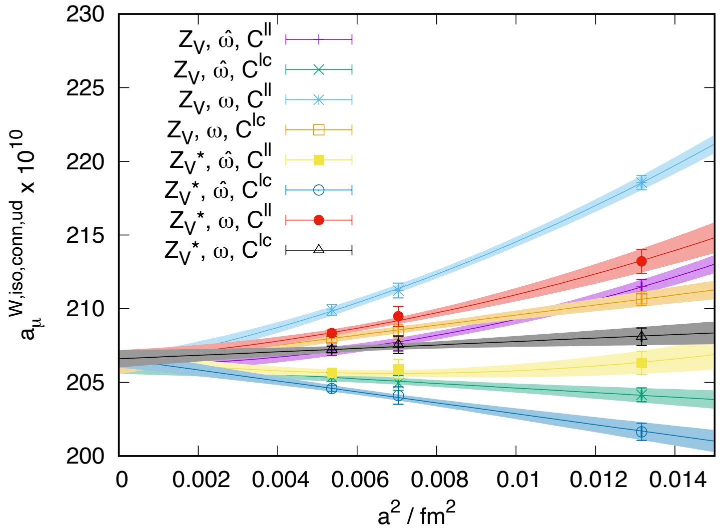
Diagnosing trivializing maps

Nobuyuki Matsumoto
(RIKEN BNL Research Center)

RIKEN R-CCS 16.02.2023

Boyle-Izubuchi-Jin-Jung-NM-Lehner-Tomiya,
PoS(LATTICE2022)229 [arXiv:2212.11387],
Work in progress

- Lattice calculation has been giving important inputs to the standard model.



g-2 window update
RBC/UKQCD (incl NM) 2301.08696

- A major source of uncertainty is the continuum extrapolation, which can be in principle reduced by adding results of fine lattices.

- However, as we reach the continuum limit, we encounter the *critical slowing down* when generating configurations, which adds at least exponential computational cost to the simple volume scaling.

Introduction (2/3)

Critical slowing down is a common pitfall of the Monte Carlo algorithm in critical statistical systems; in fact, major algorithm developments has aimed to accelerate Monte Carlo sampling:

- Overrelaxation

  Adler 81, Whitmer 84, Creutz 87

- Multigrid Monte Carlo

  Parisi 84, Goodman-Sokal 86 (see also Wolff 90)

  For Fermion Preconditioner: Wettig's talk on 15th and Peter's talk this morning

- Fourier acceleration/Riemannian manifold MC

  Parisi 84, Batrouni et al. 85,88,90 / Nguyen et al. 2112.04556

- Parallel tempering

  Swendsen-Wang 86, Geyer 91, Hukushima-Nemoto 96

  For Sign Problem: Fukuma's talk just before the break

  defect tempering:

  Hasenbusch 1706.04443, Berni-Bonanno-D'Elia 1911.03384, Bonanno-Bonati-D'Elia 2012.14000

- Cluster algorithm

  Swendsen-Wang 87, Wolff 89

- Trivializing map/normalizing flow

  Lüscher 0907.5491 / Rezende-Mohamed 15

  ML application: Akio's talk on 15th

  stochastic:

  Wu-Kohler-Noe 20, Caselle-Cellini-Nada-Panero 2201.08862

- Master field

  Lüscher 1707.09758, Bruno-Cè-Francis-Green-Hansen-Zafeiropoulos 2212.09533

  Francis' talk on 17th

- L2HMC, winding HMC, …

  Foreman-X.Y.Jin-Osborn 2105.03418, Albandea, et al. 2106.14234, …

Short timeline of the trivializing map

- Original proposal: trivializing map as a gradient flow **Lüscher 0907.5491**
- Testing the LO approximation in CP^{N-1} model **Engel-Schaefer 1102.1852**

  ⇨ performance asserted negatively

  "The reduction in the forces, ..., is compensated by the computational overhead"
- Machine learning approaches **Albergo-Kanwar-Shanahan 1904.12072, Foreman et al. 2112.01586 Bacchio-Kessel-Schaefer-Vaitl 2212.08469, Gomalizing flow**
- Wilson flowed HMC ⇨ $\times 1.5 \sim 2$ in tunneling rate in the unit of MC step **Jin LATTICE 2021 poster**
- Nonperturbatively improving the map with a Schwinger-Dyson equation including all the Wilson loops up to footprint 2

  Boyle-Izubuchi-Jin-Jung-NM-Lehner-Tomiya LATTICE2022 [2212.11387]

  ⇨ no vivid reduction in autocorrelation compared to the overhead

The fact that no visible gain has been found (apart from rapid developments in ML) suggests that neglected large loops greatly contribute to the autocorrelation.

A critical problem is that we only know little about the exact trivializing map.
More concretely, one may raise the questions:

- What does the exact flow kernel at large β look like?
  Is it really close to the Wilson flow?
- What are the appropriate basis functions to parametrize the kernel at large β?

Boyle-Izubuchi-L.Jin-Jung-NM-Lehner-Tomiya 2212.11387
(work in progress)

Aim of this work

- Using a simple 2D U(1) model, we analyze (spatially truncated) exact trivializing maps.
- This model is still far from the full QCD;
  however, there exists topological freezing at large β
  and the trivializing map is nontrivial.

➡ Good testing ground to study the properties of the exact map and to find effective approximations aiming for the full QCD.

This talk reports the ongoing study partially addressing the points:

- The convergence radius of the flow-time expansion of the flow kernel.
- Why the Wilson flow is not so effective.
- How many links need to be involved in the map to stimulate the tunneling
  (concrete results to be seen in future work).

Table of contents

- Introduction
- Review 1: Critical slowing down & topological freezing
- Review 2: Trivializing map
- Exploratory study in 2D U(1)
- Summary & outlook

Table of contents

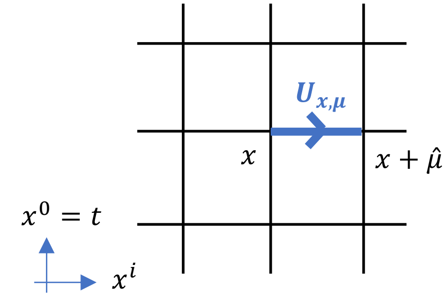
- Introduction
- Review 1: Critical slowing down & topological freezing
- Review 2: Trivializing map
- Exploratory study in 2D U(1)
- Summary & outlook

Goal Make physical predictions from the lattice path integral:

$$\langle \mathcal{O} \rangle \equiv \frac{\int (dU)\, e^{-S(U)} \mathcal{O}(U)}{\int (dU)\, e^{-S(U)}}$$

e.g., Wilson action **Wilson 74**

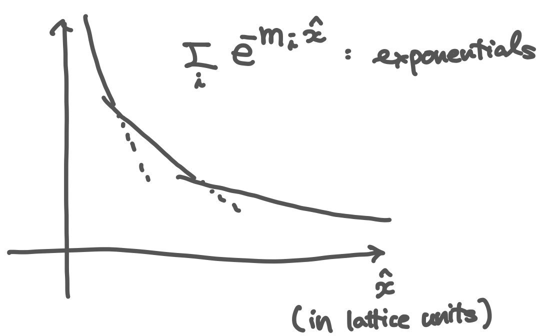
$$S(U) \equiv -\frac{\beta}{6} \sum_{x,\mu<\nu} \mathrm{Re\,tr}\left[U_{x,\mu} U_{x+\mu,\nu} U^{\dagger}_{x+\nu,\mu} U^{\dagger}_{x,\nu}\right]$$

