Playing with the kinetic term in the HMC

Alberto Ramos <alberto.ramos@desy.de>

NIC, DESY Zeuthen.

June 26, 2012

Special thanks to

- Patrick Fritzsch
- Laurent Lellouch
- Marina Marinkovic
- Alfonso Sastre
- Hubert Simma
- Rainer Sommer
- Francesco Virotta

Hybrid Monte Carlo

To sample the distribution $\propto e^{-S[U]}dU$ we use

$$\propto e^{-\frac{\pi^2}{2}-S[U]}dUd\pi = e^{-H(\pi,U)}dUd\pi$$

1. Generate random momenta $\pi_\mu(x) \sim \mathcal{N}(0,1)$.
2. Integrate e.o.m from MD time $\tau = 0$ to $\tau = \tau_L$

$$\begin{aligned} \dot{\pi}_\mu(x) &= -\frac{\delta H(\pi,U)}{\delta U_\mu(x)} = F_\mu(x) \\ \dot{U}_\mu(x) &= \frac{\delta H(\pi,U)}{\delta \pi_\mu(x)} = \pi_\mu(x)U_\mu(x) \end{aligned}$$

3. Accept with probability $\propto \min\left\{1, e^{-\Delta H}\right\}$.

$U^{(1)} \to U^{(2)} \to \ldots$ are samples of the equilirium distribution $\propto e^{-S[U]}dU$

Equilibrium properties unchanged by arbitrary kinetic term

$$H(\pi,U) = K(\pi) + S[U]$$

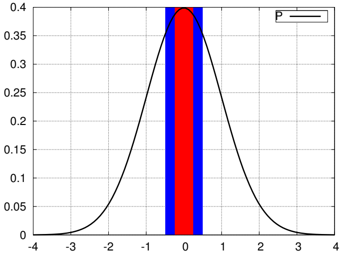
The role of the kinetic term

Kinetic term $\Longrightarrow$ Distribution for $\dot{Q}$ $\Longrightarrow$ Motion in phase space.

$$H(P,Q) = \frac{P^2}{2} + \dots$$

$$\begin{aligned} \dot{P} &= F(Q) \\ \dot{Q} &= K'(P) = P \end{aligned}$$

- $\dot{Q} \sim \mathcal{N} \propto e^{-P^2/2}$ (equilibrium distribution).
- 25% $\dot{Q} \in (-0.25, 0.25)$.
- 38% $\dot{Q} \in (-0.5, 0.5)$.



- Large $\dot{Q}$ $\Longrightarrow$ small acceptance.
- Small $\dot{Q}$ $\Longrightarrow$ large autocorelations.

Lorentz distribution

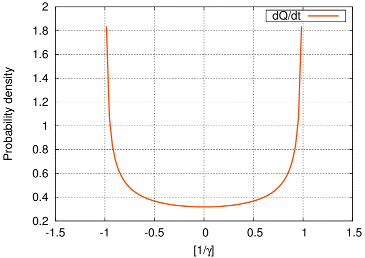
$$H(P,Q) = \log\left(1 + \frac{P^2}{\gamma^2}\right) + S(Q)$$

- Distribution function

$$f(P)dP = \frac{\gamma\, dP}{\pi[\gamma^2 + P^2]}$$

- No mean, no variance (heavy tailed distribution).
- e.o.m

$$\dot{Q} = \frac{2p}{\gamma^2 + p^2}$$



- Avoid small motion. $\mathcal{P}(-0.5 \leq \dot{Q} \leq 0.5) \approx 0.26$.
- On average move more without large motions. Cutoff in $\dot{Q}$.
- $\langle E \rangle$ twice bigger. Hepl with barriers?

