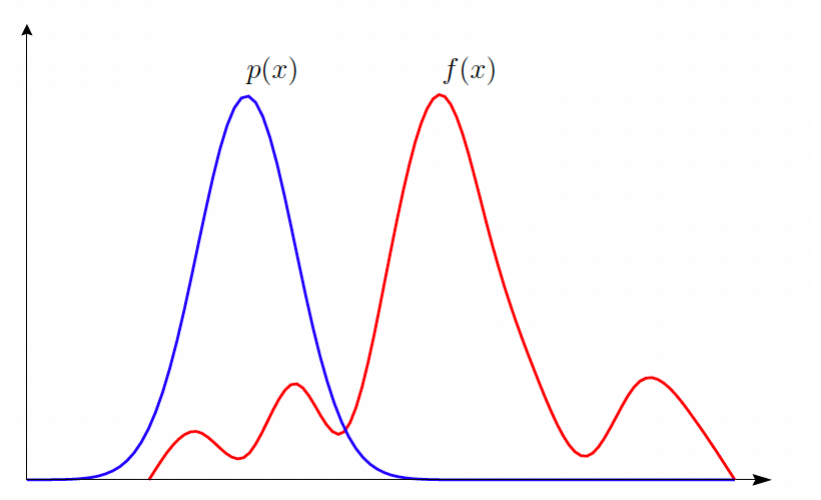
Importance sampling



Effective

A better sampling distribution allows for sampling more frequently the regions that give more contribution to the integral

$\Rightarrow$ faster convergence



However

When the sampling distribution and the function to integrate have little overlap, we say there is an overlap problem

$\Rightarrow$ very ineffective sampling

Importance sampling

Application to lattice field theory:

$$\langle \square \rangle = \int D\phi \, P_S(\phi) \square(\phi) \qquad P_S(\phi) = \frac{1}{Z} e^{-S(\phi)}$$

- Sharply peaked around configurations of minimal action
- Far too **complicated distribution** for a direct sampling
- We need method based on relative probabilities
  - avoids computation of normalization Z
- Convenient, but not necessarily optimal for all O(x)
  - may result in some large variances
- $S(\phi)$ must be real (and bounded)

Markov Chain Monte Carlo

A (discrete) Markov chain is a sequence of random variables

$$\phi^{(0)} \to \phi^{(1)} \to \cdots \to \phi^{(t)} \to \cdots \cdots \to \phi^{(N)} \qquad \phi(t) \in \Omega \leftarrow \text{state space}$$

which probability of extraction is given by a transition probability

$$T(\phi \to \phi')$$

(t is referred to as Markov time)

Properties

1. **Markovian**

$$T(\phi \to \phi') \quad \text{only depends on the current } (\phi) \text{ and future } (\phi') \text{ state}$$

2. **Time-homogeneous**

$$T(\phi \to \phi') \quad \text{is constant along the chain, i.e. t-independent}$$

3. **Probability (density)**

$$\int D\phi' \; T(\phi \to \phi') = 1 \quad \text{and} \quad T(\phi \to \phi') >= 0$$

4. **Ergodic (& irreducible)**

$$T(\phi \to \phi') > 0 \quad \forall \phi, \phi' \in \Omega$$

A chain is completely specified by the starting distribution $P_0(\phi^{(0)})$ and $T(\phi \to \phi')$

Why Markov chains ?

T is a linear map: $T : \square \to \square$, where $\square$ is the linear space of real functions on Ω.

This contains the subset of probability distributions P_Ω . Along a Markov chain

$$P_{n+1}(\phi') = (TP_n)(\phi') = \int D\phi\, P_n(\phi) T(\phi \to \phi')$$

with

$$P^{(n)} \in P_\Omega$$

Equilibrium distribution

Given an ergodic Markov chain with transition probability T, the limit

$$\lim_{t\to\infty} P^{(t)} = \lim_{t\to\infty} T^t P^{(0)} = \Pi \in P_\Omega$$

exits, is unique and independent on $P^{(0)} \in P_\Omega$.

Im particular, Π , is the unique fixed point of the chain, i.e. $(TP) = P \Leftrightarrow P = \Pi$

Remark

This is the consequence that T has a unique eigenvalue $\lambda_0 = 1$ and

$\lambda_0 > |\lambda_1| \geq |\lambda_2| \geq \ldots$, where $T v_n = \lambda_n v_n$ and $v_0 = \Pi$

$$P^{(t)} = \Pi + \sum_{n>0} c_{t,n} (\lambda_n)^n v_n \overset{t\to\infty}{=} \Pi + O(e^{-t/\tau^{exp}}) \quad \tau^{exp} = 1/\ln|\lambda_1|$$

Detailed balance condition

How can we find a T that has the desired distribution Π as equilibrium distribution ?

If T is ergodic, a sufficient (but not necessary) condition is detailed balance

$$\Pi(\phi')T(\phi' \to \phi) = \Pi(\phi)T(\phi \to \phi')$$

Proof:

Integrate both sides over ϕ and use $\int D\phi\, T(\phi' \to \phi) = 1.$

This gives the stability or fix point condition

$$\Pi(\phi') = (T\Pi)(\phi')$$

Since T is ergodic, its fixed point is unique and corresponds to its equilibrium distribution

Remarks

- If T satisfies detailed ballance or stability but is not ergodic, the convergence for large t is not guaranteed
- We can combine transition probabilities:

  $$T = T_1 \circ T_2 \circ \dots,$$

  with Ti not ergodic but satisfies detailed balance and such that T is ergodic. T will automatically satisfy the stability condition and converge to PI.