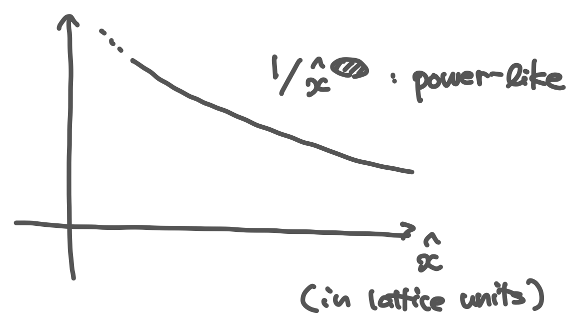


- We expect to have a finite correlation length *in physical units* in the continuum.

  infinite correlation length *in lattice units* (since $a \to 0$),
  which is a property of 2nd order phase transition. **Wilson 74**

- Generically, as we approach the critical point,
  more and more modes contribute to the correlator to give the quasi-long-range correlation.

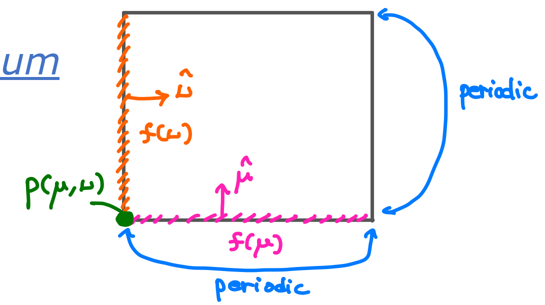




Such long correlations make the Monte Carlo simulation inefficient
because we need to move a large number of DOF
simultaneously and in a specific way to conserve the energy functional. *critical slowing down*

Further complication in QCD: *topological freezing*

Cause: nontrivial topological sectors of gauge field on T^4 *in the continuum*



- Periodicity is defined only *up to gauge transformation:*

$$A_\mu(x_\nu = L) = v_\nu(x)\left(\partial_\mu + A_\mu(x_\nu = 0)\right)v_\nu^{-1}(x).$$

't Hooft 81
cf. Dirac monopole

- The gauge function (or *transition function*) $v_\mu(x)$ encodes the topological information of the gauge field:

Lüscher 82, van Baal 82, Phillips-Stone 86
see also Kronfeld 88

$$Q \equiv \frac{-1}{16\pi}\int \mathrm{d}^4x \,\mathrm{tr}\, F_{\mu\nu}\tilde{F}_{\mu\nu}$$

$$= \frac{1}{24\pi^2}\left\{ \begin{array}{l} \sum_\mu \int_{f(\mu)} \mathrm{tr}\left(v_\mu d v_\mu^{-1}\right)^3 \\ -3\sum_{\mu\neq\nu}\int_{p(\mu,\nu)} \mathrm{tr}\,[dv_\nu^{-1}(x_\mu = L)v_\nu(x_\mu = L)v_\mu(x_\nu = 0)dv_\mu^{-1}(x_\nu = 0)] \end{array} \right\}$$

Solely expressed with $v_\mu(x)$!

- One can show that $Q \in \mathbb{Z}$ by, e.g., taking the pure gauge:

$$A_\mu dx^\mu = g^{-1}dg \qquad \left(\text{constraint:} \quad g^{-1}(x_\mu = L)g(x_\mu = 0) = v_\mu\right) \Rightarrow Q \equiv \frac{1}{24\pi^2}\int_{\partial V} \mathrm{tr}\,(g^{-1}dg)^3 \in \mathbb{Z}$$

count winding of SU(2) around S^3

- Topological sectors are disconnected
  ∵ they have $v_\mu(x)$ that cannot be continuously deformed to one another.

As the continuum limit is reached, the lattice gauge field acquires continuum-like nature. Correspondingly, configurations will be trapped in the emerging disconnected sectors during the Monte Carlo simulation (topological freezing).

More mathematical way to see the freezing is through the *geometrical definition of the lattice topological charge.* **Lüscher 82**

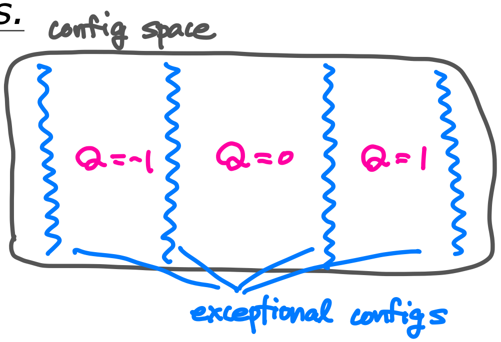
Simpler example: $U(1)$ on T^2

Phillips 85, see also Fujiwara et al. hep-lat/0001029

- Lattice topological charge: total winding of the *plaquette angles* κ_x:

$$Q^{(\mathrm{lat})} = \frac{-1}{2\pi}\sum_x \kappa_x \qquad \left(\kappa_x \equiv \frac{1}{i}\log\left(U_{x,0}U_{x+0,1}U^\dagger_{x+1,0}U^\dagger_{x,1}\right) \in [-\pi,\pi) \right)$$

Q is defined unambiguously except for the *exceptional configurations.*
s.t. $\exists x,\ \kappa_x = \pi$
($\therefore$ measure zero in path integral).

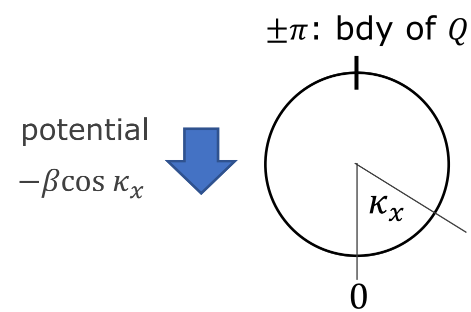


➡ Boundary of Q sectors are the exceptional configurations.

- Tunneling only occurs when the fluctuation becomes so large that the plaquette angle goes around the S^1 penetrating the potential barrier at $\pm\pi$.

However, such large fluctuation will be directly suppressed for the Wilson action at large β:

$$S(U) = -\beta\sum_x \cos\kappa_x\,.$$



➡ *Emergence of disconnected topological sectors.*

Similarly for SU(2) on T^4, exceptional configurations (= boundary of Q) consists of $\exists(\text{local Wilson loop}) = -1$, which will be suppressed at large β. **Lüscher 82**

(**Lüscher 82** gave the map from the nonexceptional gauge fields to the transition functions $v_\mu(x)$.)

