Issues in Hybrid Monte Carlo Simulations

Stefan Schaefer

STRONGnet 2012

Sources

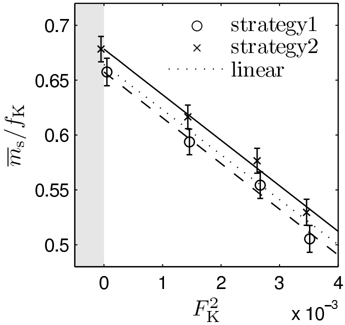
Based on work done in collaboration with Martin Lüscher

- Non-renormalizability of the HMC algorithm,
  JHEP 1104 (2011) 104
- Lattice QCD without topology barriers,
  JHEP 1107 (2011) 036
- Lattice QCD with open boundary conditions and twisted-mass reweighting,
  arXiv:1206.2809, accepted by CPC

Introduction

Continuum limit

- Continuum limit essential part of lattice computation.
- Numerical computations at various *fine* a.
  Extrapolate $a \to 0$.
- Can bring important corrections.



ALPHA'12

Overview

Rising cost as $a \to 0$

- Need more points for fixed volume
  L =const $\to N = a^{-4}$.
- Algorithms get slower.
  $\to$ physics behind this subject of this talk.

Program

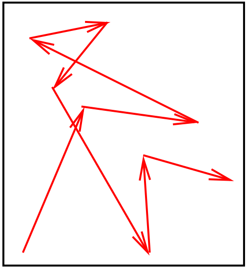
- Field theoretical framework for scaling of algorithms.
- The role of the topological charge.
- Open boundary conditions.
- Numerical results.

Lattice simulations

- Markov Chain Monte Carlo
- Sequence of field configurations

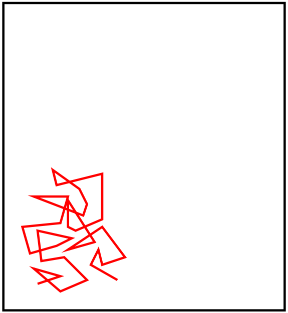
$$U_1 \to U_2 \to U_3 \to \cdots \to U_N$$

- Reliable computations need a representative sample of field space.
- Subsequent measurements are correlated.

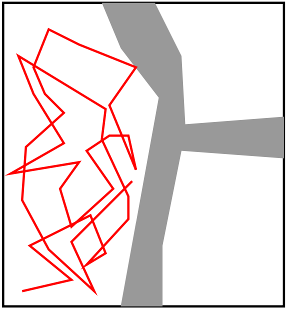


Dangers

- Algorithm is slow.
- Detectable by measuring autocorrelations.



- There are barriers in field space.
- Hard to detect.



Autocorrelations

- Sequence of field configurations

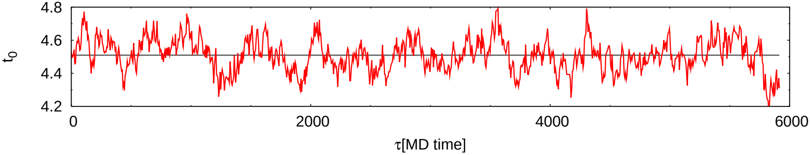
$$U_1 \to U_2 \to U_3 \to \cdots \to U_N$$

- Measurements of observables

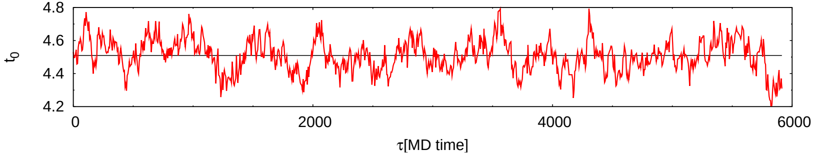
$$A_1 \to A_2 \to A_3 \to \cdots \to A_N$$

- Estimates

$$\langle A \rangle \approx \frac{1}{N} \sum_{i=1}^{N} A_i$$



Autocorrelation time



Autocorrelation function

$$\Gamma(\tau) = \langle (\boldsymbol{A}(\tau) - \bar{\boldsymbol{A}})(\boldsymbol{A}(\mathbf{0}) - \bar{\boldsymbol{A}}) \rangle$$

