Use of Schwinger-Dyson equation in constructing an approximate trivializing map

Nobuyuki Matsumoto (RBRC)



LATTICE2022 8/8/2022 @Bonn

work in collaboration with P. Boyle, T. Izubuchi, L. Jin, C. Jung, C. Lehner, and A. Tomiya BNL/Edinburgh BNL/RBRC UConn/RBRC BNL/RBRC BNL/RBRC IPUT Osaka

(work in progress)

Introduction (1/3)

QCD lattice calculation has been taking an important and crucial role in the precision test of the standard model.

e.g., HVP contribution in muon g-2



The required precision is getting in a magnificent order, and calculation on fine lattices has become more demanding.

However, as we reach the continuum limit, we face the *critical slowing down*, which is often characterized by very long autocorrelation of topological charge.

Goal of this study

Develop the idea of <u>trivializing map</u> Nicolai 1980, Lüscher 0907.5491 map between finite β theory and $\beta = 0$ theory

and utilize it to generate fine lattice configurations with reduced numerical cost.

g-2 will be extensively addressed in this conference:

Mon HS2 17:50 $\gamma^* \gamma^* \rightarrow \pi \pi$ R.Briceno **Tue HS3** 15:00 isospin M.Hoferichter 15:20 isospin M.Bruno 15:40 coord space J.Parrino 16:20 twisted mass G.Gagliardi 16:50 improved Wilson S.Kuberski 17:10 disconnected S.Bacchio 17:30 EM coupling M.T.San Jose Perez 17:50 EM coupling H.Wittig, S.J.Perez low-mode noise M.Lynch PosterA strange&charm P.Tavella PosterB Wed HS6 14:40 isospin A.Portelli **HS4** 15:00 f_{K} , f_{π} F.Stokes **Thu HS3** 9:00 *C** bc R.Gruber 9:20 C* bc A.Altherr 9:40 FV correction S.Martins 10:00 lattice R-ratio A.De Santis 10:40 π pole to HLbL G.Kanwar 11:30 η pole to HLbL S.A.Burri **HS4** 11:30 UV log *a* R.Sommer **HS7** 15:10 IR $\pi\pi$ S.Paul Fri HS3 15:30 overlap&staggered mixed A.Kotov 15:50 $F_{\pi^0/n/n' \rightarrow \gamma^* \gamma^*}$ W.Verplanke 16:40 q₁ connected S.Gottlieb 17:00 BWM update B.Toth 17:20 QED correction G.Sinha Ray 17:40 overlap valence G.Wang 18:00 staggered finite *a* M.Golterman 18:20 RBC/UKQCD update C.Lehner Please let me know if I missed your work or if I have a wrong description.

Introduction (2/3)

There are many promising approaches for solving the critical slowing down/topological freezing:

Trivializing/normalizing map

Mon HS7 14:00 Flow-based density of states for complex actions Julian Urban

14:20 Stochastic normalizing flows for lattice field theory Elia Cellini

15:00 Gauge-equivariant flow models for sampling in lattice field theories with pseudofermions Fernando Romero-Lopez

16:30 Status update on flow models for gauge field generation **Phiala Shanahan**

17:50 Machine Learning Trivializing Maps Joe Marsh Rossney

18:10 Learning trivializing flows **David Albandea**

Wed HS717:50 Generative models for scalar field theories: how to deal with poor scaling? Javad Komijani

Master field analysis

Mon HS5 16:40 Translating topological benefits in very cold master-field simulations Anthony Francis

Metadynamics

Mon HS7 17:30 Topology changing update algorithms for SU(3) gauge theory **Timo Eichhorn**

Metadynamics Surfing on Topology Barriers in the Schwinger Model **Philip Rouenhoff** Poster B

Fourier acceleration

Poster B Fourier acceleration in strongly-interacting linear sigma models **Joshua Swaim**