Topological freezing (3/3)

A detour for the topological freezing: *open boundary condition* **Lüscher-Schaefer 1105.4749**

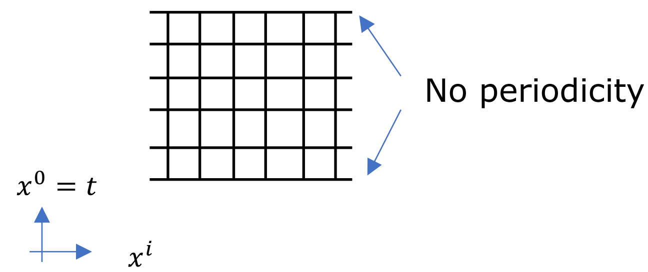
Pros

No more topological sectors in the continuum!

Cons

Need to consider the boundary effects.

In particular, translational invariance will be violated.



want to avoid if possible
∵ many statistical techniques assume the translational invariance

Regarding both critical slowing down and topological freezing,
they are rather intrinsic to the lattice simulation near the continuum (at large β).

it will be advantageous if one can use small β simulation to generate large β configurations.

trivializing map!

Table of contents

- Introduction
- Review 1: Critical slowing down & topological freezing
- Review 2: Trivializing map
- Exploratory study in 2D U(1)
- Summary & outlook

Trivializing map (1/4)

Idea **Lüscher 0907.5491**

- With a field transformation, we can obtain a new effective action:

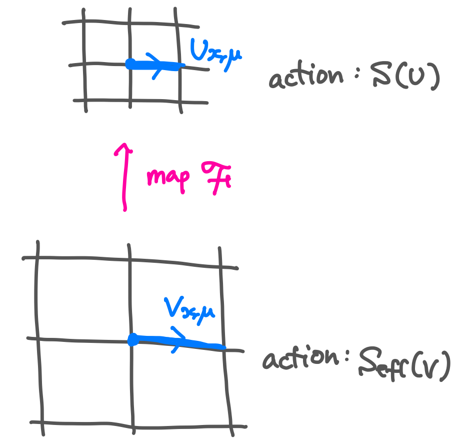
For $U = \mathcal{F}(V)$,

U: original (or "physical") variable

V: new (or "artificial") variable

$$Z \equiv \int dU\, e^{-S(U)} = \int dV \underbrace{\det \mathcal{F}_*(V)}_{\text{Jacobian}}\, e^{-S(\mathcal{F}(V))} \equiv \int dV\, e^{-S_{\text{eff}}(V)}.$$

$$\left[\; S_{\text{eff}}(V) \equiv S\big(\mathcal{F}(V)\big) - \ln \det \mathcal{F}_*(V) \;\right]$$



- We can perform the ordinary Monte Carlo sampling (e.g., HMC **Duane et al. 87**) in the V-space with the action $S_{\text{eff}}(V)$.

⇨ Prepare $\mathcal{F}$ so that the sampling in V-space becomes efficient!

- Ultimate $\mathcal{F}$: *trivializing map* **cf. in SUSY theory: Nicolai 80**

$$S_{\text{eff}}(V) = \text{const}$$

Such $\mathcal{F}$ maps the finite β theory to the strong coupling limit ($\beta = 0$), which is the opposite of where the critical slowing down occurs ($\beta = \infty$).

Trivializing map (2/4)

Writing down the Jacobian $\mathcal{F}_*(V)$

- Introduce a local parametrization $(\theta^a_{x,\mu})$ of the field space around a configuration $U_{x,\mu}$:

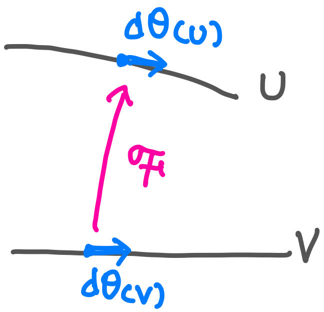
$$e^{\theta^a_{x,\mu} T^a} U_{x,\mu}.$$

T^a: $\mathrm{su}(3)$ generators. $\mathrm{tr}\left(T^a T^b\right) = -\frac{1}{2}\delta^{ab}$

- Haar measure: $(dU) \propto \prod_A d\theta^A$ $\quad A \equiv (x, \mu, a)$ labels the DOF

- $\mathcal{F}_*(V) = (\mathcal{F}_*(V)^{AB})$ can be read off from the infinitesimals:

$$d\theta^A_{(U)} = \mathcal{F}_*^{AB}(V)\, d\theta^B_{(V)}.$$



For later convenience, we also define the derivative:

$$\partial^a_{x,\mu} U_{x,\mu} \equiv \lim_{t\to 0} \frac{\left(e^{tT^a} - 1\right) U_{x,\mu}}{t} = T^a U_{x,\mu}.$$

In other words, $\partial^a_{x,\mu} = \partial_{\theta^a_{x,\mu}}|_{\theta=0}$.

<u>Comment on the convention</u>

In Lüscher 0907.5491, the symbol $\theta^a_{x,\mu}$ is used for the *Maurer-Cartan form* $\Theta^a_{x,\mu}$:

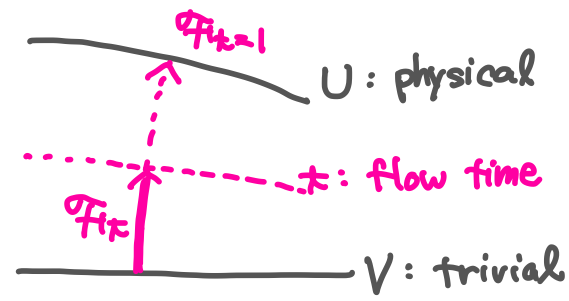
$$\Theta^a_{x,\mu} = \left(1 + O(\theta)\right) d\theta^a_{x,\mu}$$

(at each point $U_{x,\mu}$ on the group manifold).

[$\Theta^a_{x,\mu}$ is the dual of $\partial^a_{x,\mu}$: $\langle \Theta^A, \partial^B \rangle = \delta^{AB}$.] **See, e.g., Chevalley 46**

- Lüscher chose the gradient flow ansatz: **Lüscher 0907.5491**

$$\dot{\mathcal{F}}_t(U)_{x,\mu} = -T^a \partial^a_{x,\mu} K_t(U) \cdot U_{x,\mu} \,.$$



- Require that $\mathcal{F}_t$ trivializes the theory at $t = 1$:

$$S_{\mathrm{eff},\,t}(V) \overset{\Delta}{=} S\big(\mathcal{F}_t(V)\big) - \ln\det \mathcal{F}_{t*}(V) \overset{*}{=} (1-t)\, S\big(\mathcal{F}_t(V)\big)$$

requirement

NB
- $S(U)$: original action
- $S_{\mathrm{eff},t}(V)$: effective action
- $K_t(U)$: flow kernel

⬇ d/dt

$$-(\partial^A)^2 K_t + t\,\partial^A S\,\partial^A K_t \overset{*}{=} -S$$ (up to const; ignored hereafter)

from Jacobian from action

Solving the map has boiled down to solving a linear differential equation!

