Decimation map in 2D for accelerating HMC

Nobuyuki Matsumoto

RIKEN BNL Research Center
(→ BU from Sep)

LATTICE 2023
Fermilab, 07.31.2023

Based on collaboration with
P. Boyle, R. Brower, T. Izubuchi, L. Jin, C. Jung, A. Tomiya

in preparation

RIKEN's Programs for Junior Scientists
(SPDR)

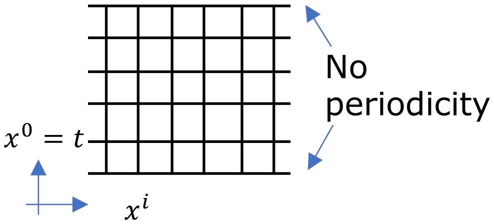
Introduction (1/2)

Goal Overcome critical slowing down & topological freezing and perform efficient simulations on fine lattices.

- Partial solution may be *open boundary condition*

  Lüscher-Schaefer 11

  But many statistical techniques assume translational invariance thus we want to keep the periodic boundary if possible.



Introduction (2/2)

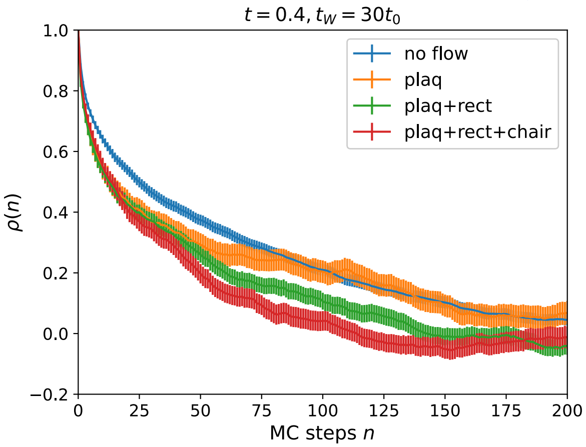
Short timeline

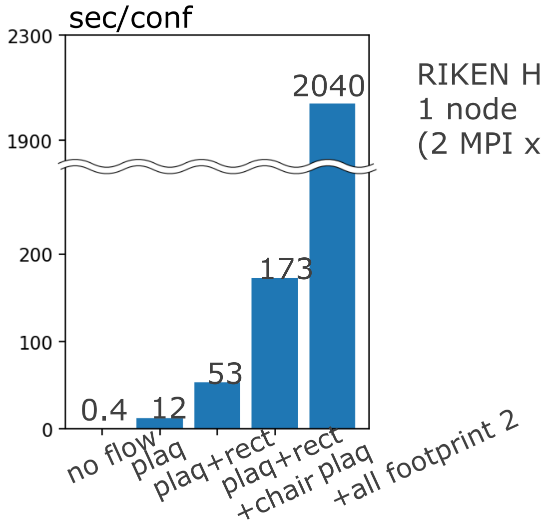
- Proposal of *trivializing map* **Lüscher 09** (Name borrowed from **Nicolai 80**)
- LO approximation w/ CP^{N-1} **Engel-Schaefer 11**

  "The reduction in the forces, ..., is compensated by the computational overhead"

- Wilson flow and its generalizations [4D SU(3) pure gauge]

  Jin LATTICE 2021 poster **Boyle-Izubuchi-Jin-Jung-NM-Lehner-Tomiya LATTICE2022 [2212.11387]**





RIKEN HOKUSAI
1 node
(2 MPI x 40 OpenMP)

Tunneling rate can increase, but still overwhelmed by overhead.

➡ Step back and reconsider strategy using 2D U(1)

- Significant developments in normalizing flow/generative models

  Albergo-Kanwar-Shanahan 19
  Bacchio-Kessel-Schaefer-Vaitl 22

  See *Plenary of G. Kanwar this morning (Mon)*

This work (1/1)

NM+ in prep

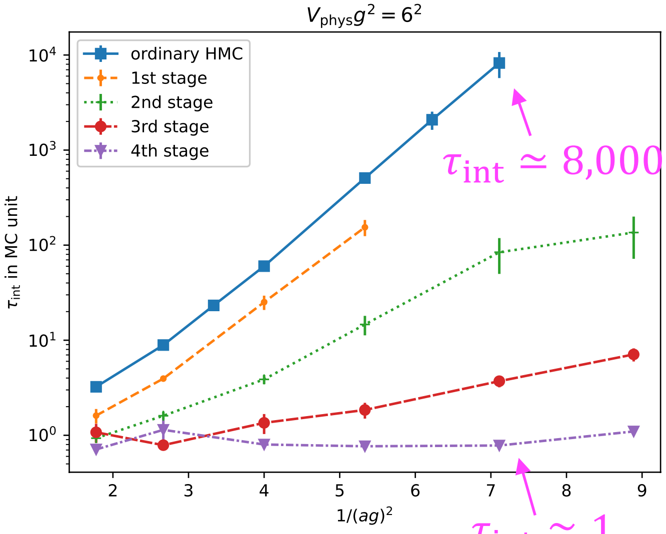
- Develop a variant of trivializing map: *"decimation map"*

