Tuning the Hybrid Monte Carlo algorithm using molecular dynamics forces' variances

Andrea Bussone

31 May 2018

Collaborators: M. Della Morte, V. Drach, C. Pica

Motivations

HMC algorithms have a large number of parameters

- Integrators: Omelyan α, ...
- Mass preconditioning: mass-preconditioning μ, ...
- Multi-time scale integration: $n, m, k, \ldots$

Optimization in QCD

[Urbach et al. Comput. Phys. Commun. 174 (2006) 87]

In BSM strongly interacting theories, changing fermion representation, forces' hierarchy may be different

Set-up

We want to solve the following problem

$$\min_{\underline{x}=(n,m,k,\mu,m_0)\in\mathbb{N}^3\times\mathbb{R}} \text{Cost}(\underline{x})$$
$$\text{subject to: } k = 10, P_{\rm acc}(\underline{x}) \gtrsim 70-80\%$$

Model under investigation:

- $\mathbf{SU}(2)$ gauge group
- unimproved Wilson fermion degenerate doublet in the fund. repr.
- $\beta = 2.2,\ m_{\rm cr} \simeq 0.77(1)$
- 3 time scales
- one mass preconditioning

Setting the stage

Why Shadow Hamiltonians?

1D Harmonic oscillator 1/2

[Yoshida Volume 152 of IAU Symposium, page 407, 1992]

Hamiltonian system to solve

$$H = \frac{p^2}{2} + \frac{q^2}{2} \quad \longrightarrow exp\left(\tau \frac{\mathrm{d}}{\mathrm{d}t}\right) = exp\left(\tau\left(-\frac{\partial H}{\partial q}\frac{\partial}{\partial p} + \frac{\partial H}{\partial p}\frac{\partial}{\partial q}\right)\right)$$

Euler scheme

$$\begin{pmatrix} p' \\ q' \end{pmatrix} = \begin{pmatrix} 1 & -\tau \\ \tau & 1 \end{pmatrix} \begin{pmatrix} p(0) \\ q(0) \end{pmatrix}$$

The energy increases at each step

$$H' = H\left(1 + \frac{\tau^2}{2}\right)$$

Symplectic Euler scheme

$$\begin{pmatrix} p' \\ q' \end{pmatrix} = \begin{pmatrix} 1-\tau^2 & -\tau \\ \tau & 1 \end{pmatrix} \begin{pmatrix} p(0) \\ q(0) \end{pmatrix}$$

Constant of motion

$$\widetilde{H} = H + \frac{\tau}{2} pq$$

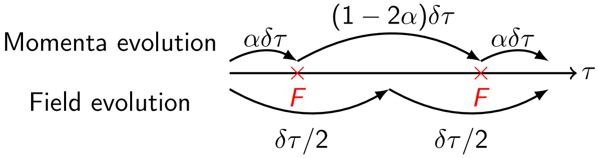
Why Shadow Hamiltonians?

1D Harmonic oscillator 2/2

We may choose[1] the Omelyan integrator

[Omelyan et al. Comput. Phys. Commun. 151 (2003) 272–314]

$$\exp\left(\alpha\delta\tau S\right)\exp\left(\frac{\delta\tau}{2}T\right)\exp\left((1-2\alpha)\delta\tau S\right)\exp\left(\frac{\delta\tau}{2}T\right)\exp\left(\alpha\delta\tau S\right)$$



Shadow Hamiltonian

$$\widetilde{H} = H + \tau^2\left(\frac{6\alpha^2-6\alpha+1}{12}\{S,\{S,T\}\} + \frac{1-6\alpha}{24}\{T,\{S,T\}\}\right) + O(\tau^4)$$

For $\alpha = 1/6$

$$\widetilde{H} = H + \frac{\tau^2}{72}F^2 + O(\tau^4)$$

[1] This is actually the only choice explored in this talk.