- For convenience we define

$$\mathcal{L}_t \equiv -(\partial^A)^2 + t\,\partial^A S\,\partial^A$$

$$\therefore\ \ \mathcal{L}_t K_t \overset{*}{=} -S$$

Existence **Lüscher 0907.5491**

- The differential operator $\mathcal{L}_t = -(\partial^A)^2 + t\,\partial^A S\,\partial^A$ is
  - elliptic ($\therefore$ bounded from below)
  - symmetric with respect to the inner product: $(\psi, \phi) \equiv \int (dU)\, e^{-t\,S(U)}\, \psi^*(U)\, \phi(U)$

    i.e., $(\psi, \mathcal{L}_t \phi) = (\mathcal{L}_t \psi, \phi)$.

$\therefore$ $\mathcal{L}_t$ shares (mostly) the same properties with the Hamiltonian in QM.

⇒ $\mathcal{L}_t$ is diagonalizable and the eigenvectors form a complete set.

- For a normalized eigenvector ψ_n, $\lambda_n = (\psi_n, \mathcal{L}_t \psi_n) = \int (dU)\, e^{-tS} |\partial^A \psi_n|^2 \geq 0$.

⇒
- eigenvalues of $\mathcal{L}_t$ are nonnegative
- zero-mode is unique, which is constant $(\because \partial^A \psi_n = 0)$

⇒ $\mathcal{L}_t^{-1}$ can be taken after removing the zero mode, and thus modulo constant.

$\therefore$ The solution of $\mathcal{L}_t K_t = -S$ exists!

<u>NB</u> Structure-wise $\mathcal{L}_t = (-\partial^A + t\,\partial^A S) \circ \partial^A$ may be related to the Fokker-Planck Hamiltonian of Langevin dynamics.

cf. In a RG flow of ϕ^4: Abe-Fukuma 1805.12094

Lüscher further gave a way to construct the map as a t-expansion: **Lüscher 0907.5491**

- Expand K_t as a Taylor series:

$$K_t = \sum_{m\geq 0} t^m K^{(m)}.$$

Plug into the equation:

$$-(\partial^A)^2 K_t + t\,\partial^A S\,\partial^A K_t = -S.$$

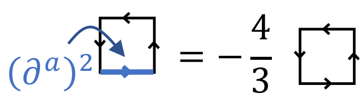
Matching the powers of t,

$$\begin{cases} \mathcal{L}_0 K^{(0)} = S, \\ \mathcal{L}_0 K^{(m)} = -\partial S \cdot \partial K^{(m-1)} \ (m \geq 1). \end{cases}$$

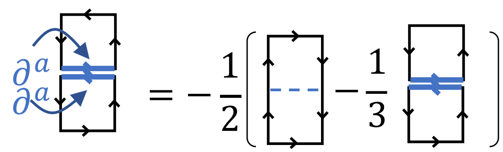
- This recurrence equation can be inverted order by order.

Basic machinery

Recall:
$\partial^a_{x,\mu} \sim T^a$ insertion

- 

$$\mathrm{tr}[(T^a)^2 \dots] = -\tfrac{4}{3}\mathrm{tr}[\dots]$$

- 

$$\mathrm{tr}[T^a A]\,\mathrm{tr}[T^a B] = -\tfrac{1}{2}\left(\mathrm{tr}[AB] - \tfrac{1}{3}\mathrm{tr}A \cdot \mathrm{tr}B\right)$$

- Radius of convergence proven to be finite (more at this point later)

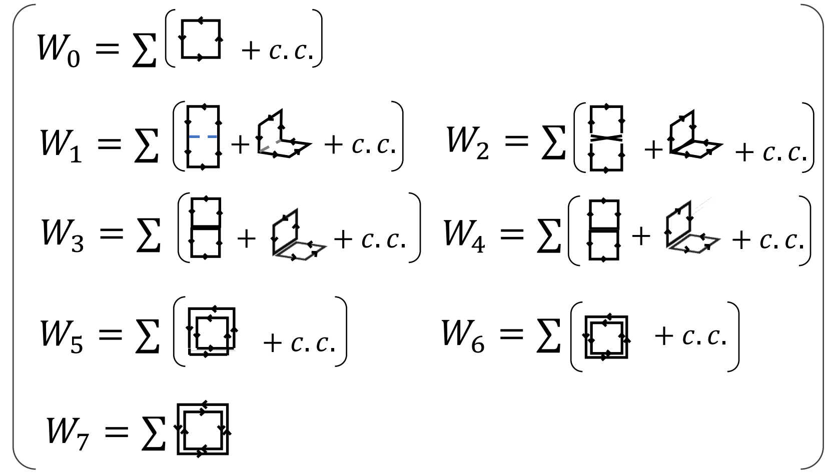
- Solution for the Wilson action case: **Lüscher 0907.5491**

$$K_t = -\frac{\beta}{32} W_0$$ ← LO: plaquette

$$+t\frac{\beta^2}{192}\left(-\frac{4}{33}W_1 + \frac{12}{119}W_2 + \frac{1}{33}W_3 - \frac{5}{119}W_4 + \frac{3}{10}W_5 - \frac{1}{5}W_6 + \frac{1}{9}W_7\right)$$

NLO: rectangle, chair, twisted rectangle …
"footprint 2 shapes"

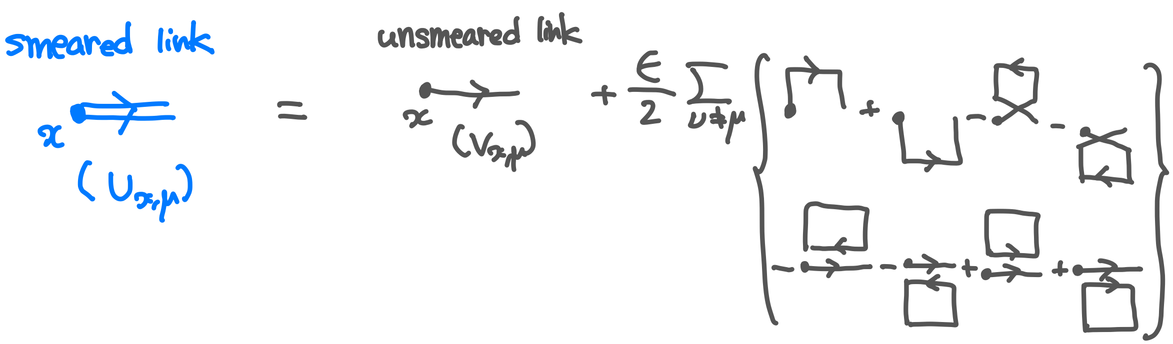
$$+O(t^2)$$



- Leading order: Wilson flow = stout smearing

Morningstar-Peardon hep-lat/0311018

For More Details: Nagatsuka's poster



Improving the map = adding more complicated shapes in RHS

- Lüscher's proposal: use the truncated kernel as an approximated kernel.
LO=Wilson flow.

