Status of simulations

Jacob Finkenrath

24.09.2019

THE CYPRUS INSTITUTE
RESEARCH•TECHNOLOGY•INNOVATION

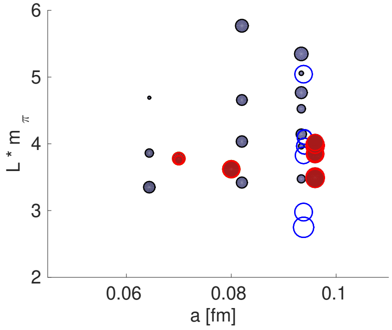
Content

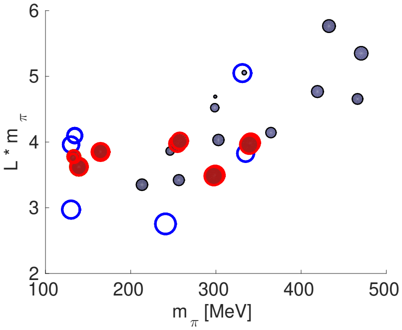
* Overview of ETMC ensembles
  - Status of cC211.06.80
  - Outlook

* Reversibility checks
  - Motivation and introduction
  - Dependence of MC observables
  - Outlook

* Optimal Integrator tuning
  - Motivation and Introduction
  - Model
  - Tuning

Overview on ETMC ensembles

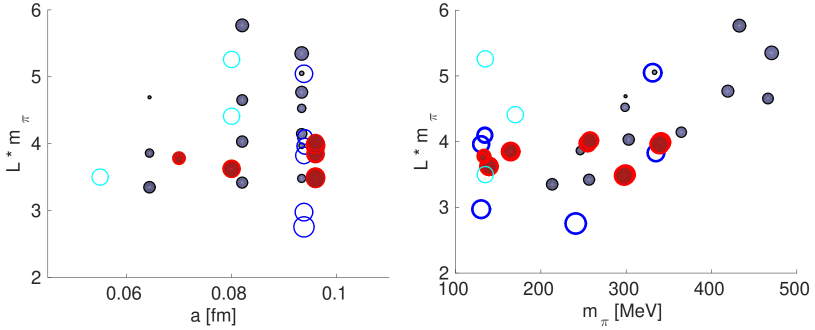




Currently running:

- L=80 at $a \sim 0.69$ fm at $m_\pi \sim 135$ MeV
- L=64 at $a \sim 0.8$ fm at $m_\pi \sim 170$ MeV
- L=24/32 at $a \sim 0.8$ fm at $m_\pi \sim 250$ MeV

Overview on ETMC ensembles



Currently running:

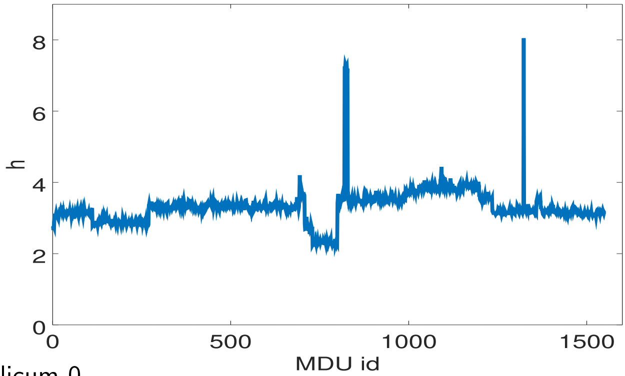
- L=80 at $a \sim 0.69$ fm at $m_\pi \sim 135$ MeV
- L=64 at $a \sim 0.8$ fm at $m_\pi \sim 170$ MeV
- L=24/32 at $a \sim 0.8$ fm at $m_\pi \sim 250$ MeV

Outlook:

- Continuation: $L = 80$ at $a \sim 0.69$ fm at $m_\pi \sim 135$ MeV (SuperMUC)
- Continuation: $L = 64$ at $a \sim 0.8$ fm at $m_\pi = 170$ MeV (Jureca-Booster)
- New Lattice : $L = 96$ at $a \sim 0.8$ fm at $m_\pi \sim 135$ MeV (Hawk: pending)
- Tuning for : $L = 96$ at $a \sim 0.55$ fm at $m_\pi \sim 135$ MeV (SuperMUC)

- Remember we also tuned for $\beta = 1.86$ which is around $a \sim 0.064$ fm

Current status of the $L = 80$ HMC simualtion



Replicum 0

- from 0 - 690 SuperMUC SB
  - at ~ 110 and ~ 230 MDU change of the integrators
  - Jul - Dec 18 on 500 Nodes
- from 691 - 1227 Marconi SKL
  - in Jan ($\sim 700 - 800$ MDU) on 250 SKL nodes
  - Feb-Apr 19 on 125 SKL nodes
- from 1227 - 1550 (now) SuperMUC NG
  - running on 125 Nodes
  - currently stable, SCRATCH not available, throughput is OKAY