Trivialization decomposed into several stages.
Each stage corresponds to integrating out local DOF

- Apply to 2D U(1) pure gauge with the guided MC (a generalization of HMC)

Horowitz 91

$\tau_{\mathrm{int}}(Q)$ in MC unit

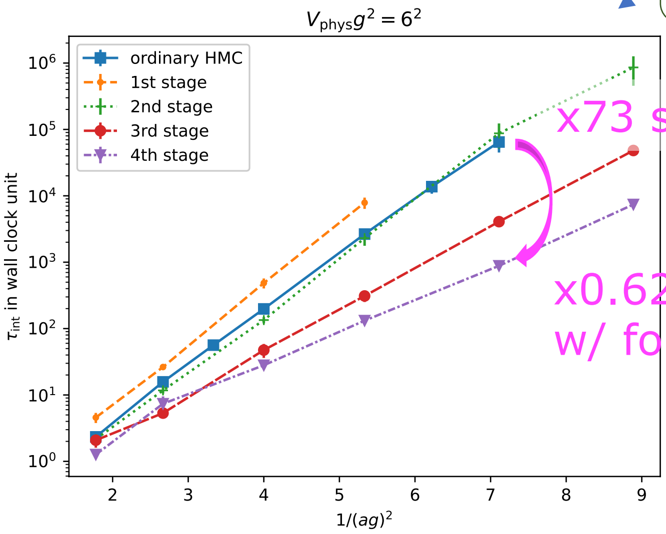


$\tau_{\mathrm{int}} \simeq 8{,}000$

$\tau_{\mathrm{int}} \simeq 1$

$\tau_{\mathrm{int}}(Q)$ in wall-clock time

Updated from ECT* ML workshop 06/28/2023



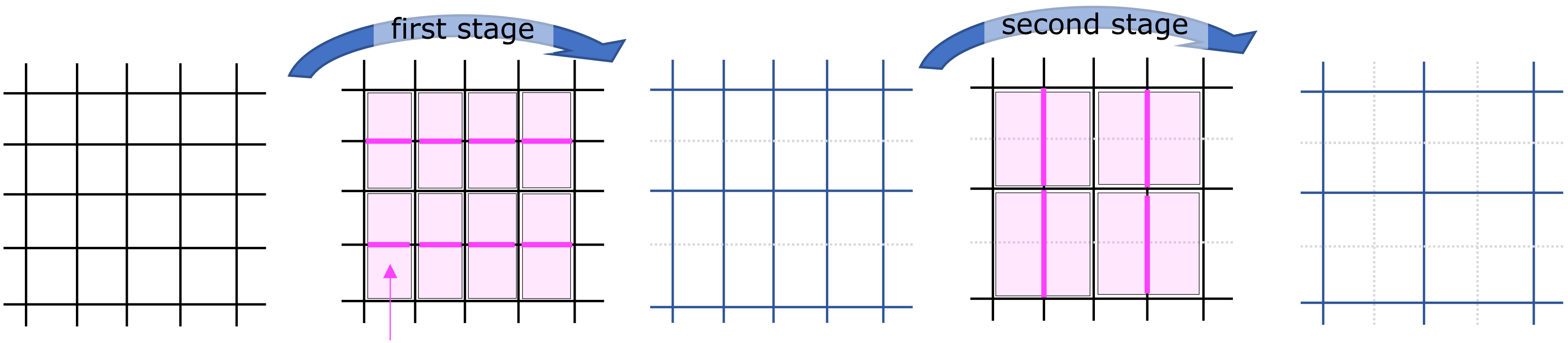
x73 speed up

x0.62 in exponent
w/ four stages

- Algorithm is exact, scalings known and controlled.

Decimation map (1/2)

- Trivialization w/ several stages



In each block, trivializing map as a gradient flow:

$$\dot{U}_t U_t^{-1} = -\partial K_t \quad \left(K_t\text{: flow kernel} \right)$$

and repeat…

cf. Lüscher 09

- K_t can be ontained from solving the linear equation: $\mathcal{L}_t K_t = -S^{(\text{loc})}$. $\left(\mathcal{L}_t \equiv -\partial^2 - t\,\partial S^{(\text{loc})} \cdot \partial \right)$

  We directly solve this with CGNE; only once as pre-calculation.