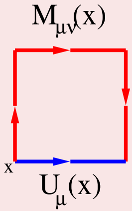
Implementation in $SU(2)$

The Hamiltonian with $U_\mu(x) \in SU(2)$ and $\pi_\mu(x) \in \mathfrak{su}(2)$

$$H(\pi, U) = \sum_{x,\mu,a} \log\left[1 + \left(\pi_\mu^a(x)/\gamma\right)^2\right] + \frac{\beta}{2}\sum_p Tr(1 - U_p)$$

Straightforward implementation

$$\begin{aligned} \dot{\pi}_\mu^a(x) &= -\frac{\beta}{4} Tr\left[T^a U_\mu(x) \sum_{\nu\neq\mu} M_{\mu\nu}(x)\right] \\ \dot{U}_\mu(x) &= \sum_a \left[\frac{2\pi_\mu^a(x)}{\gamma^2 + (\pi_\mu^a(x))^2} T^a\right] U_\mu(x) \end{aligned}$$



No modification of force computations.

Implementation in $SU(2)$

Numerical integration of e.o.m.

- Basic integrators

$$\begin{aligned} \mathcal{I}_\pi(\epsilon) &: \quad (\pi, U) \longrightarrow (\pi + \epsilon F[U], U) \\ \mathcal{I}_U(\epsilon) &: \quad (\pi, U) \longrightarrow (\pi, e^{\epsilon Z[\pi]} U) \end{aligned}$$

- $F[U]$ unchanged. $Z[\pi] \in \mathfrak{su}(2)$

$$Z^a[\pi] = \frac{2\pi_\mu^a(x)}{\gamma^2 + (\pi_\mu^a(x))^2}$$

- Any desired scheme (Leapfrog, Omelyan, 4^{th} order integrators, ...)

$$\mathcal{L}(\epsilon, N) = [\mathcal{I}_\pi(\epsilon/2)\mathcal{I}_U(\epsilon)\mathcal{I}_\pi(\epsilon/2)]^N$$

Accept-Reject based on $\Delta H(\pi, U)$ with

$$H(\pi, U) = \sum_{x,\mu,a} \log\left[1 + \left(\pi_\mu^a(x)/\gamma\right)^2\right] + \frac{\beta}{2}\sum_p Tr(1 - U_p)$$

Comparison between HMC and HMCL

Framework

- $SU(2)$ pure gauge theory.
- Two volumes 12^4 and 16^4.
- Three values of the coupling $\beta = 2.4, 2.5, 2.6$.
- Scales $a \approx 0.12, 0.08, 0.06$ fm.

Tuning

- Choose $\gamma \approx 0.8 \implies$ same P_{acc}
- Omelyan integrator with $\lambda = 0.185$ [Omelyan, Mryglod, Folk, '03].

	$V = 12^4$	$V = 16^4$
N_{steps}	20	25
ϵ	0.1	0.08
τ_L	2	2
P_{acc}	$83\% - 88\%$	$81\% - 87\%$
N_{msm}	6×10^5	10^5

Observables

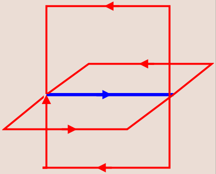
Wilson loops

- W_{1x1}: Plaquette.
- W_{2x2}
- W_{4x4}

Topological

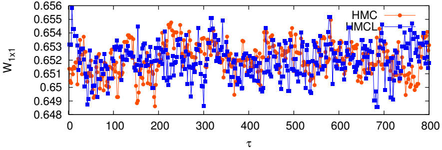
- Under-relaxed cooling

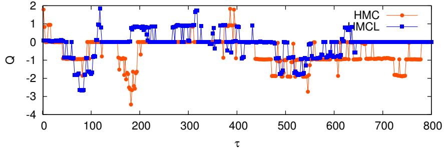
$$U_\mu(x) \to c\left(\alpha U_\mu(x) + \sum_{\nu\neq\mu} M^\dagger_{\mu\nu}(x)\right)$$

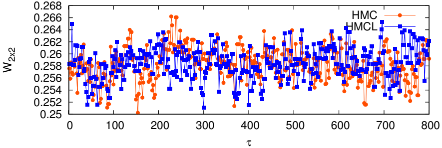


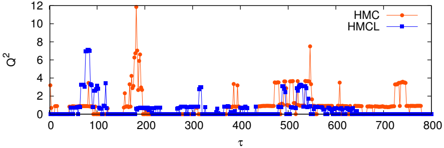
- Repeat for 60 steps ($\alpha = 2$) $\longrightarrow U_{sm}$
- Measure Q_{sm}, Q^2_{sm} (Clover definition).