Parallel tempering

Tue HS5 14:20 Towards glueball masses of large-N SU(N) Yang-Mills theories without topological freezing via parallel tempering on boundary conditions Claudio Bonanno HS1 15:20 Parallel tempering algorithm applied for the deconfinement transition of guenched QCD Ruben Kara

Multicanonical approach

Tue HS5 15:20 Topological susceptibility in high temperature full QCD via staggered spectral projectors Francesco D'Angelo

Skewed detailed balance

Thu HS7 10:40 Towards the Application of Skewed Detailed Balance in Lattice Gauge Theories Joao C. Pinto Barros

Fermion algorithms

Sat HS2 11:00 Review on Algorithms for dynamical fermions Jacob Finkenrath

Please let me know if I missed your work.

Also around BNL there are intense studies on

- Proceedings of lattice 2021
- Gauged-fixed Fourier acceleration Sheta, Zhao, Christ 2108.05486 - Riemannian manifold HMC Nguyen, Boyle, Christ, Jang, Jung 2112.04556 - L2HMC Foreman, X.Y.Jin, Osborn 2201.01582 - Machine learned trivializing map X.Y.Jin 2201.01862

 Alternatively to Luscher's trivializing flow (flow kernel obtained as a *t*-expansion), Lüscher 0907.5491
 flow time

we design the approximate trivializing flow with the <u>Schwinger-Dyson (SD) equation.</u>

Gonzalez-Arroyo, Okawa 1987, de Forcrand, Perez, Hashimoto, Hioki, Matsufuru, Miyamura,

Nakamura, Stamatescu, Tagoi, Takaishij, Umeda hep-lat/9806008

• By extending L. Jin's code of field-transformed HMC to include generic flow kernels, Jin LATTICE 2021

we perform the HMC with the resulting exact effective action after this flow.

Advantages of this method

- the basis for the flow kernel can be chosen by hand
- can be applied to the general action of interest
- the coefficients in the kernel are determined by lattice estimates of the observables; no need for analytic calculation such as *t*-expansion
- truncation effects and goodness of the flow can be measured by the force norm
- We apply our method to Wilson and DBW2 actions and show that:
 - With the SD method, we can have a better control of the effective action than *t*-expansion
 - In particular cases, faster decorrelation (in MC step unit) is observed for long-ranged observables by adding rectangle and chair to the flow
 - However, we have large algorithmic overhead, and need to check the scaling with larger statistics to confirm the actual benefits at large β

Trivializing map

Trivializing map (1/2) Lüscher 0907.5491

For a field transformation $\mathcal{F}_t: V \mapsto U = \mathcal{F}_t(V)$, we have the effective action $S_{\text{eff},t}(V)$: ٠

or $\partial_{x,\mu}^{a}U_{x,\mu} = T^{a}U_{x,\mu}$

 $\int dU \, e^{-S(U)} \mathcal{O}(U) = \int dV \det \mathcal{F}_{t*}(V) \, e^{-S(\mathcal{F}_t(V))} \mathcal{O}(\mathcal{F}_t(V))$ $\equiv \int dV \, e^{-S_{\text{eff},t}(V)} \mathcal{O}(\mathcal{F}_t(V))$ i.e.,

$$S_{\text{eff},t}(V) \equiv S(\mathcal{F}_t(V)) - \ln \det \mathcal{F}_{t*}(V)$$

 $\mathcal{F}_{t*}(V) = (\mathcal{F}_{t*}^{AB}(V))$: Jacobian matrix

 $d\theta^{A}_{(U)} = \mathcal{F}^{AB}_{t*}(V) d\theta^{B}_{(V)}$ where

 T^a : su(3) generators, tr $(T^aT^b) = -\frac{1}{2}\delta^{ab}$ Haar measure: $(dU) \propto \prod_A d\theta^A$ $A = (x, \mu, a)$ labels DOF $U_{x,\mu}' = e^{\theta_{x,\mu}^a T^a} U_{x,\mu}$

Lüscher proposed to use the *trivializing map* \mathcal{F}_t in <u>HMC</u> Duane, Kennedy, Pendleton, Roweth 1987 that makes the finite β theory mapped to the strong coupling limit $\beta = 0$ at t = 1:

 $S_{\text{eff }t=1}(V) = \text{const.}$

V will be decorrelated faster under the trivial action, and so does the configurations $U = \mathcal{F}_t(V)$



Lüscher 0907.5491

 (∂^a)

tr

tr[T

Gradient flow ansatz: ٠

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

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

 $S_{\text{eff.}t}(V) \stackrel{\scriptscriptstyle \Delta}{=} S(\mathcal{F}_t(V)) - \ln \det \mathcal{F}_{t*}(V)$ ↓ d/dt equation for the kernel:

> $-(\partial^A)^2 \tilde{S}_t + t \,\partial^A S \,\partial^A \tilde{S}_t = S \qquad (\text{up to irrelevant const})$ from Jacobian from action

Solution (for Wilson action $S = S_W$): •

$$\begin{split} \tilde{S}_t &= -\frac{\beta}{32} W_0 & \leftarrow \text{LO: plaquette} \\ &+ t \frac{\beta^2}{192} \Big(-\frac{4}{33} W_1 + \frac{12}{119} W_2 + \frac{1}{33} W_3 - \frac{5}{119} W_4 + \frac{3}{10} W_5 - \frac{1}{5} W_6 + \frac{1}{9} W_7 \Big) \\ &+ O(t^2) & \leftarrow \text{NLO: rectangle, chair, twisted rectangle ...} \\ & \quad \text{``footprint 2 shapes''} \end{split}$$

Strategy of the Schwinger-Dyson method (1/1)

 Note that the *t*-expansion is performed around *t* = 0; this corresponds to expanding *around the trivialized theory*.

In fact, the expansion parameter *t* appears in the combination βt \therefore approximation is better for small β .

However, our primary target is the large β theory; \therefore we rather want to approximate the map based on the information at large β .

- In the Schwinger-Dyson (SD) method below, we sequentially determine the flow *from the finite* β *theory*.
 - At each intermediate step, we determine the effective couplings by the SD equation.
 - We then design the flow to decrease the couplings from the lattice estimates of Wilson loops.

This determination of the flow can be seen as iteratively reexpanding the flow kernel at each step.



Schwinger-Dyson equation

• Suppose that

 $S_{\rm eff}(V) = \sum_j \beta_j W_j$

W_j: Wilson loops we omit the subscript *t* momentarily

• The couplings β_i obey the linear equation:

 $\sum_{j} \beta_{j} \left\langle \partial^{A} W_{j} \partial^{A} W_{i} \right\rangle_{S_{\text{eff}}} = \left\langle (\partial^{A})^{2} W_{i} \right\rangle_{S_{\text{eff}}} \qquad \langle \cdot \rangle_{S_{\text{eff}}}: \text{ expectation value with respect to } S_{\text{eff}}$

- Consider a variation using W_i as the flow kernel:

$$\delta V = -\epsilon T^A \partial^A W_i \cdot V$$

- The path integral is invariant under this variation (Schwinger-Dyson equation):

$$0 = \delta \int (dV) \ e^{-S_{\text{eff}}(V)} = \int (dV) \ e^{-S_{\text{eff}}(V)} \epsilon \ \frac{[-(\partial^A)^2 \ W_i}{\text{from Jacobian}} + \frac{\partial^A S_{\text{eff}} \ \partial^A W_i]}{\text{from action}}$$

- Combining this formula with the expansion of $S_{\text{eff}}(V)$:

 $\sum_{j} \beta_{j} \left\langle \partial^{A} W_{j} \right. \partial^{A} W_{i} \right\rangle_{S_{\text{eff}}} = \langle (\partial^{A})^{2} W_{i} \rangle_{S_{\text{eff}}} \square$

• Generically, we need infinite number of couplings to exactly parametrize $S_{\text{eff}}(V) = \sum_{j} \beta_{j} W_{j}$.