Integrated Autocorrelation Time

$$\tau_{\mathbf{int}}(\boldsymbol{A}) = \int_{-\infty}^{\infty} \mathbf{d}\tau\, \rho(\tau) \quad \text{with} \quad \rho(\tau) = \frac{\Gamma(\tau)}{\Gamma(\mathbf{0})}$$

Critical slowing down

Integrated autocorrelation time

$$\tau_{\text{int}}(A) = \int_0^\infty \frac{\Gamma(t)}{\Gamma(0)}\, dt$$

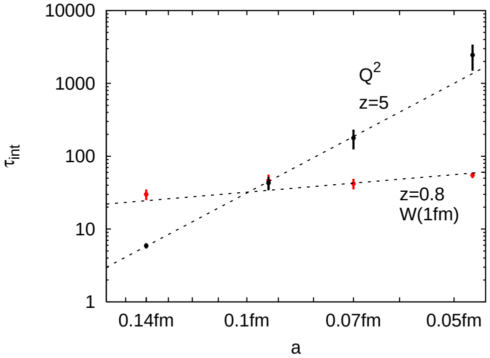
Time to make an "independent" configuration.

- How do $\Gamma(t)$ and τ_{int} scale as $a \to 0$?
- Is there universal behavior?

$$\tau_{\text{int}} \propto a^{-z}$$

- Are there barriers forming as $a \to 0$?
  $\to$topological charge

Observed scaling: Pure gauge theory

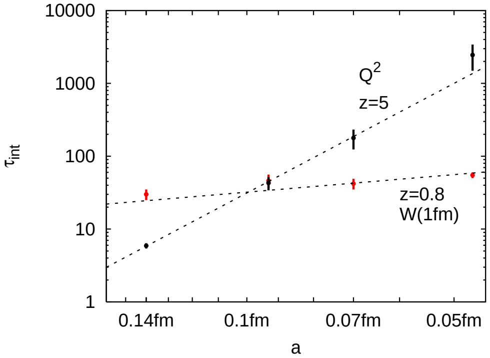


SOMMER, VIROTTA, S.S.'10

SEE ALSO DEL DEBBIO ET AL'02, LÜSCHER'10

- Pure gauge theory, Wilson action, $L = 2.4\,\text{fm}$
- $1\,\text{fm} \times 1\,\text{fm}$ Wilson loop $\rightarrow \tau_{\text{int}} \propto a^{-0.8}$
- Topological charge $Q^2 \rightarrow \tau_{\text{int}} \propto a^{-5}$

Observed scaling: Pure gauge theory



Sommer, Virotta, S.S.'10

see also Del Debbio et al'02, Lüscher'10

- Even in pure gauge theory, measurements below 0.05 fm difficult

Summary – Outlook: Part I

Summary

- An algorithm is a prescription to generate sequence of fields
$$U_1 \to U_2 \to U_3 \to \cdots \to U_N$$
- Want to study scaling as $a \to 0$.

Outlook

- Study problem with field theoretical methods.
- Example: Langevin equation as update algorithm.

Langevin equation

Field theory with a scalar field ϕ and action $S[\phi]$

$$Z = \int [d\phi] e^{-S[\phi]}$$

Langevin equation

$$\partial_t \phi = -\frac{\delta S[\phi]}{\delta \phi} + \eta_t$$

t: simulation time
η_t: Gaussian noise

- Generates fields with probability $P(\phi) \propto e^{-S[\phi]}$.
- Studied in the context of stochastic quantization. (Parisi & Wu)
- Take fields ϕ_n at time $t_n = n\tau$.

5D

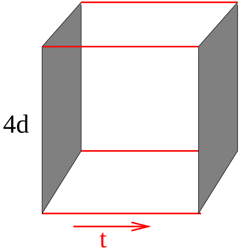
- Scaling of

$$\Gamma_A(t) = \langle A(t)A(0)\rangle - \bar{A}^2$$

- 2pt function in 5d theory
- 4d field theory + simulation time
- Analogy to 4d

$$\langle O(x)O(0)\rangle$$

- Use renormalization group analysis to study scaling of AC function.