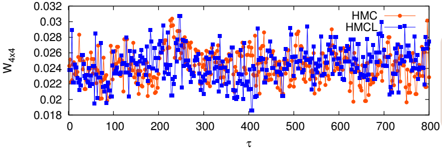
History examples: $V = 12^4, \beta = 2.5$











Key to determine which algorithm performs better is to understand auto-correlations.

Meaningful comparison

We have

- Same computational cost
- Same statistics

Compare

- τ_{int} For our observables
- Error in the quantities

Meaningful comparison

We have

- Same computational cost
- Same statistics

Compare

- τ_{int} For our observables
- Error in the quantities

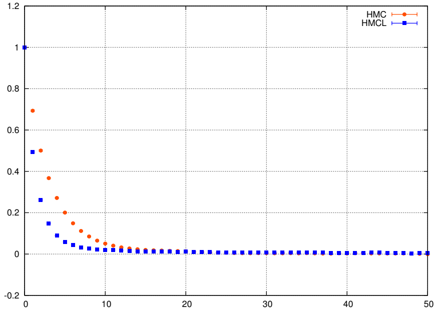


Figure: Plaquette for $V = 12^4, \beta = 2.6$

- Autocorrelation function

$$\rho_O(t) = \frac{\langle (O(0) - \overline{O})(O(t) - \overline{O}) \rangle}{\langle (O(0) - \overline{O})^2 \rangle}$$

- $\tau_{int}^{(O)}$ time to decorrelate

$$\sigma_O = \sqrt{2\tau_{int}\frac{V(O)}{N}}$$

- $\tau_{int} = 3.36(10)$
- $\tau_{int} = 2.191(53)$
- Error 23% smaller.

Meaningful comparison

We have

- Same computational cost
- Same statistics

Compare

- τ_{int} For our observables
- Error in the quantities

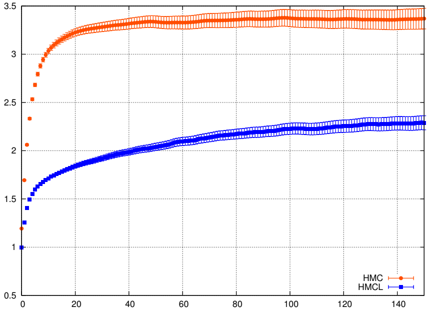


Figure: Plaquette for $V = 12^4, \beta = 2.6$

- Autocorrelation function

$$\rho_O(t) = \frac{\langle (O(0) - \overline{O})(O(t) - \overline{O}) \rangle}{\langle (O(0) - \overline{O})^2 \rangle}$$

- $\tau_{int}^{(O)}$ time to decorrelate

$$\sigma_O = \sqrt{2\tau_{int}\frac{V(O)}{N}}$$

- $\tau_{int} = 3.36(10)$
- $\tau_{int} = 2.191(53)$
- Error 23% smaller.

Meaningful comparison

We have

- Same computational cost
- Same statistics

Compare

- τ_{int} For our observables
- Error in the quantities

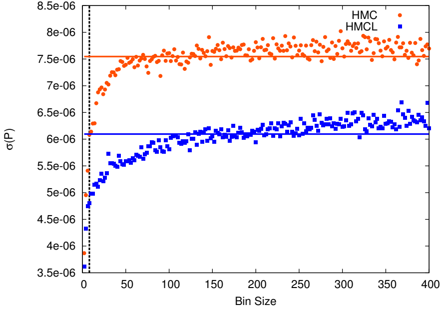


Figure: Plaquette for $V = 12^4, \beta = 2.6$

- Bin analysis with 2000 bootstrap samples.
- Use a large bin size ($Nb = 200$). With small statistics probably need different approach [Schaefer , Sommer, Virotta, '10].
- Results:
  - HMC : $\langle Tr(U_p) \rangle = 1.3401848(75)$
  - HMCL: $\langle Tr(U_p) \rangle = 1.3401944(61)$