Status of the $L = 80$ at the physical point

Two Replicas currently runing on SuperMUC SKX and SuperMUC HW

with statistics:
Rep 0 1550 MDU and **Rep 1** 950 MDU

Parameters:

κ	V	μ_ℓ	β	μ_σ	μ_δ	c_{sw}
0.13875285	160×80^3	0.0006	1.836	0.1065859	0.107146	1.6452

some stats:

m_{PCAC}	am_π	m_π/f_π	$L \cdot m_\pi$	$a(t_0)$
-2.7(8)e-5 (-3.3(14)e-5)(-2.1(6)e-5)	0.04734(6)	1.052(3)	3.79	$a = 0.06956$ fm

- PCAC mass seems to be undercontrol
- Autocorrelations
- Reversibility – not clarified

Status of the $L = 80$ at the physical point

Two Replicas currently runing on SuperMUC SKX and SuperMUC HW

with statistics:
Rep 0 1550 MDU and **Rep 1** 950 MDU

Parameters:

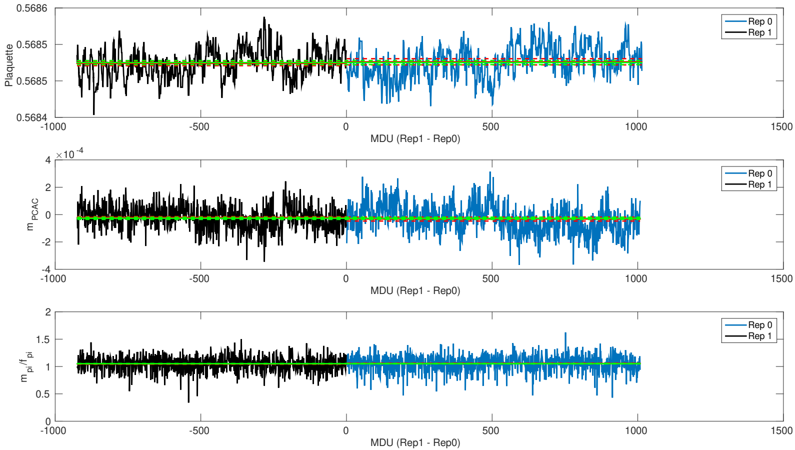
κ	V	μ_ℓ	β	μ_σ	μ_δ	c_{sw}
0.13875285	160×80^3	0.0006	1.836	0.1065859	0.107146	1.6452

some stats:

m_{PCAC}	am_π	m_π/f_π	$L \cdot m_\pi$	$a(t_0)$
-2.7(8)e-5 (-3.3(14)e-5)(-2.1(6)e-5)	0.04734(6)	1.052(3)	3.79	$a = 0.06956$ fm

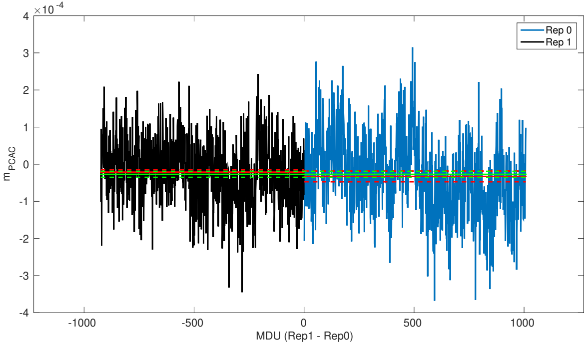
- PCAC mass seems to be undercontrol
- Autocorrelations
- Reversibility – not clarified

Progress of $V = 80^3 \times 160$ at $a \sim 0.069$ fm



Progress of $V = 80^3 \times 160$ at $a \sim 0.069$ fm

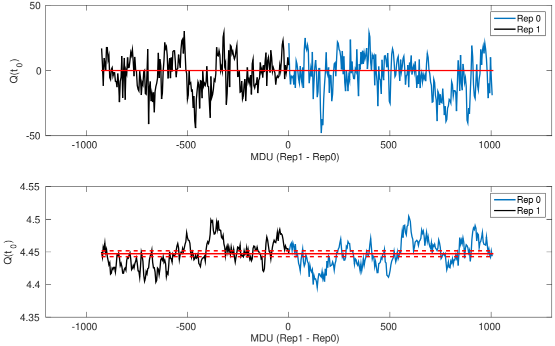
PCAC mass



- Longer oscilations
- Criterium $\frac{m_{PCAC}}{\mu} < 0.1$ is fullfilled

Topological charge and lattice spacing

Every fourth cnfg



from t_0 follows

$$\frac{t_0}{a^2} = 4.4472(45) \qquad \text{using BMW-c} \quad a = 0.0695 \text{ fm or } = 0.0673 \text{ (MILC)}$$

from f_π (simplified estimation)

$$af_0 = 0.04734(6) \qquad \text{using extrapolations with } m_\pi \quad a \sim 0.068 \text{ fm}$$