Metropolis-Hastings algorithm

A simple way to satisfy detailed balance is given by

$$T(\phi \to \phi') = P_C(\phi \to \phi')P_A(\phi \to \phi')$$

- A candidate ϕ' is proposed from ϕ with probability P_C
- ϕ' is accepted as the next step in the chain with probability P_A
- If ϕ' is rejected, ϕ is the next element, i.e. it is repeated in the chain

Acceptance probability

We can ensure detailed balance for any choice of Pc by taking

$$P_A(\phi \to \phi') = \min\left[1, \frac{\Pi(\phi')P_C(\phi' \to \phi)}{\Pi(\phi)P_C(\phi \to \phi')}\right]$$

If $P_C(\phi \to \phi') = P_C(\phi' \to \phi)$ (symmetric proposal)

$$P_A(\phi \to \phi') = \min\left[1, \frac{\Pi(\phi')}{\Pi(\phi)}\right]$$

Other PA are in principle possible but have lower acceptance.

Metropolis-Hastings algorithm

Remarks

- only the relative probabilities $\Pi(\phi')/\Pi(\phi)$ are needed to construct T

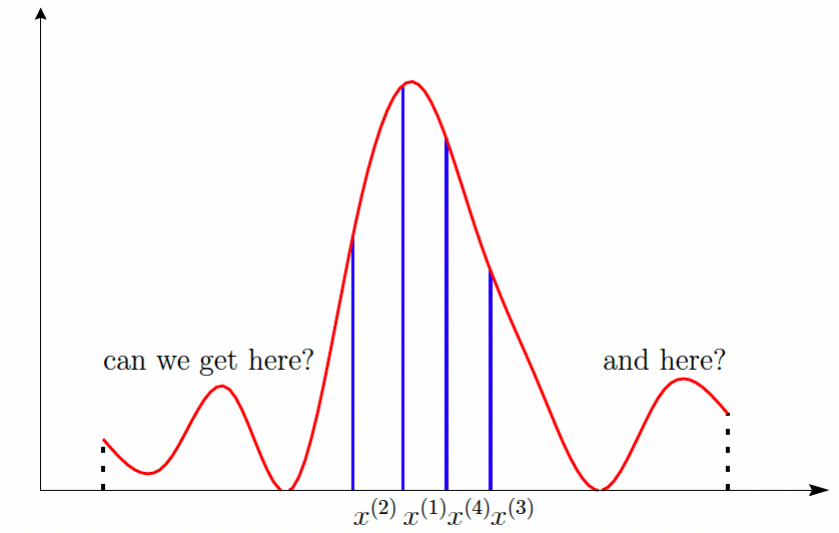
$$\Rightarrow \text{ no need for normalization of } \Pi$$

- We cannot use Markov chains to compute integrals directly, only ratios

$$\langle \square \rangle = \frac{\int D\phi \; \Pi(\phi)\square(\phi)}{\int D\phi \; \Pi(\phi)}$$

In practice there may be challenges in assuring that T is ergodic

-> can lead to improper sampling and biased results



Simulating lattice phi4-theory

Action

$$S = \sum_x \left[\sum_{\mu=0}^{D-1} \frac{1}{2} \left(\frac{\phi(x+\mu\hat{)} - \phi(x)}{a} \right) + \frac{m_0^2}{2}\phi^2(x) + \frac{g_0}{4!}\phi^4(x) \right]$$

Accept-reject step

1. Set $\phi(x) = \phi_0(x)$
2. Propose $[\Delta > 0,\ r \in [0,1)]\ \phi'(x) = \phi(x) + \Delta(r - \frac{1}{2})$
3. Accept ϕ' or keep ϕ according to

$$P_A = \min\left[1, e^{-\delta S}\right] \qquad \delta S = S(\phi') - S(\phi)$$

δS only involves $\phi'(x), \phi(x), \phi(x \pm \hat{\mu})$

1. Repeat 2. & 3. for all points x [sweep]
2. Skip k sweeps (thermalization) so that

$$P(\phi^{(t)}) \propto e^{-S(\phi^{(t)})} \qquad \Rightarrow \overline{\square} = \frac{1}{N} \sum_{t=k+1}^{N+k} \square(\phi^{(t)}) \qquad \Rightarrow \overline{\square} = \langle \square \rangle + O(1/\sqrt{N})$$

Autocorrelations

Subsequent states in a Markov chain are correlated

$$\langle\langle \square^{(k)} \square^{(l)} \rangle\rangle \neq \langle\langle \square^{(k)} \rangle\rangle \langle\langle \square^{(l)} \rangle\rangle \qquad \square \equiv \square(\phi^{(k)}) \qquad \langle\langle \cdot \rangle\rangle \equiv \text{avg. indp. chains}$$

The error on time-averges

$$\sigma^2(\overline{\square}) = \langle\langle (\overline{\square} - \langle \square \rangle)^2 \rangle\rangle = \frac{1}{N^2} \sum_{k,l=1}^{N} \langle\langle \square^{(k)} \square^{(l)} \rangle\rangle - \langle \square \rangle^2 \qquad [\langle\langle \square^{(k)} \rangle\rangle = \langle \square \rangle]$$

can be written as

$$\sigma^2(\overline{\square}) = \frac{2\tau^{int,\square} \text{var}(\square)}{N} \qquad \text{with} \quad \text{var}(\square) = \langle \square^2 \rangle - \langle \square \rangle^2$$

Integrated autocorrelation time is given by

$$\tau^{int,\square} = \frac{1}{2} \left[1 + 2 \sum_{t=1}^{N-1} \frac{\Gamma^{(\square)}(t)}{\Gamma^{(\square)}(0)} \right]$$

and the **autocorrelation function**

$$\Gamma^{(\square(t))} = \langle\langle \square^{(t+i)} \square^{(i)} \rangle\rangle - \langle \square \rangle^2$$

- Error scales via $N/2\tau^{int}$
- for time-homogeneous chains the function only depends on the distance in Markov time

Autocorrelations

Spectral decomposition

$$\Gamma^{(\square)}(t) = \sum_{n>0} b_{n,\square} e^{-t/\tau_n} \qquad \tau_n = -1/\ln|\lambda_n| \qquad [\lambda \text{ eignv. of } T]$$

- τ_n only depends on the properties of the Markov chain

  $\Rightarrow \tau^{exp} = \tau_1$ is the "slowest" mode to decorrelate
- $b_{n,\square}$ determines the coupling of O to the n-th mode