Meaningful comparison

We have

- Same computational cost
- Same statistics

Compare

- τ_{int} For our observables
- Error in the quantities

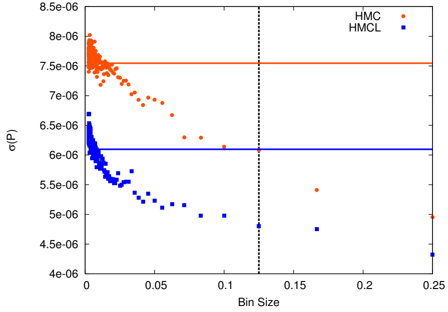


Figure: Plaquette for $V = 12^4, \beta = 2.6$

- Bin analysis with 2000 bootstrap samples.
- Use a large bin size ($Nb = 200$). With small statistics probably need different approach [Schaefer , Sommer, Virotta, '10].
- Results:
  - HMC : $\langle Tr(U_p)\rangle = 1.3401848(75)$
  - HMCL: $\langle Tr(U_p)\rangle = 1.3401944(61)$

Same game with Q^2

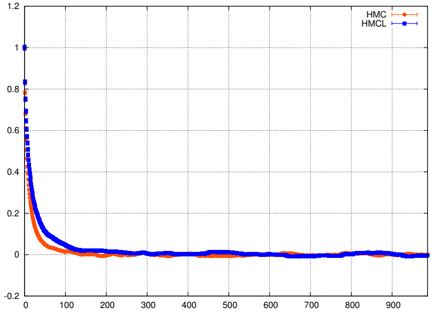


Figure: Q^2 for $V = 12^4, \beta = 2.5$

- Normalized autocorrelation function.
- HMC decorrelate faster than HMCL

Same game with Q^2

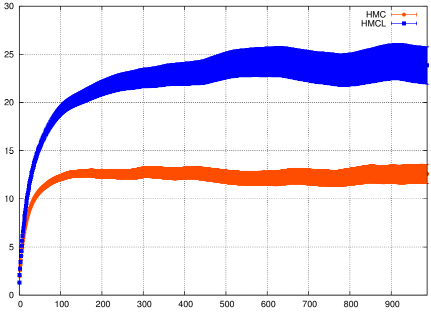


Figure: Q^2 for $V = 12^4, \beta = 2.5$

- Almost double τ_{int}
  - $\tau_{int} = 12.3(7)$
  - $\tau_{int} = 24(2)$
  - 40% bigger error for HMCL

Same game with Q^2

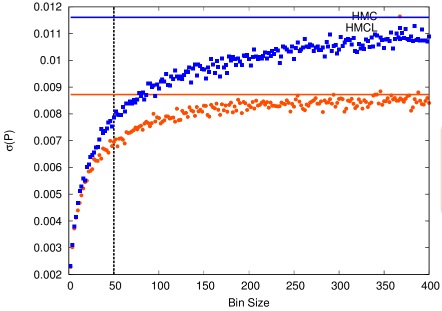


Figure: Q^2 for $V = 12^4, \beta = 2.5$

- Results
  - HMC : $\langle Q^2 \rangle = 0.8523(83)$
  - HMCL: $\langle Q^2 \rangle = 0.849(10)$

Same game with Q^2

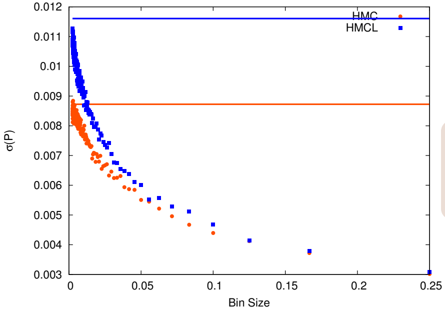


Figure: Q^2 for $V = 12^4, \beta = 2.5$

- Results
  - HMC : $\langle Q^2 \rangle = 0.8523(83)$
  - HMCL: $\langle Q^2 \rangle = 0.849(10)$

Other quantities

Very local quantities tend to decorrelate faster with HMCL, but global ("slow") quantities tend to decorrelate slower.