Going to 5D

ZINN-JUSTIN'86

Basic steps

$$Z = \int [d\phi] e^{-S[\phi]}$$

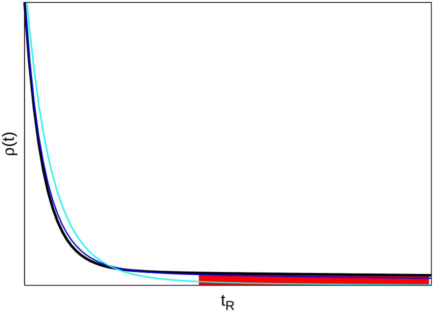
- Include differential equation constraint by delta function

$$Z' = \int [d\phi][d\eta]\delta\left[\partial_t \phi + \delta S[\phi] - \eta\right] e^{-S[\phi] - |\eta|^2/2}$$

- Replace delta function with Lagrange multipliers.
- Integrate out Gaussian noise.
- Get a renormalizable 5d field theory.

Expected scaling

- After renormalization of time, autocorrelation function scales
$$t = Z_t(g_0^2) t_R / a^2$$
- Pointwise convergence

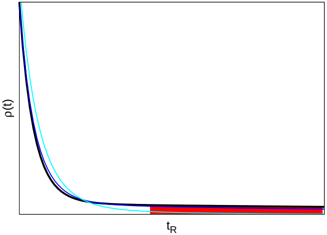


Expected scaling

- Autocorrelation time with fixed window scales.

$$\tau_{\text{int}} = \frac{1}{2}\int_0^W \rho(t)\, dt$$

- Limits $a \to 0$ and $W \to \infty$ do not necessarily commute.



Consequences for Langevin equation

$$\partial_t \phi = -\frac{\delta S[\phi]}{\delta \phi} + \eta_t$$

- Field ϕ has dimension (length)
  $\Rightarrow$ Simulation time t has dimension **(length)**^{2}
- Autocorrelations scale essentially with $\mathbf{a^{-2}}$
- Renormalizability perturbation theory also for gauge theory
  ZINN-JUSTIN&ZWANZIGER'88

In SU(3) Yang-Mills theory

$$\tau_{\text{int}} \propto g_0^{9/11} r_0^2 (1 + \mathcal{O}(g_0^2))$$ BAULIEU&ZWANZIGER'00

Towards simulations

Langevin equation

- Langevin equation not used in actual simulations
- No exact algorithm known.
- Change direction of movement on microscopic scale.
- Time has dimension **(length)**^{2}.

Hybrid Monte Carlo

- Algorithm of choice for QCD simulations.
- Exact algorithm.
- Directed movement on macroscopic scale.
- Time has dimension **(length)**.

DUANE ET AL'86

- Classical mechanics system
- Field configuration ϕ →position
- Introduce conjugate momenta
- Hamiltonian

$$H[\pi, \phi] = \frac{1}{2}\pi^2 + S[\phi]$$

Trajectory

- Choose random momentum π.
- Update by solving classical equations of motion.

$$\dot{\pi} = -\frac{\delta H}{\delta \phi} \; ; \qquad \dot{\phi} = \pi$$

Hybrid Monte Carlo

Stochastic **M**olecular **D**ynamics HOROWITZ'85-'91

$$\left.\begin{aligned}\partial_t\phi &= \pi\\ \partial_t\pi &= -\frac{\delta S}{\delta\phi} - 2\mu_0\pi + \eta\end{aligned}\right\} \Rightarrow \partial_t^2\phi + 2\mu_0\partial_t\phi = -\frac{\delta S}{\delta\phi} + \eta_t$$

η Gaussian noise with $\langle\eta_{x,t}\eta_{x',t'}\rangle = 4\mu_0\delta(t-t')\delta(x-x')$

π conjugate momenta

- $\mu_0 \to 0$ molecular dynamics
- $\mu_0 \to \infty$ Langevin equation (rescale $t \to 2\mu_0 t$)
- t has dimension (length)
- Expect $\tau_{\text{int}} \propto a^{-1}$