Outlook

Towards $L = 96$ Simulations

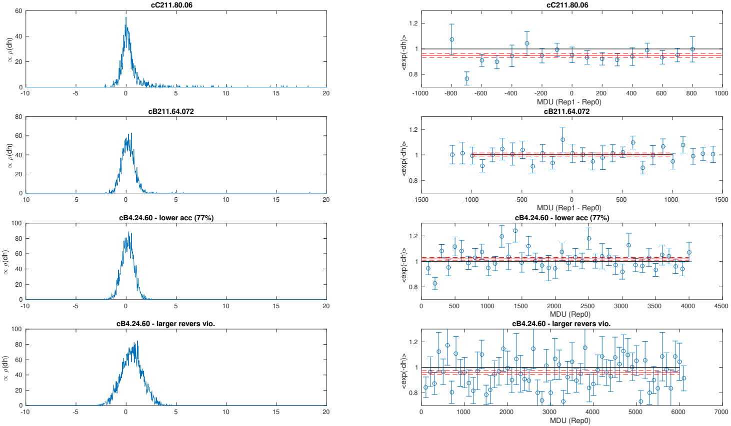
* Status of Software
  - Readiness for $L \geq 80$
  - need to coordinate the checks

* Generation of cD211-ensembles with $a < 0.06\ \mathrm{fm}$ (on SuperMUC NG)
  - Tunning effort started – $L = 32$ with $a\mu = 0.002$ is thermalizing
  - Next Steps: Tuning of $a\mu_s, a\mu_c, a\mu_\sigma, a\mu_\delta$

* Generation of cB211-ensembles with $L = 96$ (pending)
  - Few trajectories performed on 192 Skylake Nodes
  - Time per trajectory: $\sim 2.38e+04 - 3.0e+04\,\mathrm{sec}\ (6.6 - 8h)$

* Software towards exascale
  - Restructure of software to overcome strong scaling wall
  - Effort in syncronization with LyNcs, a PRACE project for Kylov Solvers towards exascale
  - we still have open positions for these efforts (at CYI/Cyprus and Bordeaux/Inria)

Reversibility



Possible reasons:

- low acceptance, large fluctuations
- Rational correction step
- Low statistics

Reversibility - Hybrid Monte Carlo algorithm:

1. Heatbath for pseudofermions η and conjugated momentas P
2. Proposal of new configuration with transition propability $T(U_{\tau=0} \rightarrow U_{\tau=1})$
3. Metropolis accept-reject step
   where $P(U_{\tau=1}) = \min\left[1, \exp\{-\delta H\}\right]$
   with $\delta H = H(U_{\tau=1}) - H(U_{\tau=0})$

Monte Carlo algrithms: **1.** and **2.+3.** will converge for $t \rightarrow \infty$

Requirement for **2.** Molecular Dynamics Integrator

- time reversable
  $\rightarrow$ $T(U_{\tau=0} \rightarrow U_{\tau=1}) = T(U_{\tau=1} \rightarrow U_{\tau=0})$ can be checked by
  $\delta\delta H = H_{U_{\tau=0\leftarrow 1\leftarrow 0}} - H_{U_{\tau=0}}$
- symplectic which keeps the phase space constant during integration
  $\rightarrow$ $\langle \exp\{-\delta H\}\rangle = 1$

Reversibility and simplecticity are related:

HMC for $SU(2)$ [Urbach, 2017]

Here, we will take a look to

$$\sigma^2(\delta\delta H) \qquad \text{and} \qquad \langle \exp\{-\delta H\}\rangle$$

for different setups

Reversibility - Hybrid Monte Carlo algorithm:

1. Heatbath for pseudofermions η and conjugated momentas P
2. Proposal of new configuration with transition propability $T(U_{\tau=0} \to U_{\tau=1})$
3. Metropolis accept-reject step
   where $P(U_{\tau=1}) = \min\left[1, \exp\{-\delta H\}\right]$
   with $\delta H = H(U_{\tau=1}) - H(U_{\tau=0})$

Monte Carlo algrithms: **1.** and **2.+3.** will converge for $t \to \infty$

Requirement for **2.** Molecular Dynamics Integrator

- time reversable
  $\to$ $T(U_{\tau=0} \to U_{\tau=1}) = T(U_{\tau=1} \to U_{\tau=0})$ can be checked by
  $\delta\delta H = H_{U_{\tau=0\leftarrow 1\leftarrow 0}} - H_{U_{\tau=0}}$
- symplectic which keeps the phase space constant during integration
  $\to$ $\langle \exp\{-\delta H\}\rangle = 1$

Reversibility and simplecticity are related:

HMC for $SU(2)$ [Urbach, 2017]

Here, we will take a look to

$$\sigma^2(\delta\delta H) \qquad \text{and} \qquad \langle \exp\{-\delta H\}\rangle$$

for different setups

Reversibility - Hybrid Monte Carlo algorithm:

1. Heatbath for pseudofermions η and conjugated momentas P
2. Proposal of new configuration with transition propability $T(U_{\tau=0} \rightarrow U_{\tau=1})$
3. Metropolis accept-reject step
   where $P(U_{\tau=1}) = \min\left[1, \exp\{-\delta H\}\right]$
   with $\delta H = H(U_{\tau=1}) - H(U_{\tau=0})$

Monte Carlo algrithms: **1.** and **2.+3.** will converge for $t \rightarrow \infty$

Requirement for **2.** Molecular Dynamics Integrator

- time reversable
  $\rightarrow$ $T(U_{\tau=0} \rightarrow U_{\tau=1}) = T(U_{\tau=1} \rightarrow U_{\tau=0})$ can be checked by
  $\delta\delta H = H_{U_{\tau=0\leftarrow 1\leftarrow 0}} - H_{U_{\tau=0}}$
- symplectic which keeps the phase space constant during integration
  $\rightarrow$ $\langle \exp\{-\delta H\} \rangle = 1$

Reversibility and simplecticity are related:

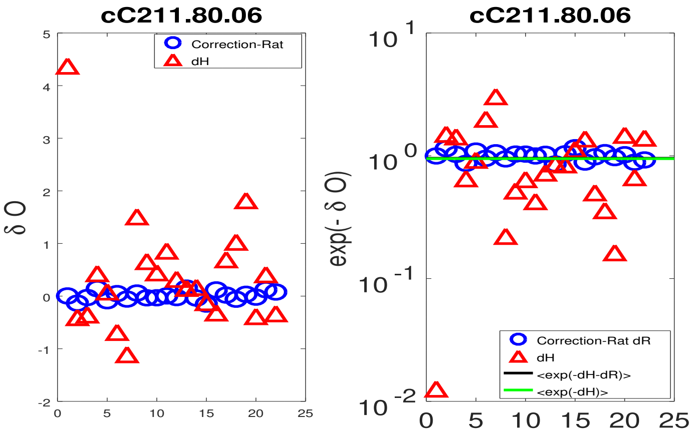
HMC for $SU(2)$ [Urbach, 2017]

Here, we will take a look to

$$\sigma^2(\delta\delta H) \qquad \text{and} \qquad \langle \exp\{-\delta H\} \rangle$$

for different setups

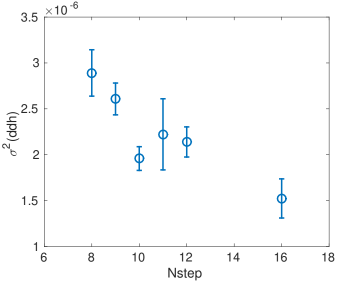
Situation for cC211.06.80

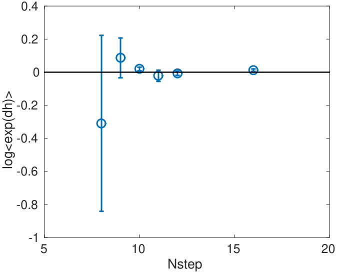


with

$$\sigma^2(\delta\delta H) = 0.00046 \qquad \text{and} \qquad \langle \exp\{-\delta H\} \rangle = 0.962(0.016); 0.952(0.022)$$

Dependence of $\sigma^2(\delta\delta H)$ and $\langle \exp\{-\delta H\}\rangle$ on N_{steps}



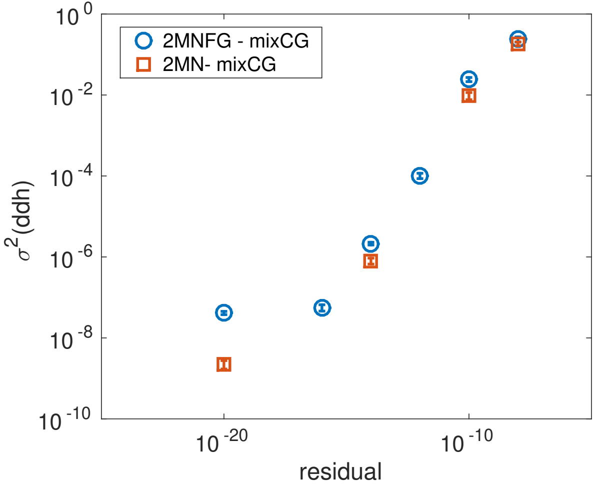


Lattice setup:

- $N_f = 4$ ensemble with $L = 24$ and $\beta = 1.778$
- $a\mu = 0.006$ and using nested integrator with 3 levels (2MNFG)
- QPhiX mixCG solver with $res^2 = 10^{-14}$.
- Acceptance between 10% to 100% - for large statistic $N_{step} = 12$