Mass preconditioning

[Hasenbusch Phys. Lett. B 519 (2001) 177]

$$D_W = D + m_0,\ Q = \gamma_5 D_W$$

We split the fermion determinant as

$$\det D_W = \det(D_W + \mu) \times \det(D_W + \mu)^{-1} D_W$$

We introduce two kind of pseudofermions for the degenerate case

$$\begin{aligned} H =& \pi^2/2 + S_G(U) \\ &+ \phi_1^\dagger \left[Q^{-1}(D_W+\mu)^\dagger (D_W+\mu) Q^{-1} \right] \phi_1 \\ &+ \phi_2^\dagger \left[\gamma_5 (D_W + \mu) \right]^{-2} \phi_2 \end{aligned}$$

In general we can put several splittings

$$H = \pi^2/2 + S_G(U) + \sum_i S_i(U, \phi_i^\dagger, \phi_i)$$

Shadow Hamiltonians

For any symplectic integrator it exists an exactly conserved Hamiltonian

[Kennedy et al. Phys. Rev. D 87 (2013) no.3]

Shadow Hamiltonian is found through BCH asymptotic expansion in $\delta\tau$

$$\widetilde{H} = H + \tau^2 \sum_i \left[\frac{6\alpha^2 - 6\alpha + 1}{3} \left(\frac{\{S_i, \{S_i, T\}\}}{\prod_{k=1}^{i} (2n_k)^2} \right) + \frac{1 - 6\alpha}{6} \left(\frac{\{T, \{S_i, T\}\} + \{\{S_i, \sum_{j>i} S_j\}, T\}}{\prod_{k=1}^{i} (2n_k)^2} \right) \right] + \mathrm{O}(\delta\tau^4)$$

By setting $\alpha = 1/6$ in each level, $T = \pi^2/2$ and only one mass-preconditioning

$$\delta H \simeq \frac{\delta\tau^2}{72} \left(|\mathcal{F}_1|^2 + \frac{|\mathcal{F}_2|^2}{4m^2} + \frac{|\mathcal{F}_G|^2}{16m^2k^2} \right), \quad |\mathcal{F}_i|^2 = \sum_{x,a,i} T_{R,i} (F_i^{a\mu}(x))^2$$

Acceptance and cost

$$\mathrm{Cost}(n,m,k,\mu) = \frac{\#\mathrm{MVM}}{P_{\mathrm{acc}}}(n,m,k,\mu)$$

[Gupta et al. Phys. Lett. B 242 (1990) 437] [Clark et al. PoS LATTICE 2008 (2008) 041]

$$P_{\mathrm{acc}}(\delta H) = \mathrm{erfc}\left(\sqrt{\mathrm{Var}\left(\delta H\right)/4}\right)$$

(Neglecting covariances)

$$\mathrm{Var}\left(\delta H\right) \simeq \frac{\delta\tau^4}{(72)^2}\left[\mathrm{Var}\left(|\mathcal{F}_1|^2\right)(\mu) + \frac{\mathrm{Var}\left(|\mathcal{F}_2|^2\right)(\mu)}{(4m^2)^2} + \frac{\mathrm{Var}\left(|\mathcal{F}_3|^2\right)(\mu)}{(16m^2k^2)^2}\right]$$

$$\#\mathrm{MVM} = (2n+1)\#\mathrm{MVM}_1(\mu) + 2n(2m+1)\#\mathrm{MVM}_2(\mu)$$

Idea: The variances and MVMs depend upon μ at fixed m_0, g_0 and V while the n, m, k dependence is known

Tests and numerical results

Hamiltonian scaling behaviour

2^{nd} order integrator: $|\Delta H| \propto \delta\tau^2$, from shadow hamiltonian: $|\delta H| \propto \delta\tau^2$

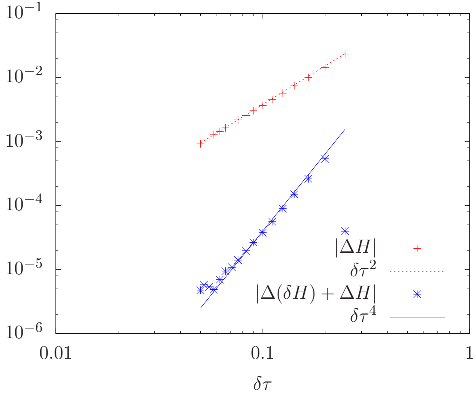
$$|\Delta(\delta H) + \Delta H| \propto \delta\tau^4$$