$$\sum_{j} \beta_{j} \left\langle \partial^{A} W_{j} \partial^{A} W_{i} \right\rangle_{S_{\text{eff}}} = \left\langle (\partial^{A})^{2} W_{i} \right\rangle_{S_{\text{eff}}} : \text{ infinite-dimensional matrix}$$

• Instead, we can try to mimic $S_{eff}(V)$ with a finite basis:

We determine β'_i by:

$$\sum_{j}^{\prime} \beta_{j}^{\prime} \left\langle \partial^{A} W_{j} \right. \partial^{A} W_{i} \right\rangle_{S_{\text{eff}}} = \left\langle (\partial^{A})^{2} W_{i} \right\rangle_{S_{\text{eff}}}$$

: finite-dimensional matrix (*i* is also restricted to the finite range)

Such β'_j give the best approximation of $S_{\rm eff}(V)$ in the sence that it minimizes the norm

$$||S_{eff} - S'_{eff}||_{S_{eff}}, \text{ where } ||S||_{S_{eff}}^2 \equiv \langle (\partial^A S)^2 \rangle_{S_{eff}}$$

L2 norm of the force
 \therefore The truncation error is calculable.

This Schwinger-Dyson method gives us a way to parametrize effective actions.

• Subtracted equation

$$\sum_{j}' (\beta_{j} - \beta_{j}') \left\langle \partial^{A}W_{j} \ \partial^{A}W_{i} \right\rangle_{S_{eff}} = 0$$
is the stationary condition:

$$\frac{\partial}{\partial \beta_{i}'} ||S_{eff} - S_{eff}'||_{S_{eff}}^{2}$$

$$= \frac{\partial}{\partial \beta_{i}'} \left\langle [\partial^{A}(S_{eff} - S_{eff}')]^{2} \right\rangle_{S_{eff}}$$

$$= -2 \sum_{j}' (\beta_{j} - \beta_{j}') \left\langle \partial^{A}W_{j} \ \partial^{A}W_{i} \right\rangle_{S_{eff}}$$

$$\equiv 0$$

• We parametrize $\frac{\tilde{S}_t}{\text{kernel function}}$ in the same truncated space:

 $\tilde{S}_t(V) = \sum_k' \gamma_{k,t} W_k$

• Differentiating the equation for $\beta'_{j,t}$:

$$\sum_{k}' \gamma_{k,t} \left\langle \partial^{B} W_{k} \partial^{B} \left[-(\partial^{A}) W_{i} + \partial^{A} S_{\text{eff},t}' \partial^{A} W_{i} \right] \right\rangle_{S_{\text{eff},t}} = -\sum_{j}' \dot{\beta}_{j,t}' \left\langle \partial^{A} W_{j} \partial^{A} W_{i} \right\rangle_{S_{\text{eff},t}}$$

$$(d/dt) \text{ acting on the Boltzmann weight} \qquad (d/dt) \text{ acting on } \beta_{j,t}'$$

This equation gives the coefficients $\gamma_{k,t}$ for a given $\dot{\beta}'_{j,t}$ (thus a trajectory of $S'_{\text{eff},t}$)

• We here require $\dot{\beta}'_{j,t} = -\frac{\beta'_{j,t}}{1-t}$ so that $\beta'_{j,t} = (1-t)\beta'_{j,t=0} = (1-t)\beta_{j,t=0}$ (linear decrease)

Linear equation for
$$\gamma_{k,t}$$
:

$$\sum_{k}' \gamma_{k,t} \left\langle \partial^{B} W_{k} \partial^{B} \left[-(\partial^{A}) W_{i} + \partial^{A} S_{eff,t}' \partial^{A} W_{i} \right] \right\rangle_{S_{eff,t}} = \frac{1}{1-t} \left\langle \partial^{A} S_{eff,t}' \partial^{A} W_{i} \right\rangle_{S_{eff,t}}$$
In practice, we use the numerical derivative with the five-point formula to calculate this matrix

Boyle, Izubuchi, L. Jin, Jung, NM, Lehner, Tomiya work in progress

Remarks (1/1)

• Some of the basis functions are not linearly independent ("Mandelstam constraints") Mandelstam 1979

We need to choose linearly independent basis in the inversion.