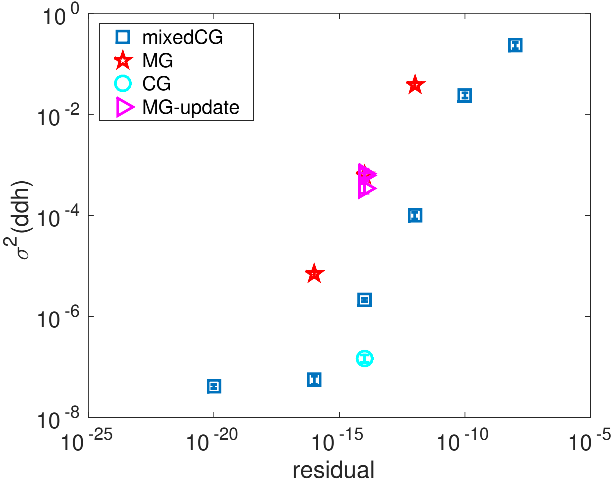
$\rightarrow$ acceptance rate has a weak influence on $\sigma^2(\delta\delta H)$ and $\langle \exp\{-\delta H\}\rangle$

Dependence of $\delta\delta H$ - Integrator comparison versus res^2



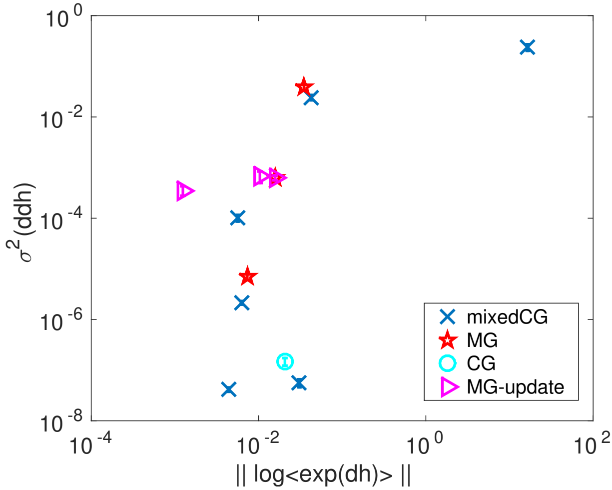
- $N_f = 4$ setup with $L = 24$ and $\beta = 1.778$
- QPhiX MixCG solver for all Hasenbusch masses
- $\sigma^2(\delta\delta H)$ shows no larger difference between 2MN and 2MNFG

Dependence of $\delta\delta H$ - Solver comparison



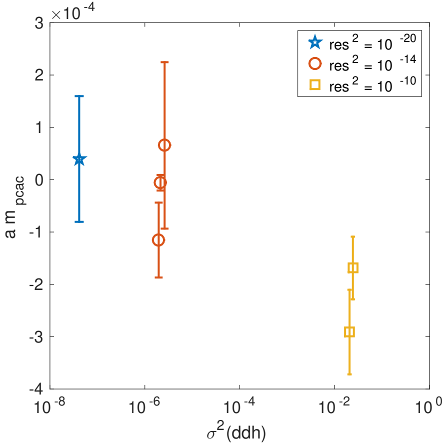
- $N_f = 4$ setup with $L = 24$ and $\beta = 1.778$
- $\sigma^2(\delta\delta H)$ of MG solver 2 magnitudes above mixCG
- $\sigma^2(\delta\delta H)$ of CG solver 1 magnitudes below mixCG

$\sigma^2(\delta\delta H)$ versus $\langle\exp\{-\delta H\}\rangle$ - PRELIMINARY no x-errors



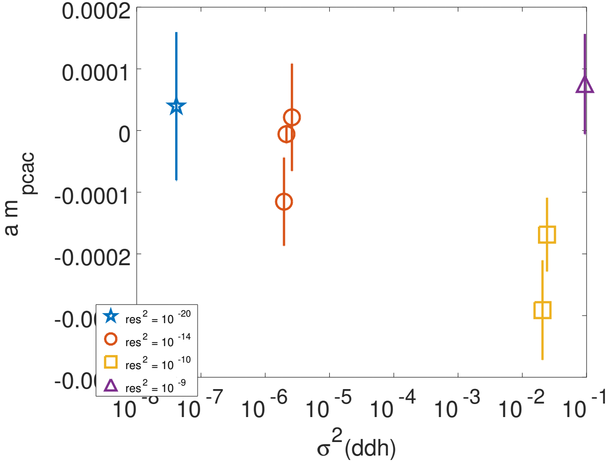
- $N_f = 4$ setup with $L = 24$ and $\beta = 1.778$
- $\langle\exp\{-\delta H\}\rangle$ starts to deviates for $\sigma^2(\delta\delta H) \gtrsim 10^{-2}$