$V = 8^4$, $\beta = 2.2$, $m_0 = -.72$

Levels

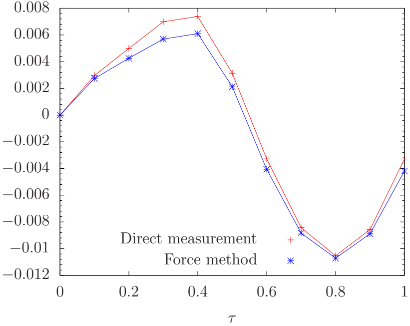
- 0: Hasenbusch $n = 4, 5, \ldots, 20$
- 1: HMC $m = 10$
- 2: Gauge $k = 10$

One trajectory from a thermalized configuration fixing conjugate momenta

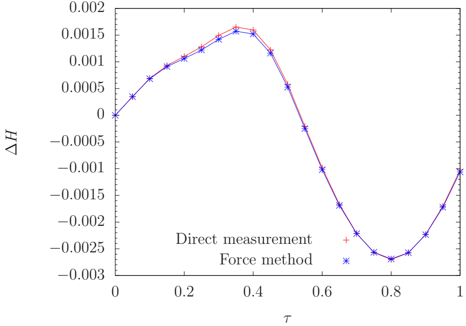


ΔH Approximation

$$\Delta H = -\frac{\delta\tau^2}{72}\Delta\left(|\mathcal{F}_1|^2 + \frac{|\mathcal{F}_2|^2}{4m^2} + \frac{|\mathcal{F}_3|^2}{16m^2k^2}\right)$$