  $\Rightarrow$ it can vary significantly among observables

Estimate of the autocorrelation function

$$\overline{\Gamma^{(\square)}}(t) = \frac{1}{N-t} \sum_{i=1}^{N-t} [(\square^{(i+t)} - \overline{\square})(\square^{(i)} - \overline{\square})]$$

Estimate of the integrated autocorrelation time

Relative error on the autocorrelation function grows exponentially -> we must choose a cutoff W

$$\tau^{int,\square,W} = \frac{1}{2}\left[1 + 2\sum_{t=1}^{W} \frac{\Gamma^{(\square)}(t)}{\Gamma^{(\square)}(0)}\right]$$

and find a compromise between statistical and systematic error

Autocorrelations

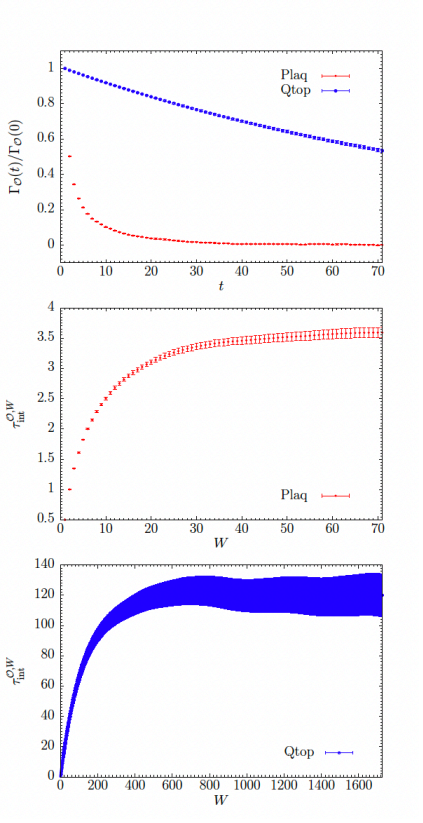
- Autocorrelations (AC) are unavoidable in Markov Chain Monte Carlo (MCMC)
- A proper estimate of AC is curical
  - no AC -> no error -> no result

Ideal:

- Length of simulation > O(100) * tau_exp
  - For thermalization O(10)*tau_exp

Otherwise risks:

- incomplete thermalization:
  - wrong sampling and a biased result
  - wrong estimation of AC, underestimation of errors



Estimation of tau^exp

Look for the observable O_slow with the largest AC, i.e. is very sensitive to slow modes

take $\tau^{\exp} \sim \tau^{int,\square^{\text{slow}}}$

Simulating Lattice QCD

Feynman Pathintegral

$$\langle \square \rangle = \frac{1}{Z} \int DU \; D\overline{\psi} D\psi \; e^{-Sg[U]} e^{-\overline{\psi} D[u] \psi} \cdot \square[U, \overline{\psi}, \psi]$$

$$= \frac{1}{Z} \int DU e^{-Sg[U]} \det(D[U]) \cdot \square[U, \overline{\psi}, \psi]$$

- pure gauge theory, quenched Simulation, det D = 1
- Action local : requires O(1) operations for a single link U update
  - Heat-bath
  - Overrelaxation
- Action non-local: requires O(V) operations for a single link uodate
  - Global update
- Global update must be **coherent** otherwise

$$\delta S \propto V \quad \Rightarrow P_A \propto \exp(-\delta S) \sim 0$$

Hybrid Monte Carlo

Add auxiliarry momentas

$$\pi(x,\mu) = T^a \pi^a(x,\mu) \in su(3) \qquad (\pi,\pi) = \sum_{x,\mu,a} |\pi^a(x,\mu)|^2$$

Hamiltonian system

$$\langle \square \rangle = \frac{1}{Z} \int DUD\pi \, e^{-S[U]} \, e^{-(\pi,\pi)/2} \qquad \text{with} \qquad \int D\pi e^{-(\pi,\pi)/2} = 1$$

$$= \frac{1}{Z} \int DUD\pi \, e^{-H[\pi,U]} \qquad \text{with} \qquad H = \frac{1}{2}(\pi,\pi) + S[U]$$

Now, we can use Molecular dynamics to update:

$$U(x,\mu) \rightarrow U(x,\mu)(t) \qquad \pi(x,\mu) \rightarrow \pi(x,\mu)(t)$$

using Hamiltons equations

$$\partial_t U(x,t) = \pi(x,\mu) U(x,\mu)$$

$$\partial_t \pi(x,\mu) = -F(x,\mu) \qquad F(x,\mu)^a = \partial_{x,\mu} S[U] \qquad [\partial^a{}_{x,\mu} \, U(y,\nu) = \delta_{xy} \, \delta_{\mu\nu} T^a U(x,\mu)]$$

Note that

$$P_H \propto exp{-H} \quad \text{and} \quad P_S \propto exp{-S} \quad \text{are equivalent for sampling } \square(U)$$

Introducing t is legitimate

$$\partial_t H = 0 \qquad \Rightarrow P_H(\pi(0), U(0)) = P_H(\pi(\tau), U(\tau))$$

Hybrid Monte Carlo

Ideal HMC algorithm

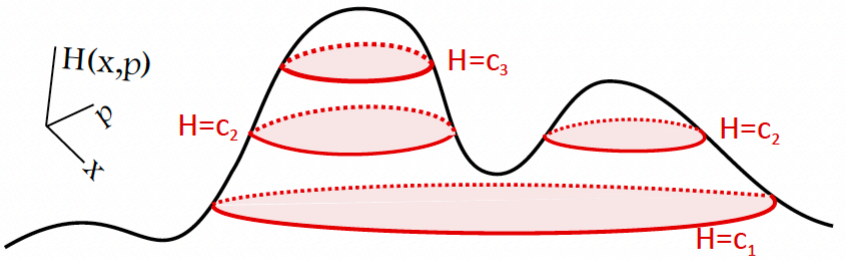
1. Start from a gauge-field U(0)
2. Sample a momentum field P(0) from the Gaussian distribution

$$P_\pi = e^{(\pi,\pi)/2}/Z$$

3. Solver Hamilton eqs. for a time t

$$(\pi(0), U(0)) \rightarrow (\pi(\tau), U(\tau))$$

4. Repeat 2. and 3. taking U(0)=U(t)



Hybrid Monte Carlo

Ergodicity:

First step is given by a heat-bath for the momenta's

$$P_\pi P_H = P_H$$

but the step is not ergodic in the total phase-space (only in the momentum part)

Second step is given by the Hamilton evolution

$$P_{MD}((\pi, U) \rightarrow (\pi', U')) = \delta(\pi' - \pi(\tau))\delta(U' - U(\tau))$$

it follows

$$P_{MD} P_H = P_H$$

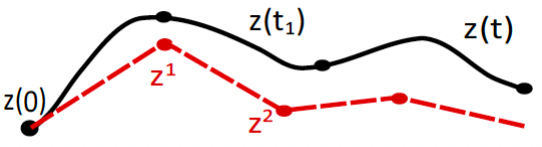
but is also not ergodic (H=const.)

However, the combination of first and second step is assumed to be ergodic

$$T = (P_{MD} P_\pi) \quad \text{has a fixed point and} \quad P_H \quad \text{is equilibrium distribution}$$

In practice

- Hamilton equations can not be solved exactly
  - use numerical integration
- H is not conserved
  - bias in equilibrium distribution



Hybrid Monte Carlo on the computer

HMC algorithm [Duane et al. 87]

1. Start from a gauge-field U(0)
2. Sample a momentum field P(0) from the Gaussian distribution

$$P_\pi = e^{(\pi,\pi)/2}/Z$$

3. Solve Hamilton eqs. numerically for a time t

$$(\pi^{(i)}, U^{(i)}) = (\pi(0), U(0)) \rightarrow (\pi(\tau), U(\tau)) = (\pi^{(f)}, U^{(f)})$$

4. Accept the configuration U'=U(t) with probability

$$P_A = \min[1, e^{\delta H}] \quad \text{with} \quad \delta H \equiv H(\pi^{(f)}, U^{(f)}) - H(\pi^{(i)}, U^{(i)})$$

   If reject start from the initial one U'=U(0)

5. Repeat 2. - 4. taking U(i)=U' and iterate

Hybrid Monte Carlo on the computer

- The numerical solution of Hamilton equations is used as a proposal in a Metropolis step
- the accept-reject step guarantees that PH is the equilibrium distribution, even if

$$\delta H \neq 0$$

- For the correctness of the HMC, the numerical integrator must preserve two **key properties** of Hamilton dynamics

  1. Time-reversibility

  $$P_{MD}((\pi, U) \rightarrow (\pi', U')) = P_{MD}((-\pi', U') \rightarrow (-\pi, U))$$

  Guarantess a symmetric proposal

  2. Phase-space measure preservation

  $$D\pi(0)\, DU(0) = D\pi(\tau) DU(\tau)$$

  in conjuction with 1. guarantees detailed balance

Molecular Dynamics integration

Hamiltonian

$$H(p,q) = \frac{1}{2}p^2 + S(q) = T(p) + S(q)$$

Time-evolution operator

$$\exp\left(\tau\frac{d}{dt}\right) f(p(t),q(t)) = f(p(t+\tau),q(t+\tau))$$

(taylor expansion) We can write

$$\exp\left(\tau\frac{d}{dt}\right) = \exp\left(\tau\left[\frac{dp}{dt}\frac{\partial}{\partial t} + \frac{dq}{dt}\frac{\partial}{\partial t}\right]\right) = \exp\left(\tau\left[-\frac{\partial H}{\partial q}\frac{\partial}{\partial t} + \frac{\partial H}{\partial p}\frac{\partial}{\partial t}\right]\right) \equiv \exp\left(\tau\hat{H}\right)$$

Hamiltonian vector field

$$\hat{H} = \left[\frac{\partial H}{\partial p}\frac{\partial}{\partial t} - \frac{\partial H}{\partial q}\frac{\partial}{\partial t}\right] = \hat{T} + \hat{S}$$

It follows

$$\exp(\tau\hat{H})H = H \quad \Rightarrow \quad \partial_t H = 0$$

and

$$\hat{T} = T'\frac{\partial}{\partial q} \quad \text{and} \quad \hat{S} = -S'\frac{\partial}{\partial p}$$

Molecular Dynamics integration

Integrable steps

$$\hat{e^{\tau T}} : f(p,q) \to f(p, q + \tau T'(p))$$

$$\hat{e^{\tau S}} : f(p,q) \to f(p - \tau S'(q), q)$$

Measure preserving/Volume preserving

$$J(\hat{e^{\tau T}}) = \frac{\partial \hat{e^{\tau T}}(p,q)}{\partial(p,q)} = \det \begin{bmatrix} 1 & \tau T''(p) \\ 0 & 1 \end{bmatrix} = 1$$

$$J(\hat{e^{\tau S}}) = \frac{\partial \hat{e^{\tau S}}(p,q)}{\partial(p,q)} = \det \begin{bmatrix} 1 & 0 \\ -\tau S''(q) & 1 \end{bmatrix} = 1$$

- $\hat{\exp(\tau T)}$ and $\hat{\exp(\tau S)}$ are exactly integrable for any τ
- can be combined to built symplectic integrators i.e. time-reversible and measure preserving