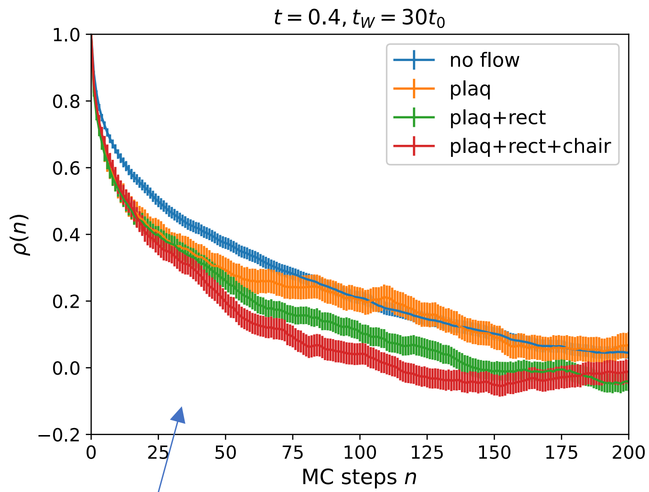
$8^3\times 16, \beta = 0.89$ DBW2 (4d) $(a^{-1} = 1.49\ \mathrm{GeV})$

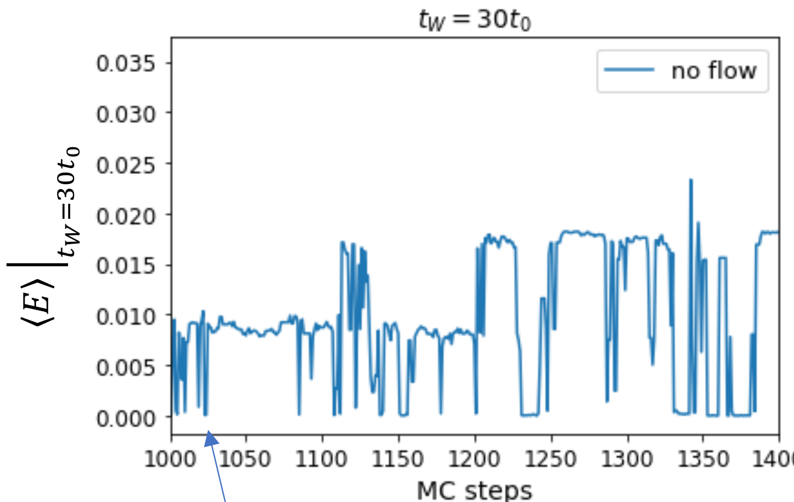
Boyle-Izubuchi-Jin-Jung-NM-Lehner-Tomiya 2212.11387

t_W: Wilson smearing flow time

$t_W^2 \langle E \rangle|_{t_W = t_0} = 0.3$ Lüscher 1006.4518

Normalized autocorrelation function $\rho(n)$ for the smeared energy density $(t_w = 30t_0)$

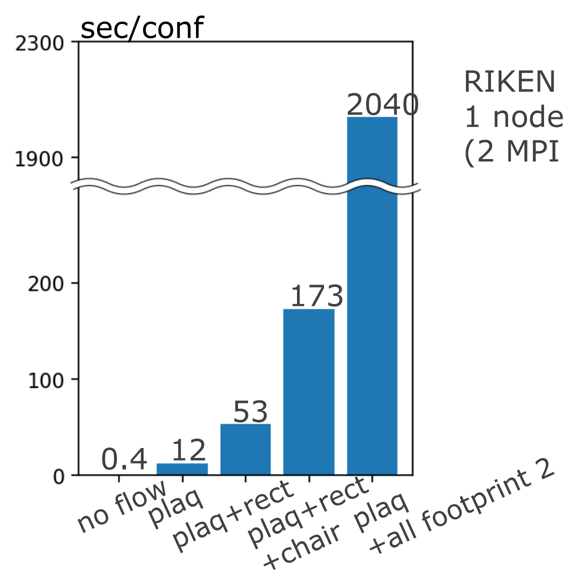




Faster decorrelation (in MC steps) by including extended loops

Smearing is performed so a long time that the energy reflects the instantons

Computational cost (1 step flow)



RIKEN HOKUSAI
1 node
(2 MPI x 40 OpenMP)

Table of contents

2D U(1) (revisited) (1/1)

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

$$= -\beta W_0$$

Topological charge: $Q = \frac{-1}{2\pi} \sum_x \kappa_x$

plaquette angle

$$\kappa_x \equiv \frac{1}{i} \log\left(U_{x,0} U_{x+0,1} U^\dagger_{x+1,0} U^\dagger_{x,1}\right),$$

$$W_0 \equiv \frac{1}{2} \sum \left[\square + c.c.\right]$$

$$= \sum \cos \kappa_x$$

(factor 2 absorbed in W_0)

Characteristic features

- Exactly solvable by character expansion **Gross-Witten 80, Rusakov 90**
- Energy density: UV divergent

$$\langle e \rangle \equiv \frac{1}{2V} \int d^2x\, F_{01}^2 \sim \frac{g^2}{2a^2}.$$

- Correlation length = 0
- topological susceptibility finite: $\chi_Q \sim \frac{g^2}{(2\pi)^2}.$

Thus in the following we deal with the topological freezing rather than the entire critical slowing down.

Though algorithmically it was difficult to see,
the t-expansion is the power-series expansion of $\frac{1}{1-x}$:

- Equation to solve: $[-(\partial^A)^2 + \beta t\, \partial^A W_0\, \partial^A] K_t = \beta W_0$
- Solution:

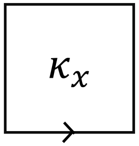
$$
\begin{aligned}
K_t &= \beta \cdot \frac{1}{-(\partial^A)^2 + \beta t \partial^A W_0\, \partial^A} \cdot W_0 \\
&= \beta \cdot \frac{1}{1 - \beta t (\partial^A)^{-2} \cdot \partial^A W_0\, \partial^A} \cdot [-(\partial^A)^{-2} W_0] \\
&= \beta \cdot \frac{1}{1 - \beta t\, \widehat{M}} \cdot W_0 \qquad \left(\widehat{M} \equiv (\partial^A)^{-2} \cdot \partial^A S\, \partial^A \right) \\
&= \beta \cdot \left(1 + \beta t \widehat{M} + (\beta t)^2 \widehat{M}^2 + \cdots \right) \cdot W_0
\end{aligned}
$$

<u>Convergence radius</u>

- Determined by the largest eigenvalue of the operator $\widehat{M}$;
  it then determines the applicable flow time t for a given β.
- Since t always appears in the product βt,
  large β trivialization can easily be out of convergence.

  (In such cases, no reason for the Wilson flow to be a good approximation.)