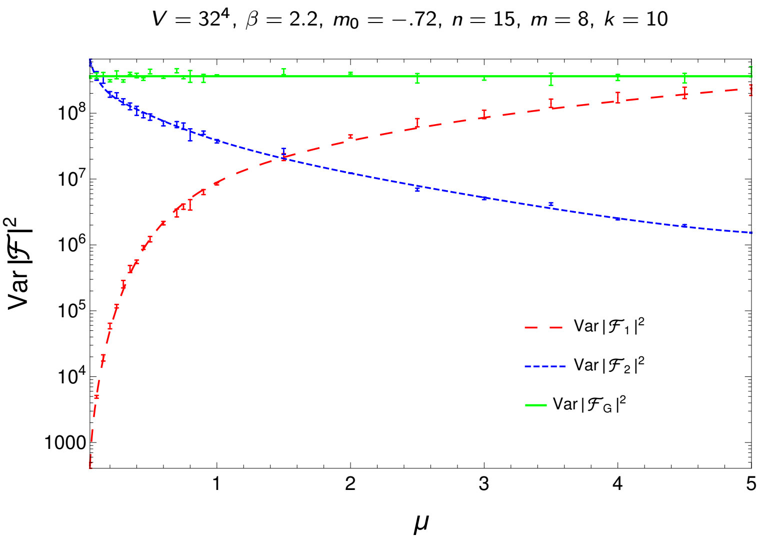
$n = 10$



$n = 20$

The measurement of Poisson brackets (forces) works well

Variances' hierarchy



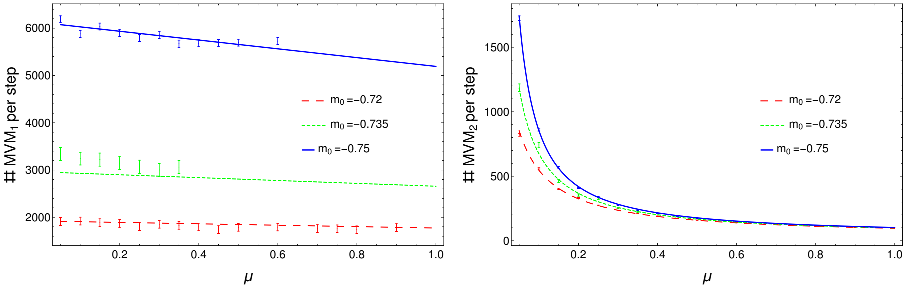
Strong dependence

Flat dependence

Matrix Vector Multiplications

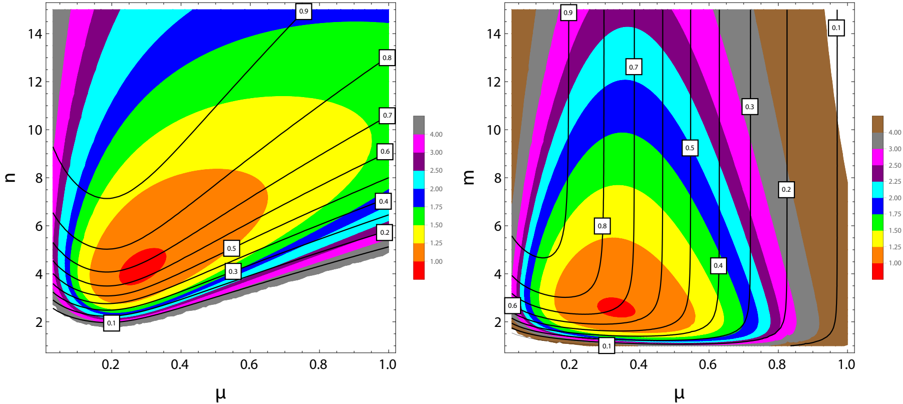
$$\#\mathrm{MVM}_1(\mu, m_0) = \frac{a + b(m_0 - m_\mathrm{c}) + c\mu}{(m_0 - m_\mathrm{c})^2} ,$$

$$\#\mathrm{MVM}_2(\mu, m_0) = \frac{a + b\mu + c\mu^2}{(\mu + m_0 - m_\mathrm{c})^2} .$$



different masses with different time scales...

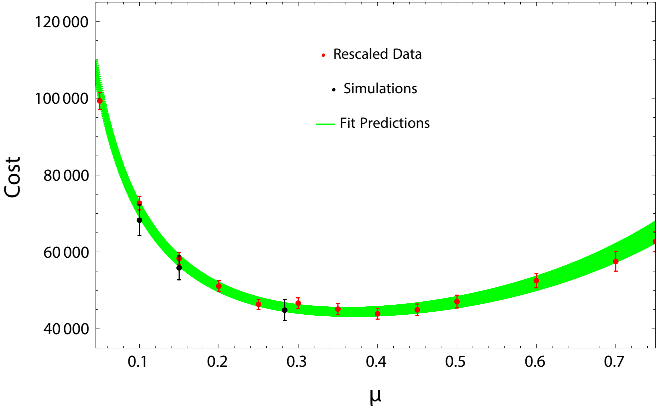
Minimizing the cost



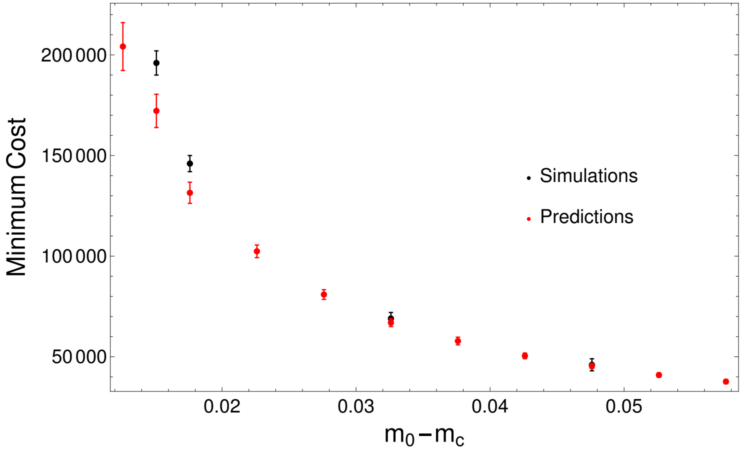
Normalized to the minimum that is found around
$(n, m, \mu) \simeq (5, 3, 0.3)$, $k = 10$ fixed
for $m_0 = -.72$

Cost predictions 1/2

There is no need for fitting the dependences



Cost predictions 2/2



Conclusions

The Shadow Hamiltonian is a clean theoretical tool to

- Optimize the parameters of the integrator
- Optimize the parameters in the determinant splitting