SMD algorithm

Algorithm — HOROWITZ'85-'91

- Momentum refreshment

$$\pi \to e^{-\gamma\delta\tau}\pi + \sqrt{1-e^{-2\gamma\delta\tau}}\eta$$

- Molecular dynamics

$$\partial_s \pi = -\frac{\delta S}{\delta \phi}; \quad \partial_s \phi = \pi$$

- Metropolis step with $\pi \to -\pi$ upon rejection

BECCARIA&CURCI'94, JANSEN&LIU'95, KENNEDY&PENDLETON'01

- $s = ta$, $\gamma = 2a\mu$
- $\gamma \to 0$, $\delta\tau \to \tau$ HMC
- $\gamma = 2a\mu$ =const $\to$ Langevin in continuum limit

SMD algorithm

Differential equation

$$\partial_t \phi = \pi \qquad ; \qquad \partial_t \pi = -\frac{\delta S}{\delta \phi} - 2\mu\pi + \eta$$

Algorithm

- Momentum refreshment

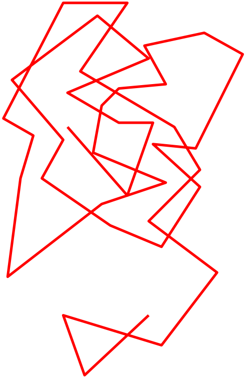
$$\pi \to e^{-\gamma\delta\tau}\pi + \sqrt{1 - e^{-2\gamma\delta\tau}}\eta$$

- Molecular dynamics

$$\partial_s \pi = -\frac{\delta S}{\delta \phi}; \quad \partial_s \phi = \pi$$

Is the HMC actually different from the Langevin?

HMC



Langevin

Renormalizability of the HMC

M. Lüscher&S.S.'11

The HMC is not renormalizable.

- UV singularity along the "light cone" at one-loop of perturbation theory.

- Not removable by local counter terms.
- Demonstrated in ϕ^4 theory.
- Most likely same in gauge theory.

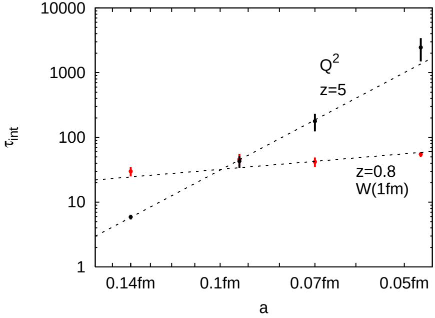
Non-Renormalizability of the HMC

- HMC is still a good algorithm.
- No statement about scaling possible.

Conjecture

- HMC and SMD fall into universality class of Langevin.
- Because of interactions, also HMC does only microscopic updates (random walk).
- SMD $\mu_0 \to \infty$ gives Langevin equation.
- Should exhibit a^{-2} scaling.

Langevin scaling?



- Topological charge.
- Smeared Wilson loop.
- No obvious scaling behavior.

Summary – Outlook: Part II

Summary

- The properties of algorithms can be analyzed using field theoretical methods.
  $\rightarrow$ renormalizable algorithms
- The Langevin equation is renormalizable.
- HMC not renormalizable $\rightarrow$ Langevin scaling.

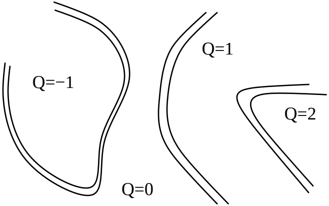
Outlook

- The special role of the topological charge.
- Numerical study.

Topological charge

$$Q = -\frac{1}{32\pi^2} \int d\,x\, \epsilon_{\mu\nu\rho\sigma} \mathrm{tr}\, F_{\mu\nu} F_{\rho\sigma}$$

- In continuum limit, disconnected **topological sectors** emerge.
- The probability of configurations "in between" sectors drops rapidly. M. LÜSCHER, '10
- Simulations get stuck in one sector.

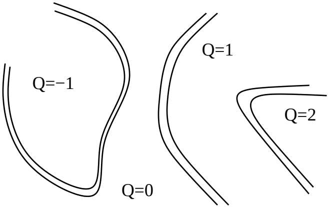


Topological charge

- **Tunneling is a cut-off effect**.
- Quasi continuous algorithms will not cure it.
- Problem for interpretation of data.
- Fixed topology introduces finite volume effects.