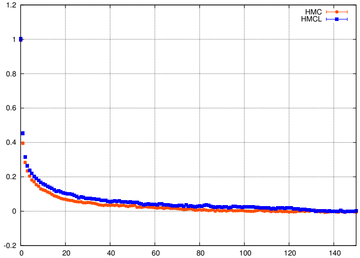


Figure: W_{4x4} for $V = 16^4, \beta = 2.5$

- $\tau_{int} = 5.50(52)$
- $\tau_{int} = 8.3(1.0)$
- HMC : $\langle Tr(W_{4\times4})\rangle = 4.8031(23) \times 10^{-2}$
- HMCL: $\langle Tr(W_{4\times4})\rangle = 4.8060(27) \times 10^{-2}$

Try with open boundary conditions

Is this a consequence of the topological modes? $\Rightarrow$ openBC [Lüscher, Schaefer, '11, '12].

Other quantities

Very local quantities tend to decorrelate faster with `HMCL`, but global ("slow") quantities tend to decorrelate slower.

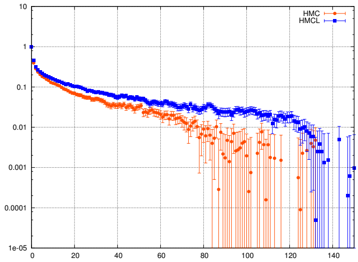


Figure: W_{4x4} for $V = 16^4, \beta = 2.5$

- $\tau_{int} = 5.50(52)$
- $\tau_{int} = 8.3(1.0)$
- HMC : $\langle Tr(W_{4\times 4})\rangle = 4.8031(23) \times 10^{-2}$
- HMCL: $\langle Tr(W_{4\times 4})\rangle = 4.8060(27) \times 10^{-2}$

Try with open boundary conditions

Is this a consequence of the topological modes? $\Rightarrow$ openBC [Lüscher, Schaefer, '11, '12].

Other quantities

Very local quantities tend to decorrelate faster with HMCL, but global ("slow") quantities tend to decorrelate slower.

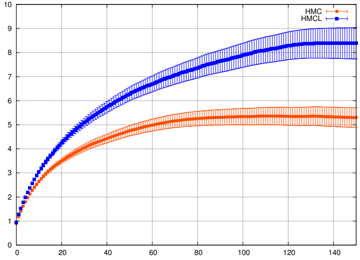


Figure: W_{4x4} for $V = 16^4, \beta = 2.5$

- $\tau_{int} = 5.50(52)$
- $\tau_{int} = 8.3(1.0)$
- HMC : $\langle Tr(W_{4\times 4})\rangle = 4.8031(23) \times 10^{-2}$
- HMCL: $\langle Tr(W_{4\times 4})\rangle = 4.8060(27) \times 10^{-2}$

Try with open boundary conditions

Is this a consequence of the topological modes? $\Rightarrow$ openBC [Lüscher, Schaefer, '11, '12].

Other quantities

Very local quantities tend to decorrelate faster with HMCL, but global ("slow") quantities tend to decorrelate slower.

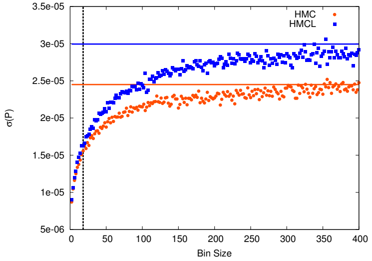


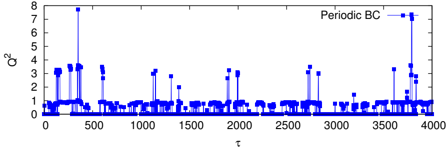
Figure: W_{4x4} for $V = 16^4, \beta = 2.5$

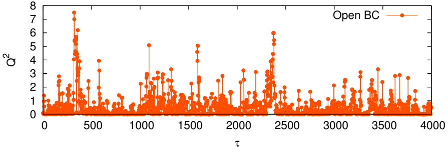
- $\tau_{int} = 5.50(52)$
- $\tau_{int} = 8.3(1.0)$
- HMC : $\langle Tr(W_{4\times4})\rangle = 4.8031(23) \times 10^{-2}$
- HMCL: $\langle Tr(W_{4\times4})\rangle = 4.8060(27) \times 10^{-2}$

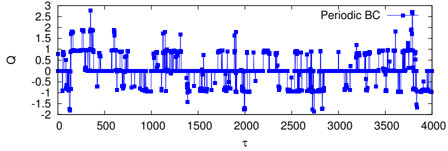
Try with open boundary conditions

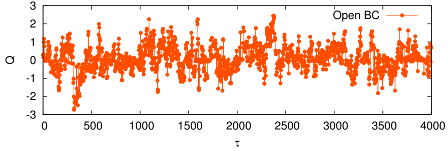
Is this a consequence of the topological modes? $\Rightarrow$ openBC [Lüscher, Schaefer, '11, '12].