Recipe

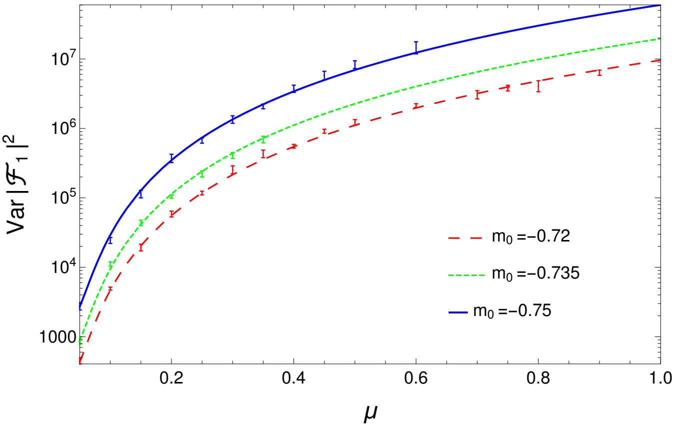
- Start with a *reasonable* choice for (n, m, k)
- Measure the *forces* in each level: $|\mathcal{F}_i|^2(\mu)$ and $\mathrm{Var}(|\mathcal{F}_i|^2)(\mu)$
- Measure the number of MVM in each level
- By fitting the dependence in μ we predict dependence of the *acceptance* and the *cost* upon (n, m, k, μ, m_0) within 10%

Outlook

- Apply this optimizing method to other strongly interacting models
- Can be generalized to more complicated cases

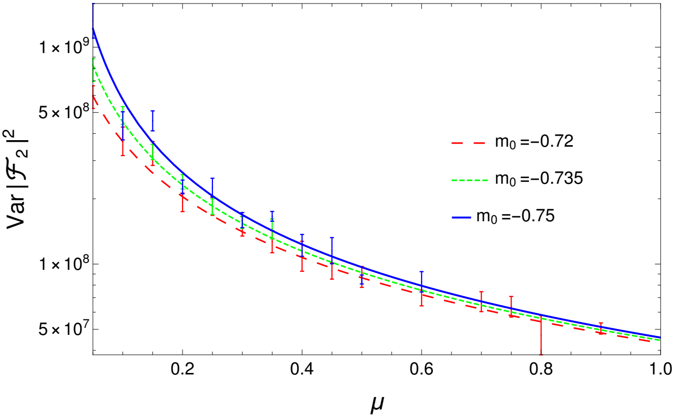
Back-up slides

Variances



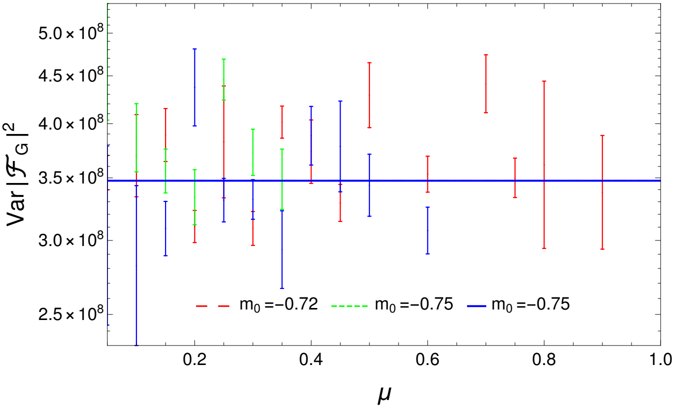
$$\mathrm{Var}(|\mathcal{F}_1|^2)(\mu, m_0) = \frac{a\mu + b\mu^2 + c\mu^3}{(m_0 - m_c)^2}$$

Variances



$$\mathrm{Var}(|\mathcal{F}_2|^2)(\mu, m_0) = \frac{a + b\mu + c\mu^2}{(\mu + m_0 - m_\mathrm{c})^2}$$

Variances



$$\mathrm{Var}(|\mathcal{F}_{\mathrm{G}}|^2)(\mu, m_0) \quad = \quad a$$