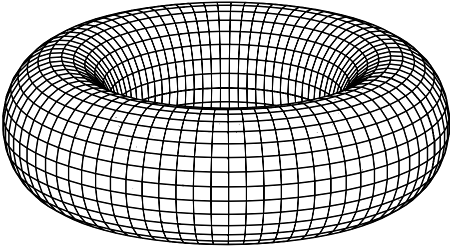
$$\langle A \rangle = \langle A \rangle_{Q=Q_0} \cdot \{1 + \mathcal{O}(V^{-1})\}$$

- Prevents simulations on fine lattices.



Boundary conditions

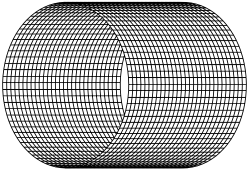
- Periodic boundary conditions do not let charge flow out of the volume.
- Field space is disconnected in continuum.



Open boundary conditions

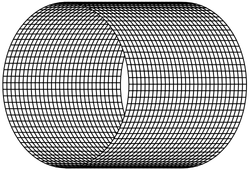
Proposed solution

- open boundary condition in time direction
  $\rightarrow$ same transfer matrix, same particle spectrum
- periodic boundary condition in spatial directions
  $\rightarrow$ momentum projection possible



Open boundary conditions

- Lattices of size $T \times L^3$.
- Neumann boundary conditions in time.



- Gauge fields

$$F_{0k}|_{x_0=0} = F_{0k}|_{x_0=T} = 0, \quad k = 1,2,3$$

- Fermion fields

$$P_+\psi(x)|_{x_0=0} = P_-\psi(x)|_{x_0=T} = 0 \qquad P_\pm = \frac{1}{2}(1 \pm \gamma_0)$$

$$\bar\psi(x)P_-|_{x_0=0} = \bar\psi(x)P_+|_{x_0=T} = 0$$

Goals of numerical study

Langevin scaling

- Demonstrate that HMC falls into universality class of Langevin.
- Find a^{-2} scaling.

Benefit of the boundaries

- For $T \to \infty$ boundary has no effect.
- Does it improve the situation for a typical sized lattice? How does this depend on a?

Numerical study

Lattice

- pure gauge theory, Wilson action
- L^4 lattices
- $L = 1.6$fm from $r_0 = 0.5$fm
- $a = 0.1$fm, 0.08fm, 0.067fm, 0.05fm, 0.04fm
- longer lattices for T dependence

Algorithms

- HMC
- SMD at fixed $\gamma = 2a\mu_0 \rightarrow$ Langevin as $a \rightarrow 0$.

Observables

Requirements

- Arguments based on renormalization.
- Need to consider quantities with continuum limit.
  $\rightarrow$ does not apply to previous plot.

- Noise can cover autocorrelations.
- Use low-noise observables.

Gradient flow

NEUBERGER'06, LÜSCHER'10, LÜSCHER&WEISZ'11

- Smoothing with **gradient flow** at fixed flow time $t = t_0$.

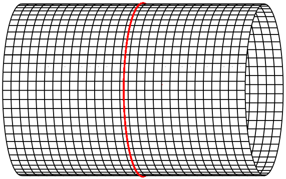
$$\partial_t V_t(x,\mu) = -g_0^2 \left[\partial_{x,\mu} S(V_t)\right] V_t(x,\mu); \quad V_t(x,\mu)|_{t=0} = U(x,\mu)$$

- Gaussian smoothing over $r \sim \sqrt{8t}$.
- Renormalized quantities with continuum limit.
- Smooth observables $\rightarrow$ long autocorrelations.

$$\overline{E} = -\frac{a^3}{2L^3} \sum_{\vec{x}} \mathrm{tr}\, G_{\mu\nu} G_{\mu\nu} \big|_{x_0=T/2}$$

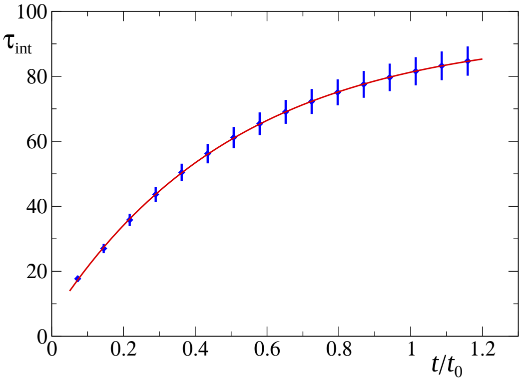
$$\overline{Q} = -\frac{a^3}{32\pi^2} \sum_{\vec{x}} \mathrm{tr}\, \tilde{G}_{\mu\nu} G_{\mu\nu} \big|_{x_0=T/2}$$

$$Q = -\frac{a^4}{32\pi^2} \sum_{x} \mathrm{tr}\, \tilde{G}_{\mu\nu} G_{\mu\nu}$$