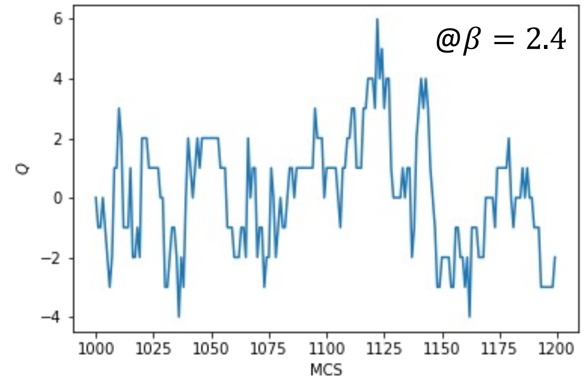
- System:

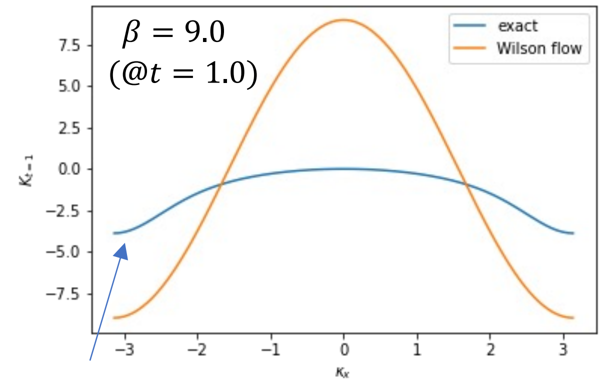


- Radius of convergence: $\beta t < 2.40$

- We compare the Wilson flow kernel with the exact trivializing map obtained by an inversion w/ CG

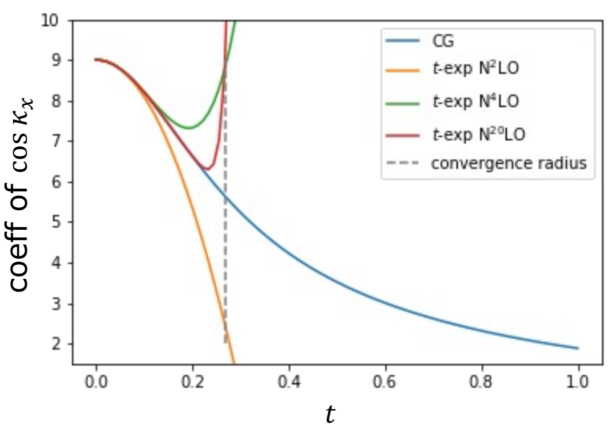
History of Q in 2D 16x16 ordinary HMC

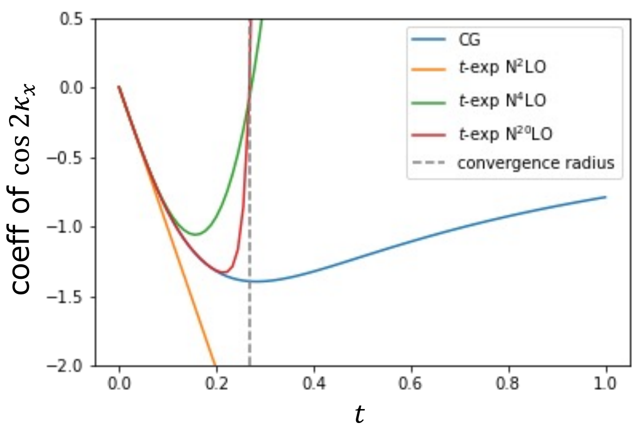


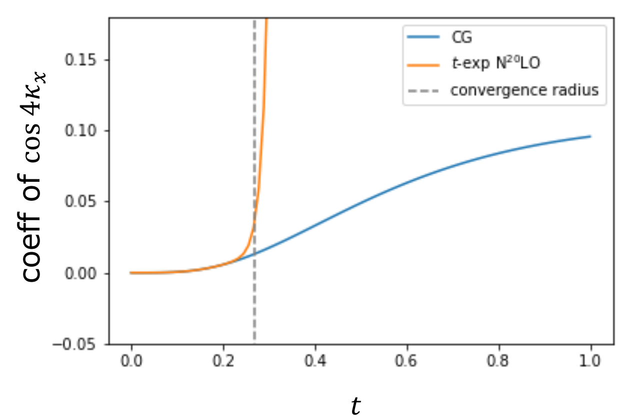


Angular dependence quite modest

Comparison of the coefficients

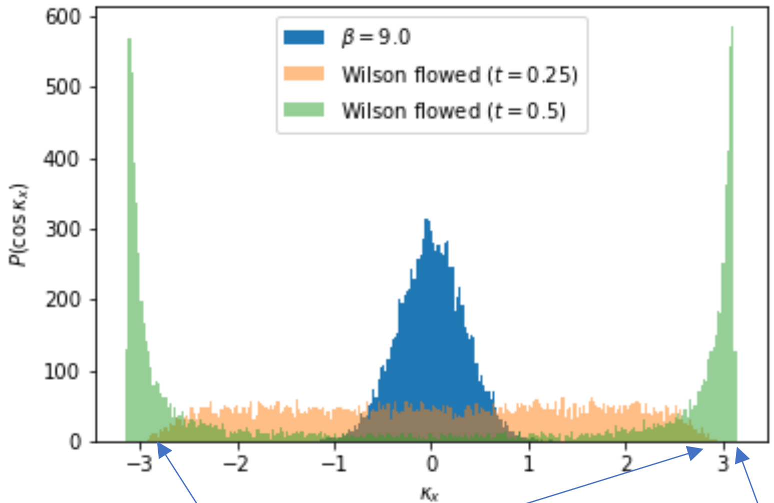






Does the exact flow kernel close to the Wilson flow? --- Apparently not.

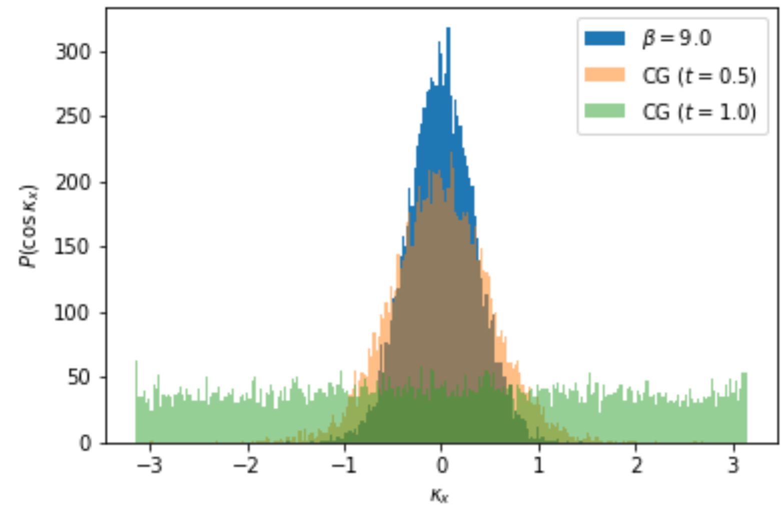
w/ Wilson flow kernel



peaky structure
just before $\theta=\pi$

$\theta=\pi$ has
vanishing probability

w/ kernel obtained by CG



($\epsilon=0.025$ in both cases)

- As expected from the gradient, Wilson flow overshoots (at least for this simplest system)
- The peaky structure can hinder the tunneling.
- Higher windings are essential to control the shape of K_t over the entire $U(1)$.