- Dividing the system into $\mathcal{T}$ and $\mathcal{R}$,

  to be trivialized remainder

local trivialization
(change of variables)

$$\begin{tikzcd} Z = \int (d\mathcal{T})(d\mathcal{R})\, e^{-S(\mathcal{T},\mathcal{R})} \arrow[r, phantom, "="] \arrow[d, "\text{integrate}"'] & \int (d\mathcal{T})(d\mathcal{R})\, e^{-S_{\text{eff}}(\cancel{\mathcal{T}},\mathcal{R})} \arrow[d, "\text{integrate}"] \\ \int (d\mathcal{R})\, e^{-S_{\text{eff}}(\mathcal{R})} \arrow[r, phantom, "\overset{!}{=}"] & \int (d\mathcal{R})\, e^{-S_{\text{eff}}(\mathcal{R})} \end{tikzcd}$$

must be the same!

Decimation map (2/2)

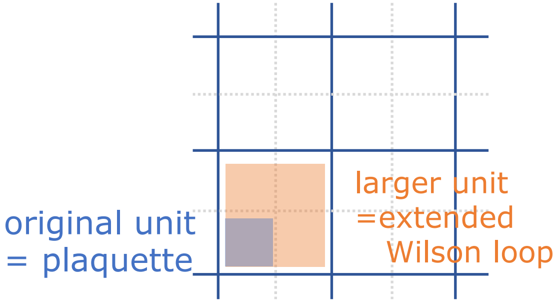
NM+ in prep

- This map leaves the physics in large units exactly unchanged.

  ∴ can be regarded as a coarse-graining map

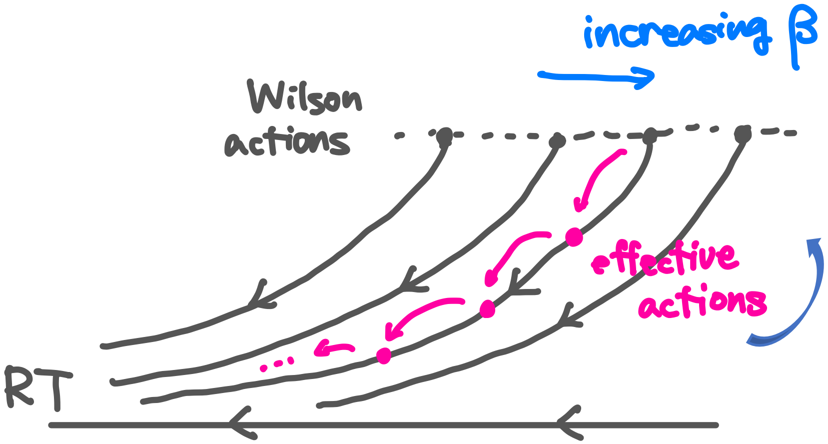
cf. Wilson-Kogut 74, Kadanoff 75, Migdal 76

see also *U. Wenger Mon, R. Abbott Mon*



original unit = plaquette

larger unit =extended Wilson loop



Integrate in the decimated links
when calculating observables.

Updates in coarse-grained lattice.

- Acceleration expected because of
  - Increase of the trivialized region
  - Coarser action for the large-unit variables

2D $U(1)$ (1/1)

Wilson action

$$S(U) \equiv -\beta \sum_x \cos\kappa_x$$

$\kappa_x \equiv \frac{1}{i}\log\left(U_{x,0}U_{x+0,1}U^{\dagger}_{x+1,0}U^{\dagger}_{x,1}\right)$: plaquette angle

$-\pi \leq \kappa_x < \pi$

$\beta = \frac{1}{(ag)^2}$

Solvable system, exact formulas from Fourier expansion.

see also *D. Hoying Tue*

Characteristic features

- Topological charge: $Q = \frac{-1}{2\pi}\sum_x \kappa_x \in \mathbb{Z}$ Simplest lattice gauge field theory with topology
- correlation length zero (ultralocal)

  deal with topological freezing rather than entire critical slowing down.
- topological susceptibility: $\chi_Q \sim \frac{g^2}{(2\pi)^2}$. $\therefore$ typical instanton size $\sim (2\pi)^2/g^2$.