Leap Frog

$$[I_{LPFR}\ (h)]^n = \left(\hat{e^{\frac{h}{2}S}}\hat{e^{hT}}\hat{e^{\frac{h}{2}S}}\right)^n$$

which is reversible by construction and volume preserving

Integration error of symplectic integrators

Using Baker-Campbell-Hausdorff (BCH) formula

$$\ln(e^A e^B) = (A + B) + \frac{1}{2}[A, B] + \frac{1}{12}([A, [A, B]] - [B, [A, B]]) + \ldots$$

we find

$$[I_{LPFR}\ (h)]^n = \left(\exp[(\hat{T} + \hat{S})h - \frac{1}{24}([\hat{S}, [\hat{S}, \hat{T}]] + 2[\hat{T}, [\hat{S}, \hat{T}]])h^3 + O(h^5)]\right)^{\tau/h}$$

$$= \left(\exp[\tau((\hat{T} + \hat{S}) - \frac{1}{24}([\hat{S}, [\hat{S}, \hat{T}]] + 2[\hat{T}, [\hat{S}, \hat{T}]])h^2 + O(h^4))]\right)$$

$$\equiv \exp(\tau \tilde{H}) = \exp(\tau(\hat{T} + \hat{S}) + O(h^2))$$

Shadow Hamiltonian

The commutator of two Hamiltonian vector fields is a Hamiltonian vector field

$$\tilde{H} = \frac{\partial \tilde{H}}{\partial p}\frac{\partial}{\partial q} - \frac{\partial \tilde{H}}{\partial q}\frac{\partial}{\partial p}$$

Symplectic intergrators exactlu conserve a shadow Hamiltonian

Shadow Hamiltonian

It holds

$$[\hat{H_1}, \hat{H_2}] = \hat{H_3}$$

and its follows

$$H_3 = (H_1, H_2)_p = \frac{\partial H_1}{\partial p}\frac{\partial H_2}{\partial q} - \frac{\partial H_1}{\partial q}\frac{\partial H_2}{\partial p}$$

Now to find the shadow Hamiltonian, replace the commutators with Poisson brackets and it follows

Leap Frog

$$\Delta H_{LPFR} = \frac{1}{24}[(S,(S,T))_p + 2(T,(S,T))_p]h^2 + O(h^4)$$

$$= -\frac{1}{24}(S'^2(q) - 2p^2 S''(q))h^2 + O(h^4)$$

with

$$S'^2(q) = F^2 \quad \text{and} \quad S''(q) = F'$$

Remarks:

- BCH only gives an asymtotic expansion for tildeH
- existence of a conserved Hamiltonian tildeH along the trajectory means

$$\delta H = (H^{(f)} - \tilde{H}^{(f)}) - (H^{(i)} - \tilde{H}^{(i)}) = (\Delta H^{(f)} - \Delta H^{(i)}) = O(h^2)$$

Generalization of MD integrators

Second minimal norm scheme (OMF2)

$$I_{OMF2}(h) = e^{\lambda h \hat{S}} e^{h/2 \hat{T}} e^{(1-2\lambda) h \hat{S}} e^{h/2 \hat{T}} e^{\lambda h \hat{S}}$$

with

$$\Delta H_{OMF2} = (c_1(\lambda)(S,(S,T))_p + c_2(\lambda)(T,(S,T))_p) h^2 + O(h^4)$$

$$\text{Minimizing } c_1^2 + c_2^2 \quad \text{gives} \quad \lambda \approx 0.19$$

Fourth order integrator (OMF4) with 11 stages

$$I_{OMF4}(h) = e^{r_0 h \hat{S}} \cdots e^{r_0 h \hat{S}} \quad \Delta H_{OMF4} = O(h^4)$$

Remarks

- Measuring $var(\Delta H)$ in simulations and minimizing it allows for a systematic optimization

  [Clark et al. 11]

Multiple time-scale integration

Multiple actions

$$H(p,q) = \frac{1}{2}p^2 + S_1(q) + S_2(q) \qquad ||F_2|| \ll ||F_1||$$

If Cost(F2) >> Cost(F1) it may be convenient to use different step sizes h

Nested integrators [Sexton, Weingarten 92]

$$I(h) = e^{\frac{h}{2}\hat{S}_2}\left(e^{\frac{h}{2m}\hat{S}_1}e^{\frac{h}{m}\hat{T}}e^{\frac{h}{2m}\hat{S}_1}\right)^m e^{\frac{h}{2}\hat{S}_2}$$

Shadow Hamiltonian

$$\Delta H = [\alpha F_2^2 + \beta F_2' + \beta F_1 F_2 + \frac{1}{m^2}(\alpha F_1^2 + \beta F_1')]h^2 + O(h^4)$$

Remarks

- Correlation term between F1 and F2 is not suppressed by m

  -> efficiency depends on correlation between forces

- in lattice QCD,

  $$||F_G|| \gg ||F_{F,1}||$$

  and opposite for their cost

Some Remarks on MD integration

Gauge group integration

$$e^{h\hat{T}} : U(x,\mu) \rightarrow e^{h\pi(x,\mu)}U(x,\mu) \quad \pi(x,\mu) \rightarrow \pi(x,\mu)$$

$$e^{h\hat{S}} : U(x,\mu) \rightarrow U(x,\mu) \quad \pi(x,\mu) \rightarrow \pi(x,\mu) - hF(x,\mu)$$

Measure preservation

$$\langle e^{-\delta H} \rangle = 1 \quad \delta H = H^{(f)} - H^{(i)}$$

Reversibility:

$$\Delta = ||U' - U|| \quad (\pi', U') = F \circ [I(h)]^n \circ F \circ [I(h)]^n(\pi, U)$$

with

$$F(\pi, U) = (-\pi, U)$$

is violated by rounding errors

$$\Delta \propto h^{\nu} \quad \text{with} \quad \nu > 0 \quad \text{Liapunov exponent}$$

A too large h

$$\nu \propto h$$

and MD integration becomes unstable

Remarks on MD integration

Acceptance probability

$$P^{acc} = \langle \min[1, e^{-\delta H}] \rangle \overset{V\to\infty}{=} \operatorname{erfc}(\sqrt{\sigma^2(\delta H)/8})$$

With

$$\sigma^2(\delta H) = \langle (\delta H)^2 \rangle - \langle \delta H \rangle^2 \propto V h^{2n}$$

To tune the algorithm:

- Select stable integrator
- Minimize cost per trajectory at constant acceptance rate
  - Requires stable integrator
- $$P^{acc} = const. \Rightarrow \sigma^2 = const. \Rightarrow h \propto V^{-1/2n}$$