- It is also possible to make a $1/(\beta t)$ expansion:

$$\begin{aligned}
K_t &= \beta \cdot \frac{1}{-(\partial^A)^2 + \beta t\, \partial^A W_0\, \partial^A} \cdot W_0 \\
&= \beta \cdot \frac{1}{1 - \frac{1}{\beta t}(\partial^A W_0\, \partial^A)^{-1}(\partial^A)^2} \cdot \frac{1}{\beta t}(\partial^A W_0\, \partial^A)^{-1} \cdot W_0 \\
&= \beta \cdot \left(1 + \frac{1}{\beta t}\widehat{M}^{-1} + \frac{1}{(\beta t)^2}\widehat{M}^{-2} + \cdots\right) \cdot \frac{1}{\beta t}(\partial^A W_0\, \partial^A)^{-1} \cdot W_0
\end{aligned}$$

$$\left(\widehat{M} = (\partial^A)^{-2} \cdot \partial^A S\, \partial^A \right)$$

Convergence region complementary to the βt expansion.

- Recall that, in the strong coupling regime,
the plaquettes may be thought of as a spin-like collective variable.

By analogy, in the weak coupling regime, the function:

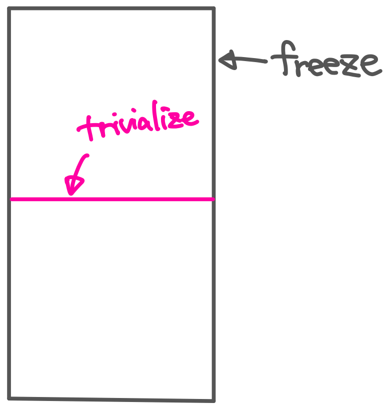
$$(\partial^A W_0\, \partial^A)^{-1} \cdot W_0$$

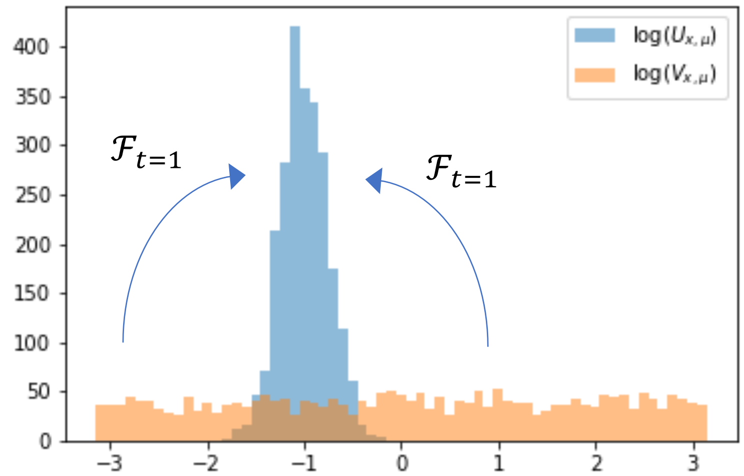
may be regarded as a (gauge-invariant) wave-like collective variable;
at least, its gradient gives the direction in which β decreases at $\beta = \infty$.

What are the appropriate basis functions to parametrize the kernel at large β?
--- The above Krylov basis!

- Since the growth of basis functions seems unavoidable,
  it should be a good strategy to trivialize a local region, not the entire system.

- As a first step, we trivialize one link in the 2D:



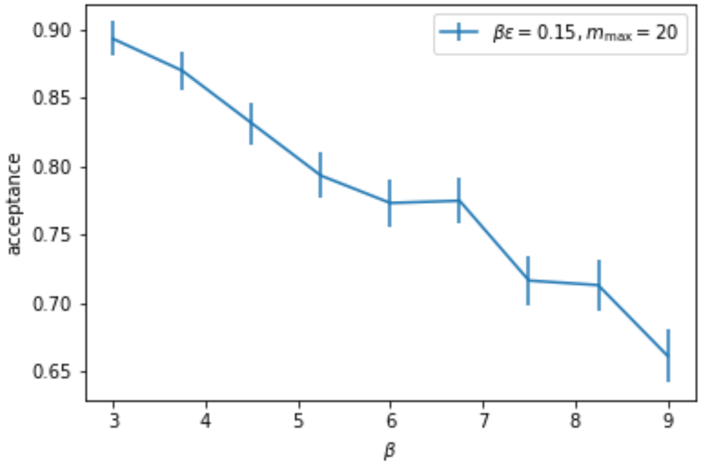


likeliness of appearance in U-space = geometrical area in V-space
(cf. cumulative distribution function)

- Still, we can update a quarter of the links simultaneously.

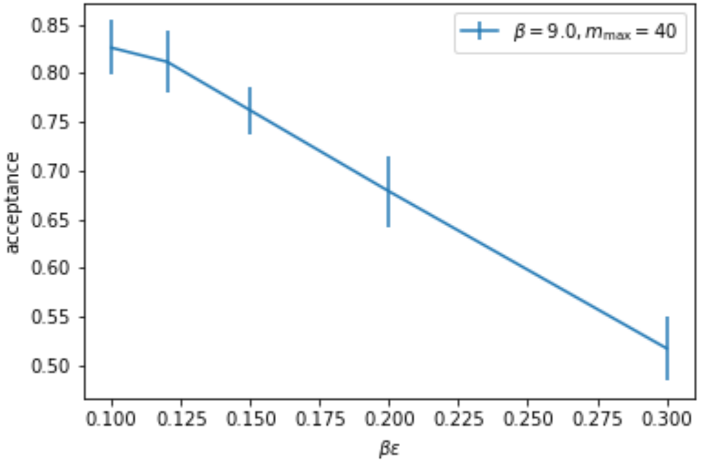
We measure the acceptance rate when using the simple Metropolis
(updates performed locally in parallel; if map is exact, acceptance=100%)

varying β

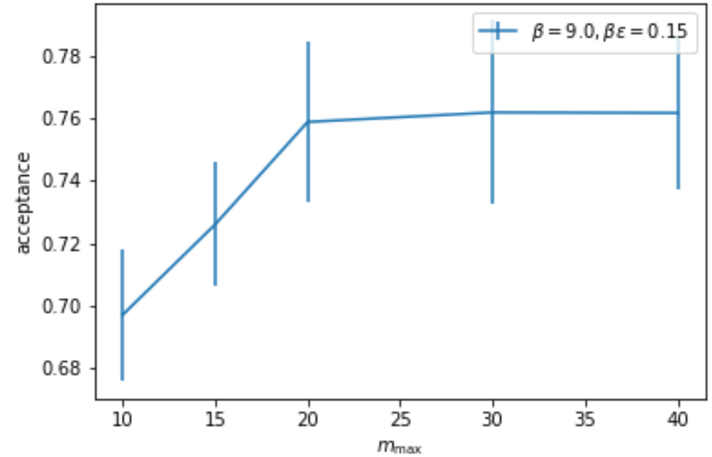


Larger function space required for larger β

varying step size ϵ



varying #irrep to include



Systematically improvable!