Update algorithm (1/1)

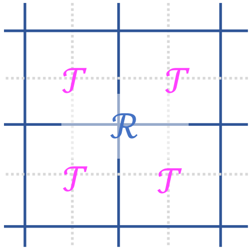
Guided-Hamiltonian Monte Carlo (a variant of HMC):
Horowitz 91 **Duane et al. 87**

- Replace the action in H by an approximate effective action (calculated in CG).
- Detailed balance still holds $\because$ Liouville theorem and reversibility

Pros

- Simplifies the calculation of force
- Transparent separation of UV/IR

$\mathcal{T}$ $\mathcal{R}$



Cons

- Acceptance rate needs to be controlled additionally by the flow step size ϵ: $\delta H \propto \epsilon^n$

($n = 2$ below)

- no gauge fixing
- periodic boundary
- Flow equation solved with the midpoint integrator ($\therefore$ net discretization error = $O(\epsilon^2)$ as mentioned)
- Bijective ensured by keeping ϵ in a bound. **cf. Lüscher 09**

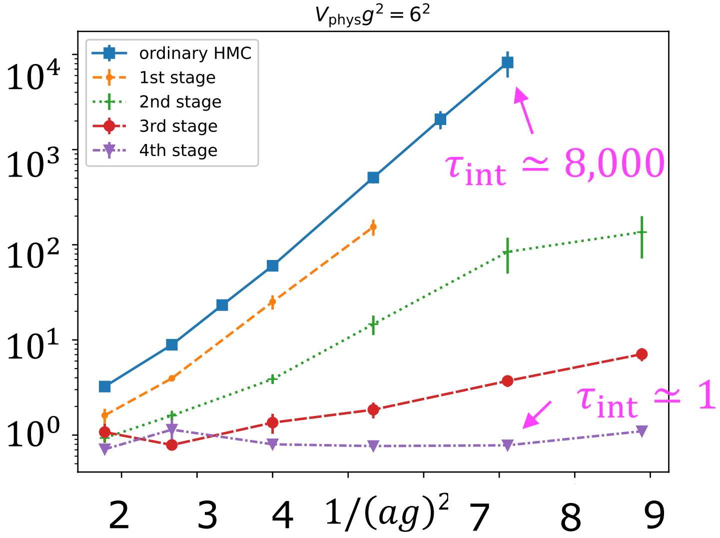
Cost scaling towards the continuum limit (1/1)