Effect of the smoothing

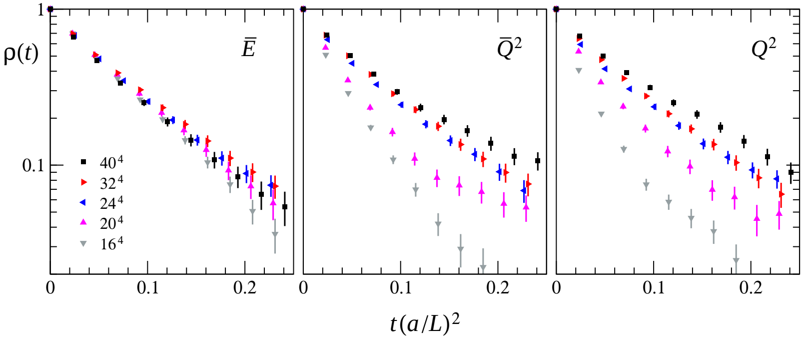
Autocorrelation time of $\bar{E}$ vs smoothing range (a=0.05fm).



- $\sqrt{8t}$ smoothing radius $\rightarrow t = t_0$ smoothing over $r \approx r_0$
- $\tau_{\rm int}$ saturates with $\tau_{\rm int} = 93 + ae^{-c/t}$.

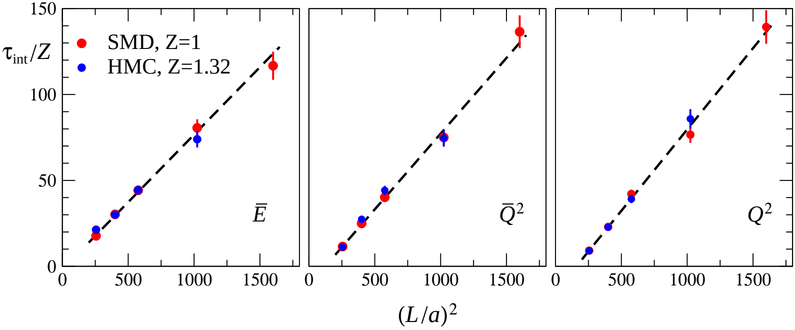
Scaling towards continuum limit

Autocorrelation function vs scaled MC time



- Energy (on time slice) shows very good scaling.
- Large cut-off effects in topological observables.

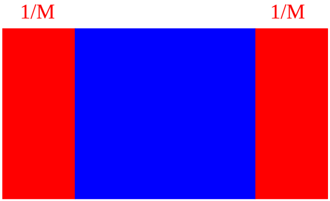
Scaling towards continuum limit: $\tau_{\rm int}$ vs a^{-2}

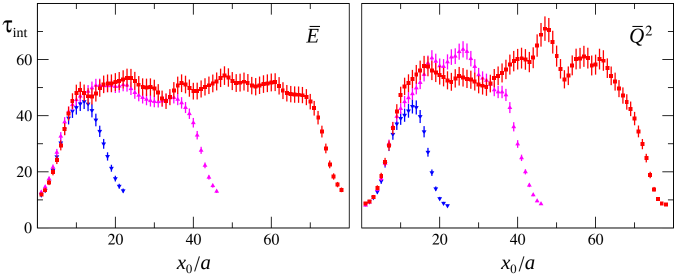


- HMC and $\mathrm{SMD}_{0.3}$ show same scaling up to a constant. $\rightarrow$ universal behavior
- Topological observables well described by $\tau_{\rm int} = c_1 + c_2/a^2$
- **Also Q^2 and $\bar{Q}^2$ show a^2 scaling for $a \rightarrow 0$.**

Large T

- Simulations so far on relatively short lattices.
- Bound. cond. break time translation invariance.
- Effect from boundaries exponentially suppressed with distance.
- Need slightly longer lattices.