Dependence of observables



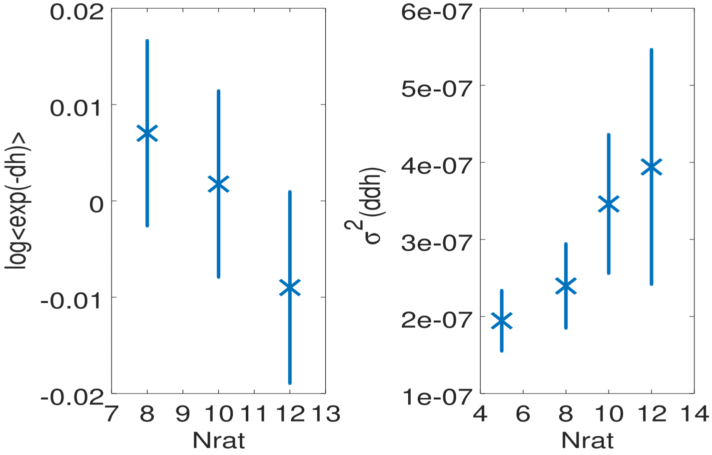
- No significant deviations found
- – only PCAC mass could show some sampling problems

Dependence of observables



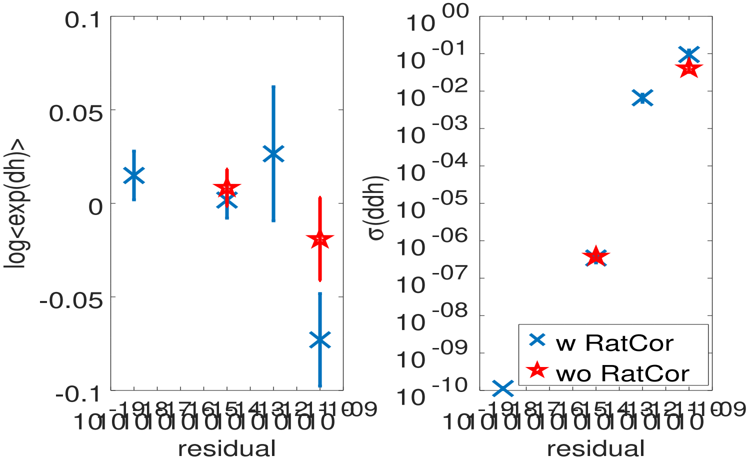
- No significant deviations found
  - only PCAC mass could show some sampling problems

Status of $N_f = 2 + 1 + 1$ runs



- $N_f = 2 + 1 + 1$ using $\beta = 1.933$ with $L = 24$ (used for tunning cD211)
- no significant tendency found

Status of $N_f = 2 + 1 + 1$ runs



- $N_f = 2 + 1 + 1$ using $\beta = 1.933$ with $L = 24$ (used for tunning cD211)
- no significant tendency found

Outlook

Note that the Reversibility of cC211.06.80 is fairly good

$$\sigma^2(\delta\delta H) < 10^{-3}$$

Still unclear is $\langle \exp\{-\delta H\} \rangle$

- possible: Low statistics or large volume/small mass effects

We will further investigation

- higher statistics for Reversibility checks for cC211.06.80
- Volume and mass scaling is missing

Possible overlap with effort by Carsten

Motivation

Order 4 integration schemes have a better volume scaling cost of HMC:

$$cost \propto V^{(1+\alpha)}$$

- order 2 schemes: $\alpha = 1/4$
- order 4 schemes: $\alpha = 1/8$

→ in general many LQCD groups using oder 4 schemes like CLS, BMW, etc. its implemented in Chroma, grid, ect...

→ symplectic integrators based on Omelyan, Mryglod, Fold (2003) where more than 10 fourth order integrators are introduced (which approach is optimal ?)

- Solite study on $SU(3)$ about tuning is missing (there are a couple but mainly only for two level and 2MN)
- Target: Tuning model
  - * minimize computational costs for large scale simulations on hybrid machines
  - * optimal integrator setup with Hasenbusch masses
  - * Comparison of different integrator approaches

inspired by a work in the Schwinger Model
[D. Shcherbakov, M. Ehrhardt, M. Günther, J. Finkenrath, F. Knechtli, M. Peardon]

Leap frog

We have to integrate a system (U,P) with an action H using Hamiltons equation

and we can defined the operators with step size h

$$e^{Ah} \ : \ U \to U' = \exp(iPh)U$$

and

$$e^{Bh} \ : \ P \to P' = P - ihF(U)$$

then the Leap-frog is given by

$$\Delta(h) = e^{h\frac{\hat{B}}{2}} e^{h\hat{A}} e^{h\frac{\hat{B}}{2}}$$

and the minimal norm scheme order 2

$$\Delta(h) = e^{h\alpha\hat{B}} e^{\frac{1}{2}h\hat{A}} e^{(1-2\alpha)h\hat{B}} e^{\frac{1}{2}h\hat{A}} e^{\alpha h\hat{B}}$$

there is a so called Shadow-Hamiltonian for which the 2NM integerator is order 4 the Shadow Hamiltonian is given by [Kennedy,Clark,Silva,Joo] 2009/2011

$$\tilde{H} = H_0 + \left(\frac{6\alpha^2 - 6\alpha + 1}{12} \{S, \{S, T\}\} + \frac{1 - 6\alpha}{25} \{T, \{S, T\}\} \right)$$