NM+ in prep

Physical volume fixed to $6\times 6/g^2$, ϵ scaled to keep acceptance rate ~0.9

~typical instanton size

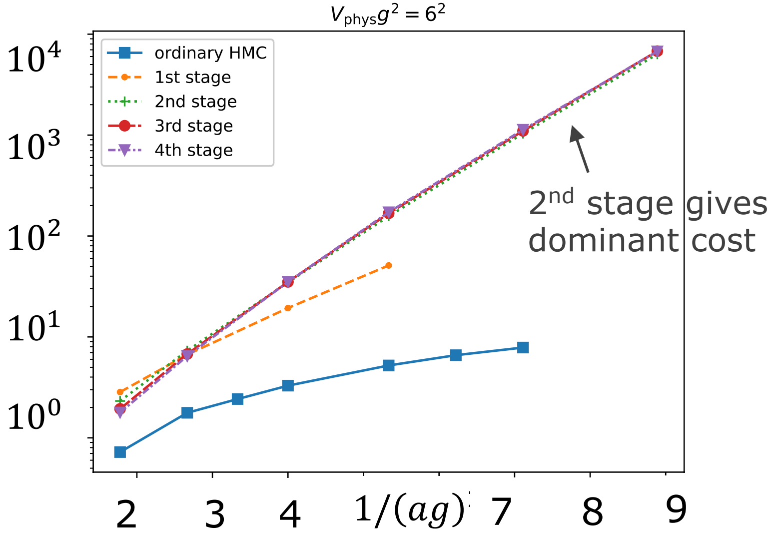
$\tau_{\rm int}(Q)$ in MC unit



$\tau_{\rm int} \simeq 8{,}000$

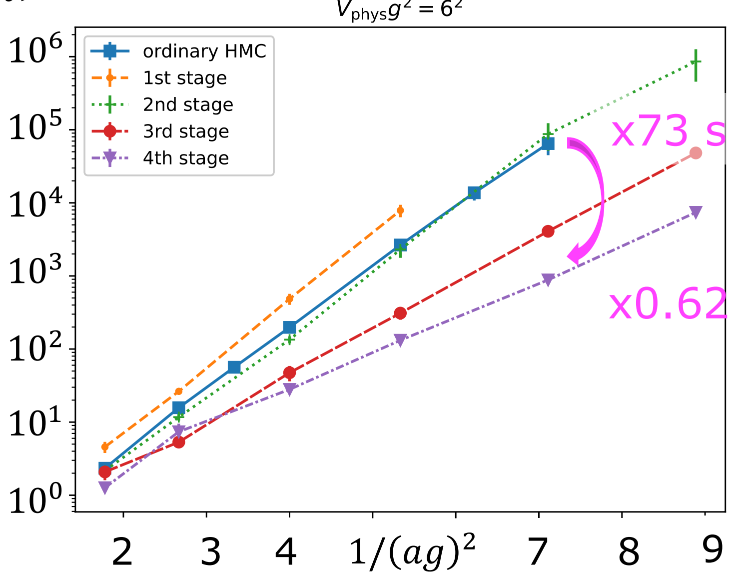
$\tau_{\rm int} \simeq 1$

Execution time/conf (no parallelization)



2nd stage gives dominant cost

$\tau_{\rm int}(Q)$ in wall-clock time



x73 speed up

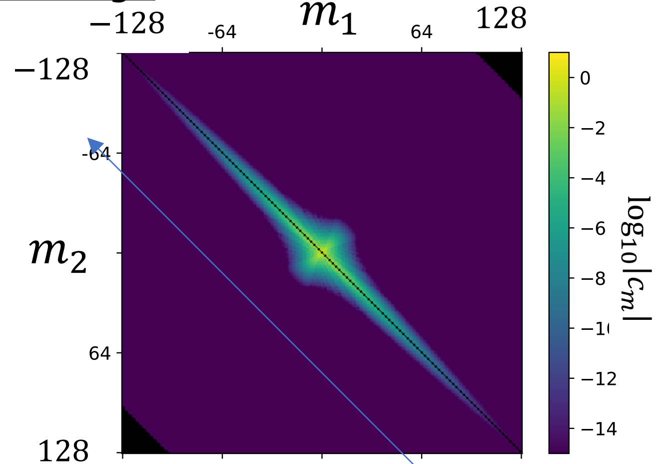
&

x0.62 in the exponent w/ four stages

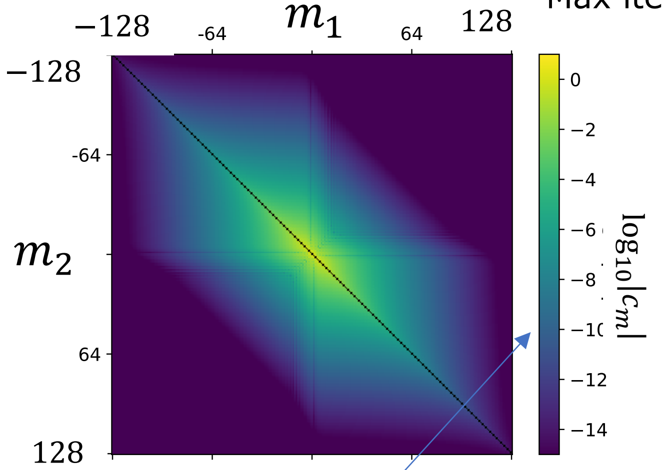
Kernel shapes @β=8.89 (1/1)