Critical slowing down

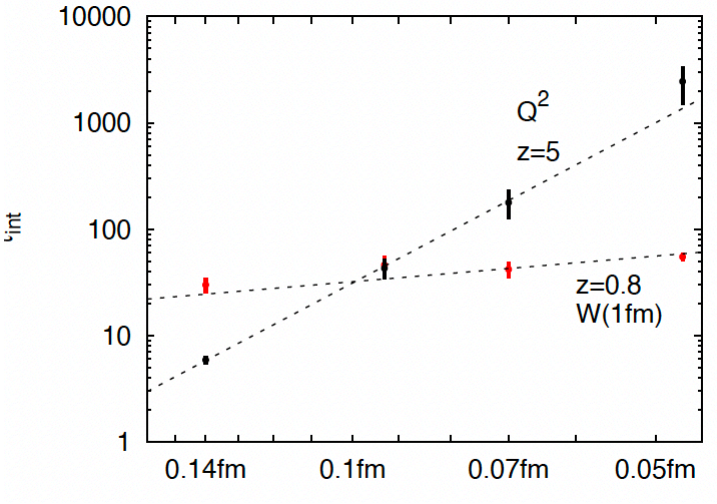
Approaching the continuum limit

$$\tau^{int,\square} \propto a^{-z}$$

where z depends on the algorithms

Status

- certain algorithms can be analyzed as QFTs :
- Simulation time is the (D+1)th dimension [Parisi, Wu 81, Zinn-Justin 86]
- HMC can not be analyzed this way: empirically z=2
- this might be true if topology issue is absent

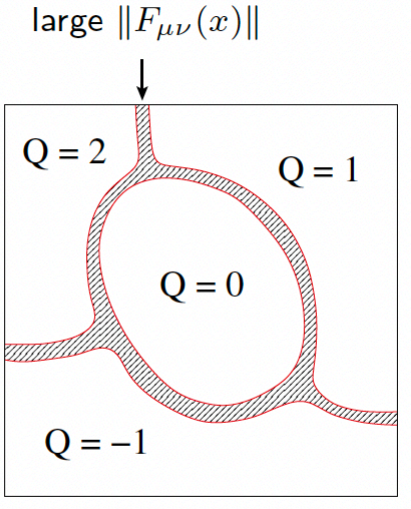


Topology freezing

- continuum path-integral with periodic bcs. for Fmunu is a sum over disconnected topological sectors
- on the lattice the field space "between" sectors rapidly vanishes as a->0
- Probability of changing sector in simulations rapidly deteriorating

  $\Rightarrow \tau^{int,Q}$ increases dramatically

- Ergodicity in simulations is compromised and so the results for any observablee
- One possible way, is to open the boundaries in time
- ...



field space

Fermions in simulations

Lattice QCD path-integral

$$\langle \square \rangle = \frac{1}{Z} \int DU \; D\overline{\psi} \; D\psi \; e^{-Sg[U]} e^{-\overline{\psi} D[U] \psi} \square[U, \overline{\psi}, \psi]$$

$$= \frac{1}{Z} \int DU \; e^{-Sg[U]} \det(D[U]) \square[U]$$

Fermionic observable

$$\square[U, \overline{\psi}, \psi] \rightarrow \square'[U] = \square[U, \frac{\partial}{\partial \eta}, \frac{\partial}{\partial \overline{\eta}}] e^{\overline{\eta} D^{-1}[U] \eta}$$

Determinant as observable:

$$\langle \square \rangle = \frac{\langle \det(D[U]) \square'[U] \rangle}{\langle \det D[U] \rangle}$$

- Calculation of det(D[U]) requires O(V^3) operations
- Overlap problem -> large statistical fluctuations

Fermions in simulations

Lattice QCD path-integral

$$\langle \square \rangle = \frac{1}{Z} \int DU \; D\overline{\psi} \; D\psi \; e^{-Sg[U]} e^{-\overline{\psi} D[U] \psi} \square[U, \overline{\psi}, \psi]$$

$$= \frac{1}{Z} \int DU \; e^{-Sg[U]} \det(D[U]) \square[U]$$

Fermionic observable

$$\square[U, \overline{\psi}, \psi] \rightarrow \square'[U] = \square[U, \frac{\partial}{\partial \eta}, \frac{\partial}{\partial \overline{\eta}}] e^{\overline{\eta} D^{-1}[U] \eta}$$

Determinant in Metropolis

$$P_A \propto e^{-Sg[U'] - Sg[U]} \frac{\det D[U']}{\det D[U]}$$

- for a single link update

  $$\det(D[U + \delta U] D[U]^{-1})$$

  requires O(V) operation

  for a full sweep follows propotional to V^2
- det D[U] must be real and positiv

Fermions in simulations

Lattice QCD path-integral

$$\langle \square \rangle = \frac{1}{Z} \int DU \; D\overline{\psi} \; D\psi \; e^{-Sg[U]} \, e^{-\overline{\psi} D[U] \psi} \, \square[U, \overline{\psi}, \psi]$$

$$= \frac{1}{Z} \int DU \; e^{-Sg[U]} \det(D[U]) \square[U]$$

Fermionic observable

$$\square[U, \overline{\psi}, \psi] \rightarrow \square'[U] = \square[U, \frac{\partial}{\partial \eta}, \frac{\partial}{\partial \overline{\eta}}] e^{\overline{\eta} D^{-1}[U] \eta}$$

Determinant as effective action

$$S^{eff} = -\text{tr} \; \ln(D[U]) \qquad \Rightarrow \qquad F^{eff} = -\text{tr}(D[U]^{-1} \, \partial D[U])$$

- Difficult and impractical to mae the algorithm efficient (and exact)
- det D[U] must be real and positiv

Fermions in simulations

Pseudo-fermions [Weingarten, Petcher 81]

$$\det(D)^2 = \det(Q^2) \propto \int D\phi^\dagger D\phi e^{-\phi^\dagger Q^{-2}\phi} \qquad [Q = \gamma_5 D = Q^\dagger]$$