- Various T for $a = 0.067$fm.
- Effect of boundary small at $x_0 \sim 0.7$fm.
- Behavior in center of short lattices representative of large T.

Large T

Finite volume

- For $T \to \infty$ the effect of the b.c. vanishes.
- But also the effect on observables vanishes as V^{-1}.

Dependence on T

- Width of distribution of Q is $\propto \sqrt{TL^3}$.
- Change of charge through boundary $\propto \sqrt{L^3}$.
  $\to$ expect $\tau_{\text{int}} \propto T$, for random walk
- For each T, there is an a from which the boundary tunneling dominates over the bulk tunneling.

Dynamical fermions

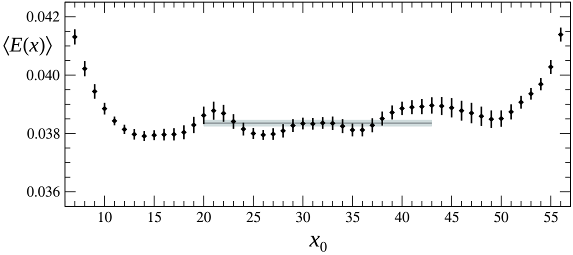
Action

- $N_{\mathrm{f}} = 2 + 1$ 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

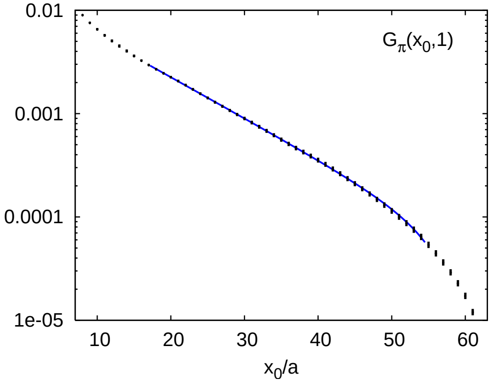
- Reweighting to avoid stability problems.
- Generated with new public `openQCD` code.
  `http://cern.ch/luscher/openQCD`

Effect of the boundary: gauge observables



- Wilson flow time $t = t_0$
- Smoothing radius $r = \sqrt{8t} \approx 0.5$ fm.
- Correlation length $1/(am_\pi) \approx 11$
- Plateau starting ~ 1 fm from boundary.

Fermions and open boundary conditions



source at $y_0/a = 1$

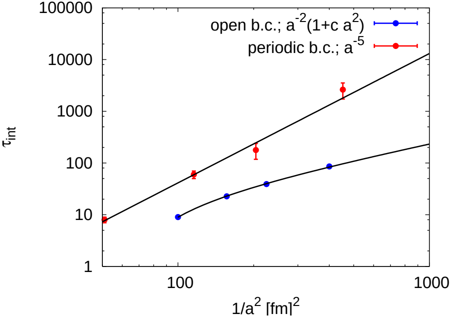
- Chiral perturbation theory with Dirichlet b.c.

$$G(x_0, y_0) \propto \sinh(m(T - x_0)) \sinh(m y_0) \qquad \text{for} \quad y_0 < x_0$$

- Valid if sufficiently away from boundary (≈ 0.5 fm).

Summary I: Open boundary conditions

- To remove the regulator from, a series of fine lattices has to be simulated.
- Particularly important for e.g. heavy quarks.
- With periodic boundary conditions, topology gets stuck
  →use open boundary in time



Summary II: Renormalizable algorithms

- Algorithms can be studied using field theory.
- The free field scaling of the HMC does not hold with interactions.
- Likely in Langevin universality class.
- No macroscopic steps.

Summary III: Cost of simulation

- Fixed volume

$$a^{-4} \text{ points}$$

- Fixed acceptance (2nd order integrator).

$$a^{-1} \text{ step size}$$

- Scaling of $\tau_{\text{int}} \to a^{-2}$ length

$$\text{Total cost } \propto a^{-7}$$

Total cost

$$\text{Total cost } \propto a^{-7}$$

Factor 2 in lattice spacing $\leftrightarrow$ factor 128 cost.