NM+ in prep

1st stage



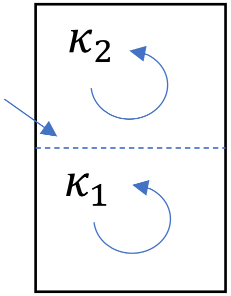
2nd stage



Solves with about an hour using GPU.
Max iteration count: ~9000

$$K_t = \sum_{\mathbf{m}} c_{t,\mathbf{m}} \cos(m_1\kappa_1 + m_2\kappa_2)$$

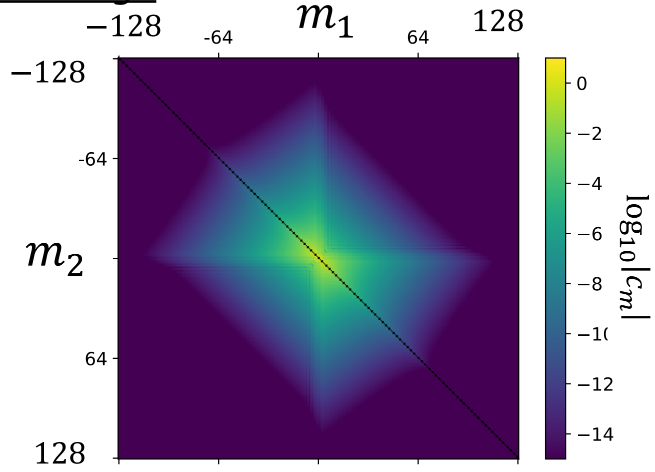
trivialized link



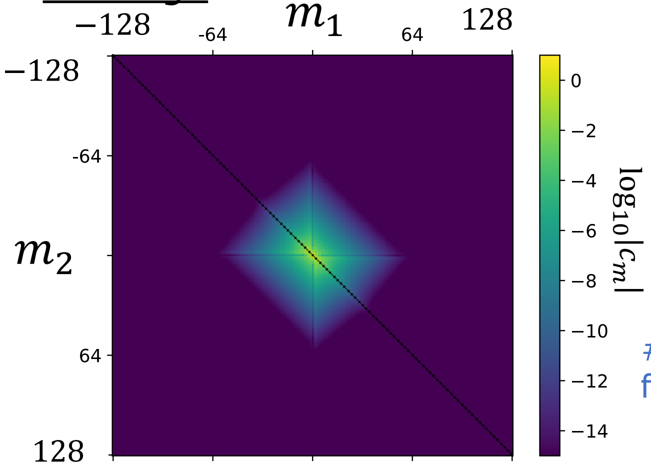
more spread out as $\beta \to \infty$

Due to higher-order loops in S_{eff}

3rd stage



4th stage

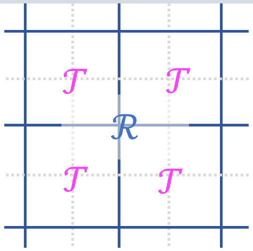


#(relevant loops) shrinks
for higher stages=coarse actions

- 2nd stage is the most expensive simply because ∃many terms.
- From $1/\beta$ expansion of the kernel, suppression of large winding is seemingly power-law.

Analyzing the tunneling (1/1)

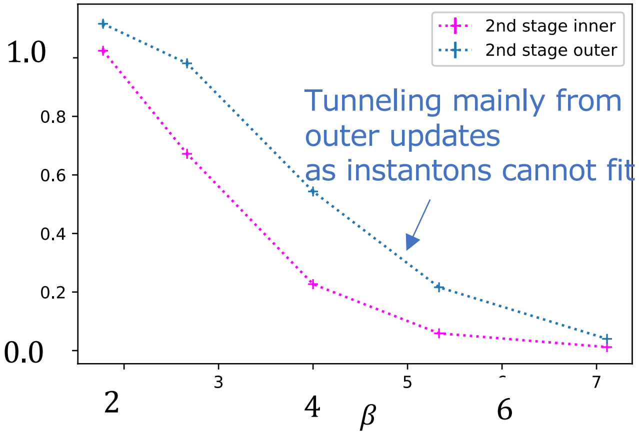
NM+ in prep



- By switching on/off the inner/outer updates alternatively, we investigate how tunneling is induced.

2nd stage

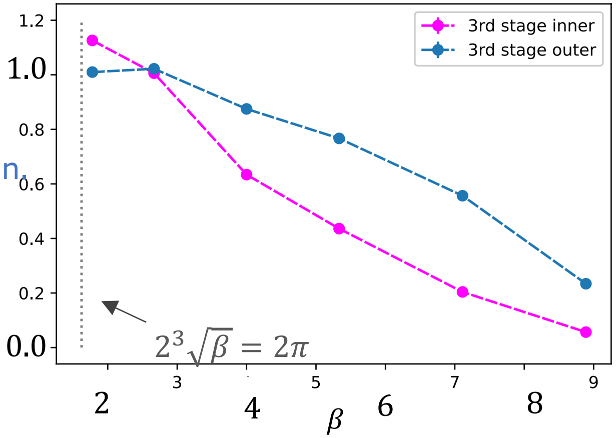
tunneling rate ($\equiv \langle |Q_n - Q_{n+1}| \rangle_{MC}$)



Tunneling mainly from outer updates as instantons cannot fit in.