- det Q^2 is expressed in terms of a bosonic Gaussian integral with pseudo-fermions interacting non-locally
- $$\det(Q)^* = \det(Q^\dagger) = \det Q \in \mathbb{R}$$

  but for Wilson quarks not necessarily positiv
- positivity of the fermion kernel is needed for the convergence of the integral
- two degenerate quarks guarantee positivity and allow for an easy pseudo-fermion generation
- more difficult for single quarks

HMC with pseudo-fermions

Heat-bath

$$P_\eta[\eta] \propto \int D\eta^\dagger D\eta \; e^{-\eta^\dagger \eta} \qquad \Rightarrow \qquad \phi = Q\eta \qquad \Rightarrow P_\phi[\phi]$$

- At the beginning of a trajectory, we generate pseudo-fermions from Gaussian fields
- Averaging results over many trajectories effectively samples the contribution from the pseudo-fermion integral

Hamiltonian

$$H = \frac{1}{2}(\pi, \pi) + S \qquad S = S_G + S_{pf} \qquad S_{pf} = (Q^{-1}\phi, Q^{-1}\phi)$$

Dynamics

$$\partial_t U(x,\mu) = \pi(x,\mu)U(x,\mu) \qquad \partial_t \pi(x,\mu) = -F(x,\mu) \qquad F^a(x,\mu) = \partial^a{}_{x,\mu} S$$

Pseudo-fermions are held fixed during the Hamiltonian evolution

Fermionic forces

$$F_{pf}{}^a{}_{x,\mu} = \partial^a{}_{x,\mu}(\psi, \psi) = -2\mathrm{Re}(\chi, (\partial^a{}_{x,\mu} Q)\psi)$$

with

$$\psi = Q^{-1}\phi \qquad \text{and} \qquad \chi = Q^{-1}\psi$$

Challenges of simulating fermions

Solving the Dirac equation

- Computing Spf and Fpf requires solving linear systems:

$$D\chi = \eta \quad \Rightarrow \quad \chi = D^{-1}\eta$$

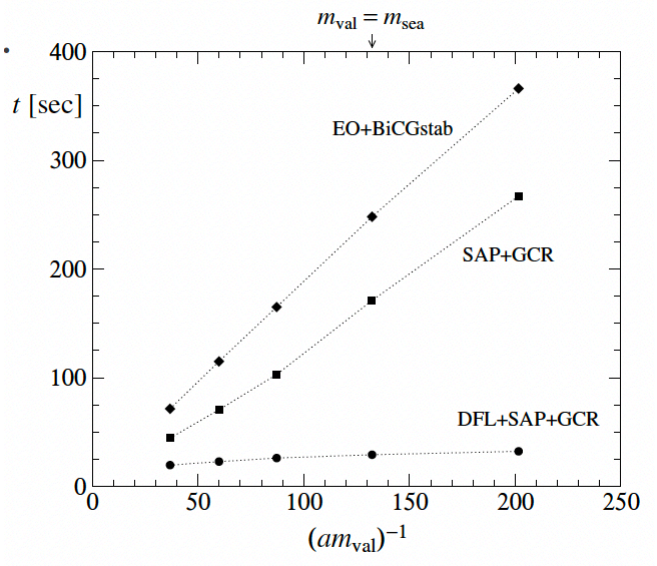
- D is a sparse matrix :

$$D\chi \quad \text{requires O(V) operations}$$

  iterative solvers are an effictive method to solve these stsmts

- Computational cost of conventional solvers rapidly grows with (a mq)^{-1}, system is quickly ill conditioned
- Multi-grid solvers are pratically solving this issues
  - lead to a speed up of O(10) - O(100)

Usually main computational challenge



Challenges of simulating fermions

Fermionic forces

- Single pseudofermion HMC is not competitive
- Consider

$$F_{pf} = (\phi, \partial Q^{-2}\phi) \qquad vs. \qquad F_F^{eff} = -2\partial \mathrm{trln}(Q)$$

  Fpf is a stochastic estimate of Feff at the start of the trajectory

$$\langle F_{pf} \rangle = F_F^{eff}$$

  - Fpf has very large fluctuations

1. $$||F_{pf}|| \gg ||F^{eff}{}_F||$$
2. $$var(||F_{pf}||) \gg var(F^{eff}{}_F)$$

  - Reall the shadow Hamiltonian

$$\Delta H_{OMF2} = (c_1 ||F||^2 + c_2 \pi^2 \partial^2 S) h^2 + O(h^4)$$

$$\text{large } var(||F_{pf}||) \quad \Rightarrow \text{ large } var(\Delta H) \quad \Rightarrow \text{ low } P_A \quad \Rightarrow \text{ small } h$$

  - large foces also trigger intabilities more easily -> limits the step size h

Even-odd reduction/preconditioning

If D only connects nearest-neighboring sites the fermionic problem can be effectively reduced to half the lattice

Checkerboard decomposition

Even or odd point

Schur decomposition

$$D = \begin{bmatrix} D_{ee} & D_{eo} \\ D_{oe} & D_{oo} \end{bmatrix} \qquad \det D = \det \hat{D} \det D_{oo} \qquad \hat{D} = D_{ee} - D_{eo} D^{-1}{}_{oo} D_{oe}$$

Even-odd action

$$S_{pf} \rightarrow S_{pfe} + S_{det} = (\hat{Q}^{-1} \phi_e, \hat{Q}^{-1} \phi_e) - 2\mathrm{tr} \ln(Q_{oo})$$

- Spfe involves pseudo-fermions residing only on the even sites of the lattice, helps in speed and reducing fluctuations
- Sdet and corresponding force can be evaluated exactly, i.e., without introducing pseudofermions

Hasenbusch mass-preconditioning

Frequency splitting [Hasenbusch 01, Hasenbusch, Jansen 03]

$$\det(Q^2) = \det(Q^2+\mu^2)\prod_{k=0}^{n-1}\det\left(\frac{Q^2+\mu_k^2}{Q^2+\mu^2{}_{k+1}}\right) \qquad 0=\mu_0<\ldots<\mu_n$$