this motivates the force gradient schemes and the choice $\alpha = 1/6$

Force Gradient term

The 2MNFG scheme with force gradient term is given by

$$\Delta(h) = e^{h\frac{1}{6}\hat{B}} e^{\frac{1}{2}h\hat{A}} e^{\frac{2}{3}h\hat{B} - \frac{1}{72}h^3 C} e^{\frac{1}{2}h\hat{A}} e^{\frac{1}{6}h\hat{B}}$$

where the force gradient term

$$C = 2 \sum_{x=1,\nu=0}^{V,3} \frac{\partial S}{\delta U_\nu(x)} \frac{\partial^2 S}{\delta U_\nu(x) \delta U_\mu(x)}$$

which is given by a second derivative

- Force gradient term can be calculated via an additional inversion of $DD^\dagger$ [Lin and Mawhinney 2011]

Lets summarize the different integeration schemes which will we analyze

2MN: 2 Minimal Norm scheme (4 inversions per step)
2MNP: 2 Minimal Norm scheme (Position driven) (4 inversions per step)
2MNFG: 2 Minimal Norm scheme with approx. Force Gradient (6 inversions per step)
OMF4: 4th order OMF scheme (10 inversions per step)
OPT4FG: 4th order OPT4FG scheme (position version) (14 inversions per step)
OMF4FG: 4th order OMF4FG scheme (stage 5) (12 inversions per step)

inversions are counted as inversions of the single operator D

Model - Metropolis Algorithm

Acceptance rate for Metropolis accept–reject steps is given by (if lognormal)

$$P_{acc} = \mathrm{erfc}(\sqrt{\sigma^2/8}) \quad \text{with} \quad \sigma^2 = \mathrm{var}(\delta H)$$

Minimization

$$\min(\mathrm{Cost}) \qquad \text{with} \qquad \sigma^2 = const$$

It follows for δH for a generic integrator

$$\begin{aligned}\delta H \quad = \quad & (\nu-1)A + (\sigma-1)B + \alpha[A,[A,B]] + \beta[B,[A,B]] + \gamma_1[A,[A,[A,[A,B]]]] + \\ & \gamma_2[A,[A,[B,[A,B]]]] + \gamma_3[B,[A,[a,[A,B]]]] + \gamma_4[A,[A,[B,[A,B]]]] + \xi_1 \ldots\end{aligned}$$

where $\nu = 1$ and $\sigma = 1$ (symplectic)

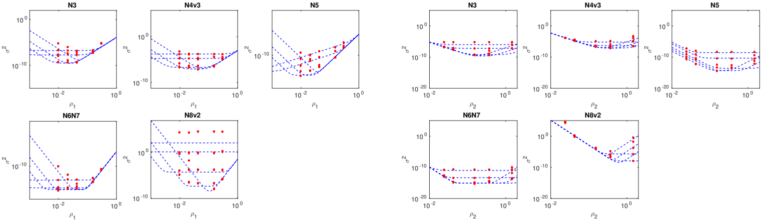
- Operator with (α,β) are of order h^3
- Operator with (γ_1,γ_2,γ_3,γ_4) are of order h^5

- choose parameters which set ($\alpha = 0$,$\beta = 0$)
  $\Rightarrow$ Integrator of fourth order
- choose parameters which set ($\alpha = 0$,$\beta = 0$,$\gamma_1 = 0$,$\gamma_2 = 0$,$\gamma_3 = 0$,$\gamma_4 = 0$)
  $\Rightarrow$ Integrator of sixth order

Goal: We will measure the variances of the different terms by choosing different parameter sets

Measuring higher order terms

Using two Hasenbusch mass terms with ρ_1 and ρ_2



with fit-ansatz

$$\sigma^2 = A_1 \left(N_{stage(0)} \frac{(\rho_1^2 - \mu^2)^{k/2}}{\mu^k} + N_{stage(1)} \frac{(\rho_2^2 - \rho_2^2)^{k/2}}{\rho_1^k} \right) + \frac{B}{\rho_1^k} + C$$

we found:

$k = 4$ for α, β

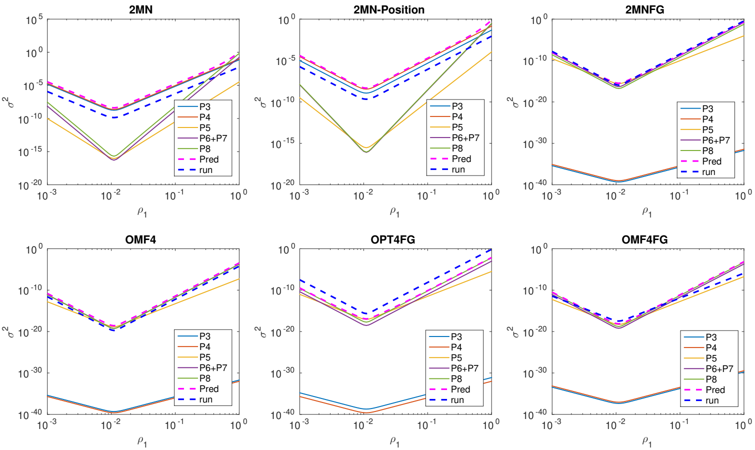
$k = 6$ for γ_1

$k = 8$ for $\gamma_2 + \gamma_3, \gamma_4$

Cross-check by directly measuring σ^2 for

- 2MN, 2MNp, 2MNFG, OMF4, OMF4FG

Model Prediction



Prediction by measuring higher order terms matches the model prediction of integrators