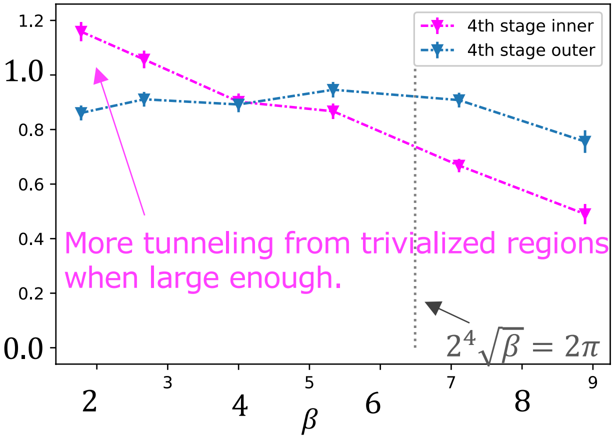
3rd stage

tunneling rate



4th stage

tunneling rate

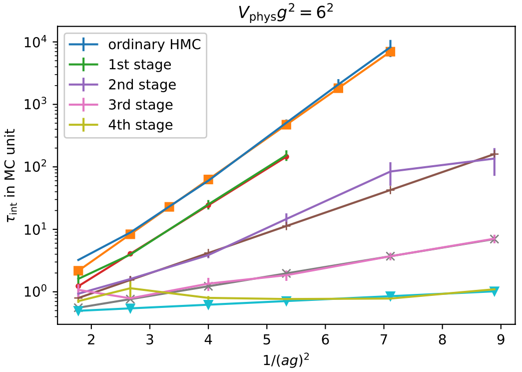


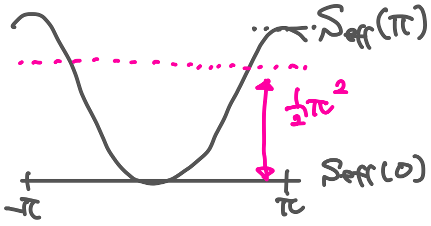
More tunneling from trivialized regions when large enough.

- At the large β regime, the ratio of the tunneling rates (in MC units) can be understood from the probability:

$$P\left(\frac{\pi^2}{2} > S_{\text{eff}}^{(\text{loc})}(\pi) - S_{\text{eff}}^{(\text{loc})}(0)\right), \pi \sim \mathcal{N}(0,1)$$

From fit	From prob
1: 0.886(52)	1: 0.856
2: 0.493(40)	2: 0.507
3: 0.235(14)	3: 0.242
4: 0.067(42)	4: 0.101





Aiming for large β, it seems more effective to get with field transformations a coarse-grained theory than enlarging completely trivialized regions.

Towards QCD (1/1)

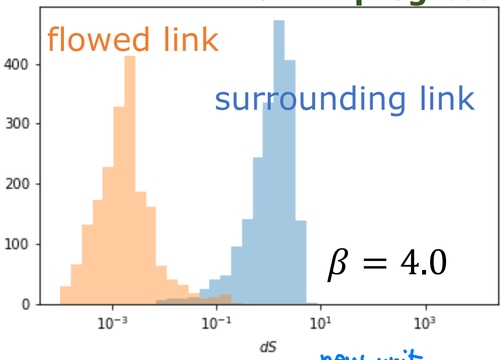
Including fermion

- We can expect exponential speed-up to remain when using exact S_{eff} in HMC (sometimes referred to as "FT-HMC")
  see Lüscher 09

  especially when fermion is present because of its additional cost.
  cf. Engel-Schaefer 11, *P.Boyle Plenary Tue*
- However, effect of this peculiar smearing on fermion is nontrivial.
  see also *J. Finkenrath Mon*

Forces (1^{st} stage, exact S_{eff})

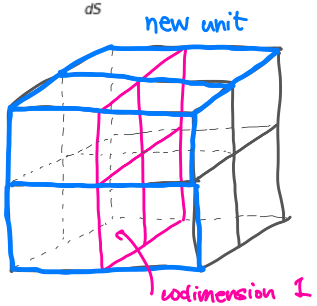
NM+ work in progress



Higher-dimension

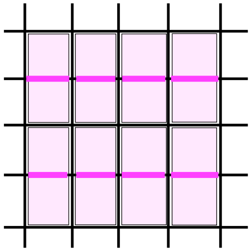
- trivializing map for codimension one surfaces → theory of 1x2 Wilson loops
  ∵ gauge invariance

  2D Wilson is special that links on a line are independent;
  much larger function space required for higher dimensions.



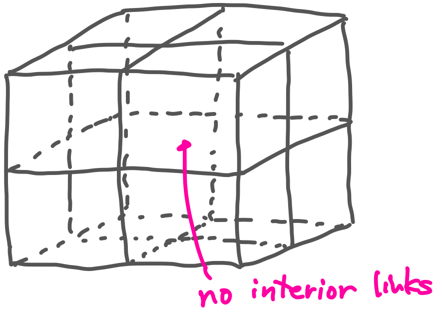
- Another possibility: freeze links surrounding local volume & trivialize interior.
  → "cage action", which can lower potential barriers for the remaining links.