Actions and forces

$$S_{pf,n} = (\phi_n, (Q^2+\mu_n^2)^{-1}\phi_n)$$

$$S_{pf,k} = (\phi_k, (Q^2+\mu^2{}_{k+1})(Q^2+\mu^2{}_k)^{-1}\phi_k) \qquad k=0,\ldots,n-1$$

$$(F_{pf,k})^a(x,\mu) = -2(\mu^2{}_{k+1}-\mu_k^2)\mathrm{Re}(\chi_k, (\partial^a{}_{x,\mu}Q)\psi_k)$$

- A proper tuning of mu can lead to a significant improvement
- $\text{smaller } var(||F_{pf}||) \Rightarrow \text{smaller } var(\Delta H) \Rightarrow \text{larger } P_A \Rightarrow \text{larger } h$
- for QCD, at the physical point,
- $$\mu_1 \sim m_q; \quad \mu_{k+1} \sim 10\mu_k$$

  Typically

  $$||F_{pf,k+1}|| \gg ||F_{pf,k}|| \quad \text{while opposite for their cost}$$

  Multiple time-step integration can be useful

Rational HMC

Single-quark determinante [Kennedy et al. 98, Clark, Kennedy 03]

$$\det(Q) \rightarrow \det(\sqrt{\overline{Q}^2}) = \det(W)\ \det R^{-1} \qquad W = |Q|R$$

where

$$R = r_b R^{n,\epsilon}(r_b^{-2} Q^2) \qquad \epsilon = (r_a/r_b)^2 \qquad \lambda(|Q|) \in [r_a, r_b]$$

Rational approximation

$$R^{n,\epsilon}(x) = A \frac{(x + a_1) \cdots (x + a_{2n-1})}{(x + a_2) \cdots (x + a_{2n})} \approx \frac{1}{\sqrt{x}} \qquad a_1 > \ldots > a_{2n} > 0$$

- Zolotarev rational function of degree (n,n)
- Guarantees smallest possible

$$\delta = \max_{\epsilon \le x \le 1} |1 - \sqrt{x} R^{n,\epsilon}(x)|$$

Frequency splitting (e.g. n=10)

$$\det(R^{-1}) \propto \det(P^{-1}{}_{1,4}) \det(P^{-1}{}_{5,7}) \det(P^{-1}{}_{8,10})$$

with

$$P_{k,l} = \prod_{j=k}^{l} \frac{Q^2 + \nu_j^2}{Q^2 + \mu_j^2} = 1 + \sum_{j=k}^{l} \frac{\rho_j}{Q^2 + \mu_j^2} \qquad \mu_1 > \ldots > \mu_n$$

Rational HMC

Actions and forces

$$S_{pf,k,l} = (\phi_{k,l}, P_{k,l}\,\phi_{k,l}) \qquad (F_{pf,k,l})^a(x,\mu) = \partial^a{}_{x,\mu} S_{pf,k,l}$$

Remarks

- ,We have to ensure \;$\det(Q) > 0$, this is ensured by measure the spectral range $[r_a, r_b]$ of Q^2

  If that fails, sign flip need to be included as reweighting factor (on some CLS ensembles thats the case [Mohler, Schaefer 19])

- Choose a large enough n to have a good approximation of R
- For Wilson quarks the RHMC is typically used for heavy quarks, (charm and strange)
- det(W) can be included in the accept-reject step or in the observable as a reweighting factor. It can be estimated stochastically

$$\langle \square \rangle_{|Q|} = \frac{\langle \square W \rangle_{R^{-1}}}{\langle W \rangle_{R^{-1}}} \qquad W = \langle e^{-\eta^\dagger [(1+Z)^{-1/2} - 1]\eta} \rangle_P$$

Costs

Approaching the continuum limit at constant physics, the cost to obtain a set of statistical independent configurations scales like

$$C \propto V a^{-4} \qquad V^{1/8} a^{-1/2} \qquad a^{-z} \propto V^{9/8} a^{4.5+z}$$

where the first term comes from the solver, the second from the integrator and the third from the autocorrelation

- OBC: z=2, PBC z=5

On a GPU machine with 4A100 this leads to

- O(1000) trajectories
- A physical volume of L=5.5 fm and a=0.05 fm
- Nf=2+1+1 twisted mass fermions at physical masses

$$C \approx 0.5 \text{ Mi node hours}$$

More techniques

- Pure-gauge algorithms
- Domain Decomposition
- Force gradient integrators
- Reweighting techniques
- Simulating chiral fermions
- Multi-level sampling
- …

References

This talk is based on Mattia Della Brida's Lattice Practice Contribution of 2021

Other references:

M. Lüscher, Computational Strategies in Lattice QCD, in Les Houches Summer School: Session 93: Modern perspectives in lattice QCD: Quantum field theory and high performance computing, pp. 331–399, 2, 2010, 1002.4232.

A. D. Kennedy, Algorithms for dynamical fermions, hep-lat/0607038.

S. Schaefer, Simulations with the hybrid Monte Carlo algorithm: Implementation and data analysis, in Les Houches Summer School: Session 93: Modern perspectives in lattice QCD: Quantum eld theory and high performance computing, pp. 401–422, 8, 2009.

S. Schaefer, Simulation Algorithms, Lattice Practices 2018

S. Schaefer, Status and challenges of simulations with dynamical fermions, PoS LATTICE2012 (2012) 001, [1211.5069].

S. Duane, A. D. Kennedy, B. J. Pendleton and D. Roweth, Hybrid Monte Carlo, Phys. Lett. B 195 (1987) 216–222.

A. Barp, F.-X. Briol, A. D. Kennedy and M. Girolami, Geometry and Dynamics for Markov Chain Monte Carlo, 1705.0289

ALPHA collaboration, U. Wolff , Monte Carlo errors with less errors, ComputPhys. Commun. 156 (2004) 143–153, [hep-lat/0306017].

ALPHA collaboration, S. Schaefer, R. Sommer and F. Virotta, Critical slowingdown and error analysis in lattice QCD simulations, Nucl. Phys. B 845 (2011) 93–119, [1009.5228].