Note:

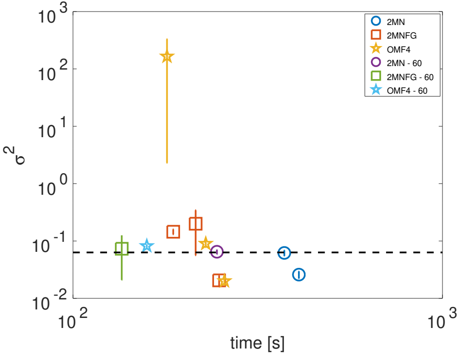
- Preliminary data: different integrators behavior fittig by using similar approach
- Parameter predicitions still preliminary

Tuning

We collect more data for different level setups:

- 2MN and 2MNFG predicition works resonable well
- OMF4 still need additional fine tuning

For test: we choose $N_f = 2$ and $L = 24$ with two masses $\mu = 0.003, 0.006$



2MN: 5 Fermions with lvl 6: 0.00205, 0.0161, 0.16, 2.24

2MNFG 5 Fermions with lvl 4: 0.0031, 0.028, 0.16, 1.62

OMF4: 3 Fermions with lvl 3: 0.011, 0.091

Conclusion

Simulations

- for the $L = 80$ we have around 2000 MDUs ($\sim 90 days$ to reach 3000 MDUs)
- tuning for cD211-lattices ongoing
- Readiness for $L = 96$?

Reversibility

- deviation of $\sigma^2(\delta\delta H)$ to $\langle \exp\{-\delta H\} \rangle$ for cC211.80 unclear
- still investigation in $n_f = 2 + 1 + 1$ ongoing

Tuning of integrator

- or OMF4 still investgation necessary
- other OMF integrator still missing

approximative Force Gradient scheme

Lin and Mawhinney came to a simple but genius–like idea
the second derivative term can be approximated

then the approximtive force gradient term is given by

0. Save U
1. Update the gauge field temporarily according to $U_0 = exp(-h^2/24 F_j T^j)U$.
2. Apply $e^{Bh} \,:\, P \rightarrow P' = P - ihF(U_0)$ and
0. Restore U and discard U_0.

and is straightforward to implement.

Advantages:

- needs only an allocation of mem for an additional gaugefield
- reduces the inversion of D from 3 to 2 during the force gradient calculation

but it does not solve that all parts of the actions are mixing

however maybe this is not needed if the most expensive terms of the Hamiltonian are dominating the integrations

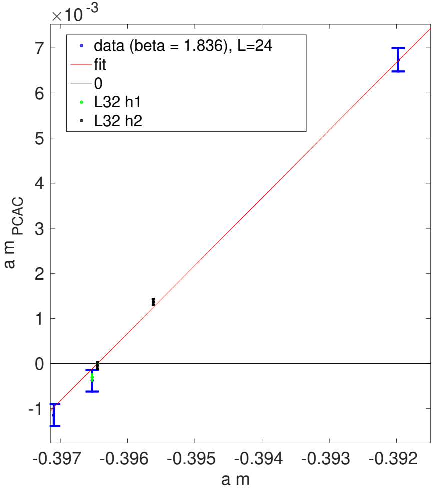
Light Quark Sector

Tuning of κ_{crit}:

$$\frac{Z_A m_{PCAC}(\kappa)}{\mu} \lesssim 0.1$$

such that $a\Lambda_{QCD}|\epsilon/\mu| \lesssim a^2\Lambda_{QCD}^2$ [P. Boucaud et al. 2008]

- Estimate κ_{crit} using $L = 24$ and $\mu = 0.0035$
- Retune μ_{sigma} and μ_δ using $L = 32$ and $\mu = 0.002$
- Estimate κ_{crit} using $L = 32$ and $\mu = 0.002$
- Retune μ_{sigma} and μ_δ using $L = 32$ and $\mu = 0.002$
- Estimate κ_{crit} using $L = 40$ and $\mu = 0.00125$



Light Quark Sector

Tuning of κ_{crit}:

$$\frac{Z_A m_{PCAC}(\kappa)}{\mu} \lesssim 0.1$$

such that $a\Lambda_{QCD}|\epsilon/\mu| \lesssim a^2\Lambda^2_{QCD}$ [P. Boucaud et al. 2008]

Summary:

- Using different combinations of fits the slope is around
- $slope \sim 1.7$