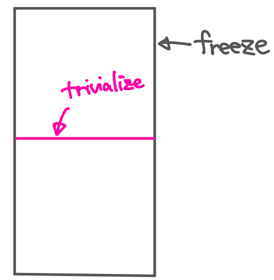
Trivializing a local region in 2D (3/3)

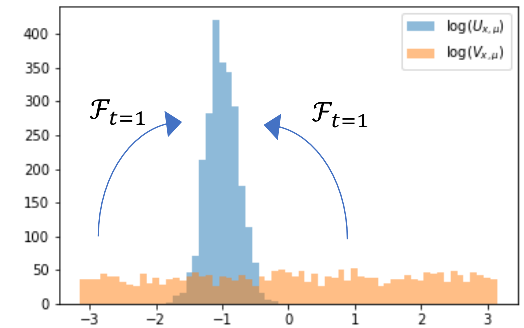
- Caveat: one link is not sufficient for the tunneling

  ∵ The value of the link is almost determined by the surrounding links (situation similar to a local heat-bath)

  Suppose the surrounding links are fixed to 1;

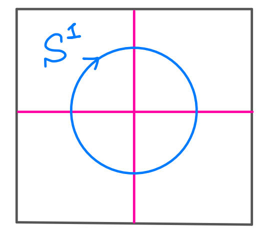
  then the *active link* can fluctuate only within the width determined by the plaquette distribution.





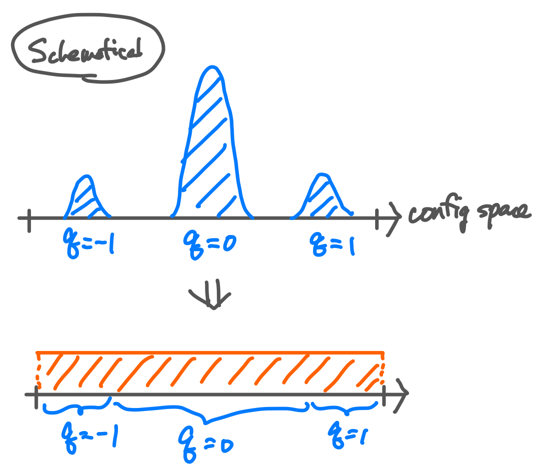
- How many links would we need to make an instanton?

  → With four links, one can create an instanton even when the surrounding links are fixed to 1.



The islands of configurations with nonvanishing probabilities will be again stretched to form a flat distribution.

In this way, the trivializing map can make up a global update algorithm (when having islands with nontrivial topology).



- We should however need a larger trivialization region for large β.

  ∵ appearance probability of nontrivial topology is determined by the topological susceptibility in the U-space; there is an obvious volume scaling.

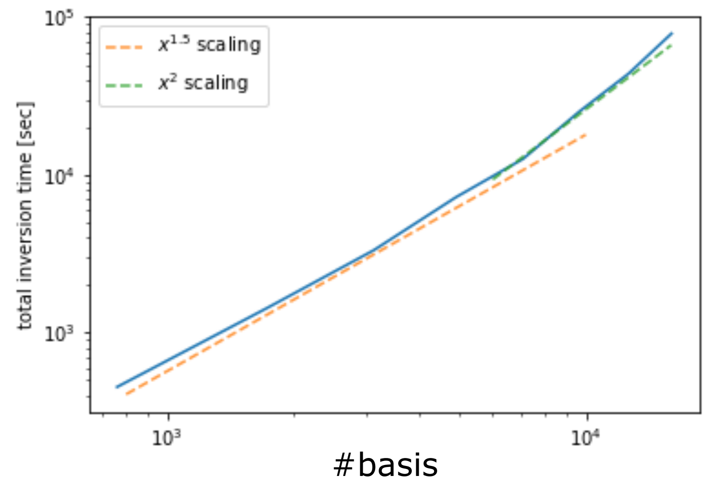
Since a matrix representation of $\mathcal{L}_t$ is sparse, we use, e.g., GMRES for the inversion.

$\mathcal{L}_t$ can be symmetrized by a preconditioning (as for the Fokker-Planck Hamiltonian), but then the treatment of zero-modes becomes tricky especially for multiple DOF.

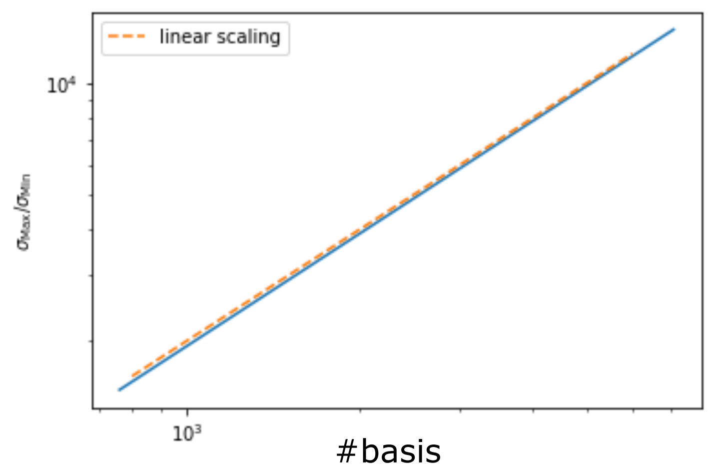
1 link trivialization case

- Time required to obtain the total map $\propto$ $\#\mathrm{basis}^{1.5}$

  (unless #basis unnecessarily large)



- Condition number $\propto \#\mathrm{basis}^1$



Seemingly pursuable problem! (coding mostly done)

Summary

- Using a simple 2D U(1) model, we analyzed (spatially truncated) exact trivializing maps.
- We reported the ongoing study partially addressing the points:
  - The convergence radius of the flow-time expansion of the flow kernel.
  - Why the Wilson flow is not so effective.
  - How many links need to be involved in the map to stimulate the tunneling.

Outlook

- Study field-transformed HMC with an effective exact trivializing map!
- Towards this goal, there are many technical issues not discussed here:
  - ϵ needs to be taken small when including higher irreps to ensure the map is one-to-one
    ($\because$ force can become larger for a kernel with higher windings)
    Adaptive step size? Higher order Runge-Kutta?
  - Even with the kernels obtained by BiCGStab,
    S_{eff} has a very thin peak at $\kappa = \pi$ due to $O(\epsilon^2)$ effect,
    which migrates to large forces in field-transformed HMC.
    Make the target action slightly differ from the flat distribution?
- Include fermion / develop algorithm that is capable of it. *Peter's talk this morning*

Thank you.