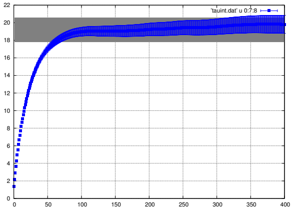
Periodic vs. Open boundary conditions $V = 12^4, \beta = 2.5$

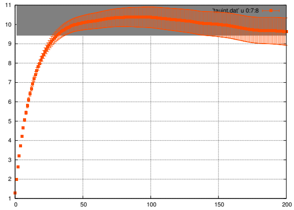




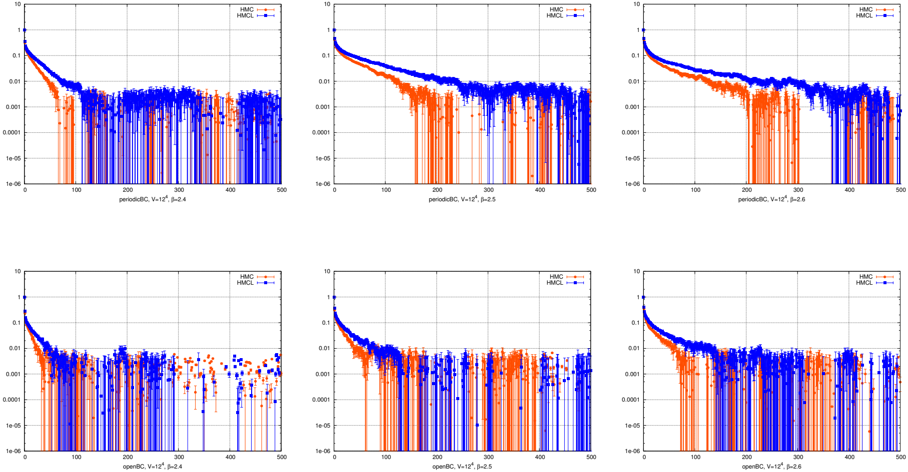








$W_{4\times4}$ for periodic vs. open boundary conditions



Some numbers with open boundary conditions

$V = 12^4$	$\beta = 2.4$	$\beta = 2.5$	$\beta = 2.6$	.alg.
$W_{1\times1}$	0.595398(10)	0.6175535(81)	0.6365164(76)	HMC
	0.5953973(97)	0.6175420(79)	0.6364912(65)	HMCL
$W_{2\times2}$	0.181980(19)	0.212554(16)	0.239240(14)	HMC
	0.181980(21)	0.212542(20)	0.239196(16)	HMCL
$W_{4\times4}$	0.0079110(77)	0.014772(11)	0.022583(12)	HMC
	0.0079129(90)	0.014777(13)	0.022552(15)	HMCL
Q^2	1.334(11)	0.4849(71)	0.1409(39)	HMC
	1.336(14)	0.4850(87)	0.1453(54)	HMCL

Conclusions

- Lorentzian kinetic term allows control on $\dot{Q} \Longrightarrow$ Choose motion in phase space.
- Comparison of two different motions in phase space by changing the kinetic term in the HMC
  - `HMCL`: All coordinates moves similarly, avoiding "small" and "large" updates.
  - `HMC`: Mix of large and small changes in the coordinates.
- Fast modes decorrelate faster for `HMCL`, slow modes decorrelate faster with `HMC`.
- Except for very local quantities, `HMC` performs better. $[-20\%, +40\%]$ in final errors for observables.
- Tests with open boundary conditions: slow modes are moved away from $\lambda_0 = 1$.
- Effect of open BC visible even in error estimation of local quantities.
- Questions:
  - Different γ's ?
  - What happens with fermions ?