⋮



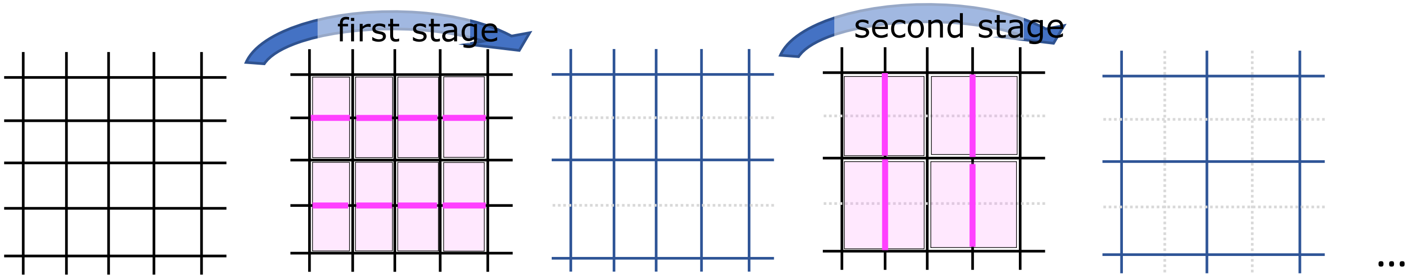
Non-Abelian groups

- We can work in the Wilson loop space to construct the function basis; labels may again be the winding numbers
  (though we need to take into account the noncommutativity and traces).
- Though there are Mandelstam constraints, **Mandelstam 79**
  use of overcomplete basis seems possible when using CG.

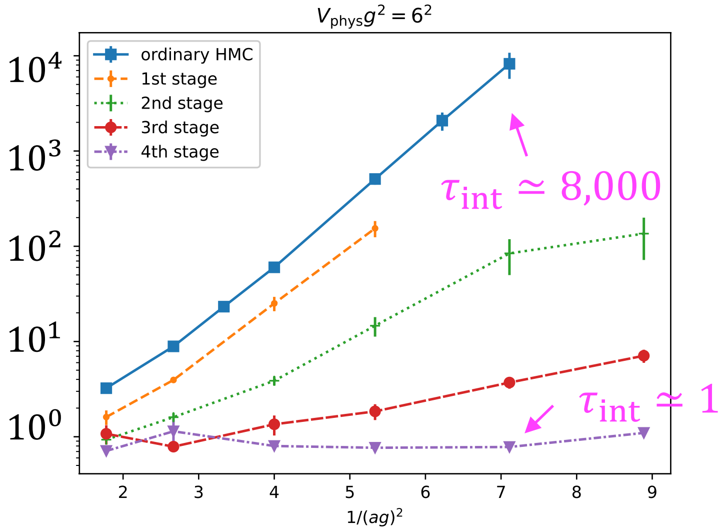


Summary (1/1)

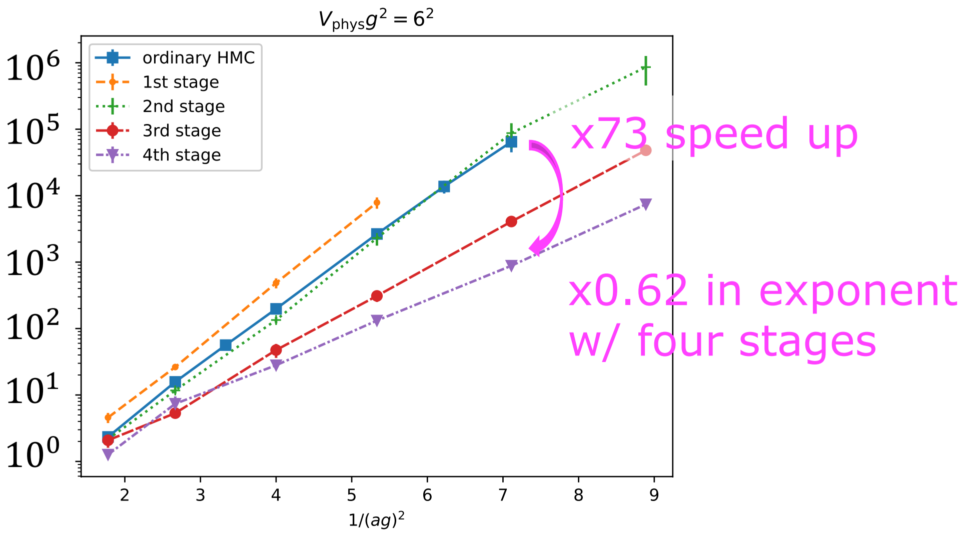
- We considered *decimation map* that can be regarded as a coarse-graining transformation.



$\tau_{\rm int}(Q)$ in MC unit



$\tau_{\rm int}(Q)$ in wall-clock time (no parallelization)



- It is true that the current investigation uses special features of 2D and U(1);

  however, we believe that having a method
  that works on this simplest model and has possible generalization directions
  will be a good starting point for developing algorithms for QCD.

Thank you.