• We use the HMC algorithm with the exact transformed action $S_{\text{eff}}(V)$ developed by Luscher. Luscher 0907.5491

Advantages of designing the trivializing map with the Schwinger-Dyson equation

- the basis for the flow kernel can be chosen arbitrarily by hand
- can be applied to the general action of interest
- the coefficients in the kernel, $\gamma_{k,t}$, are determined by lattice estimates of the observables; no need for analytic calculation such as *t*-expansion
- truncation effects and goodness of the flow can be measured by the force norm

Software (1/1)

 Fully parallelized code based on qlat software (C++ codebase) <u>https://github.com/waterret/Qlattice</u>

We extended L. Jin's code of field-transformed HMC Jin LATTICE 2021 to include generic flow kernels including all the footprint 2 loops.

• Most costly part is the multiplication of a matrix including the Hessian $\partial^A \partial^B \tilde{S}_t$ in propagating the force.



 By dividing the directions of the flowed links and appropriately coloring/masking the lattice for each type of loops, we can run the multiplication in parallel. Lüscher 0907.5491, Boyda,

Kanwar. Pacaniere, Rezende, Albergo, Cranmer, Hackett, Shanahan 2008.05456



Volume scaling is basically linear O(vol) for local Wilson loop bases

Results

Computation resources (1/1)

• RIKEN HOKUSAI





• Univ of Tokyo Oakforest-PACS (retired)



picture taken from HD of C

 USQCD facility at BNL (KNL) funded by US DOE



picture taken from HP of CCS

We are grateful for these resources.

Evolution of effective action [Wilson action] (1/1)

 $8^4, \beta = 6.13$ Wilson $(a^{-1} = 2.56 \text{ GeV})$ Ce-Consonni-Engel-Giusti 1506.06052 Boyle, Izubuchi, L. Jin, Jung, NM, Lehner, Tomiya work in progress



Difference from the target trajectory of effective action

Evolution of effective action [Wilson action] (1/1)

 $8^4, \beta = 6.13$ Wilson $(a^{-1} = 2.56 \text{ GeV})$ Ce-Consonni-Engel-Giusti 1506.06052 Boyle, Izubuchi, L. Jin, Jung, NM, Lehner, Tomiya work in progress



Naively adding the rectangle term to the LO *t*-expansion makes the deviation more significant

Autocorrelation [DBW2 action] (1/3)

 $8^3 \times 16, \beta = 0.89$ DBW2 $(a^{-1} = 1.49 \text{ GeV})$ $(c_1 = -1.4008)$ Necco hep-lat/0309017

Boyle, Izubuchi, L. Jin, Jung, NM, Lehner, Tomiya work in progress





The increase is due to the increase of the nonzero matrix elements in $\partial^A \partial^B \tilde{S}_t$.

It seems essential to design the flow with a few types of loops to circumvent large algorithm overhead.

Difference from the target trajectory of effective action

15/19

Autocorrelation [DBW2 action] (2/3)

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





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



Autocorrelation [DBW2 action] (3/3)

Faster decorrelation is observed by the extended loops, but the autocorrelation is not controlled completely: (and statistics should be increased for definite conclusions)



Discussion (1/1)

• Integrated autocorrelation time of the topological charge increases exponentially with β (\therefore with a power of a^{-1}).

If not complete trivialization, we can have large benefits by changing the exponent.

Possible strategies?

- SD method itself has a large room for optimization choose a different path of β'_t (e.g., exponential decrease, decrease c₁), t-dependent step sizes, change basis at each step ...
- On the other hand, the relatively small improvements show that the truncated large loops contribute to increasing the autocorrelation.

In fact, we need an *infinite* number of loops to obtain the exact trivializing flow, Lüscher 0907.5491 with which we expect to decrease the autocorrelation of *all modes* of the system;

since we need to deal with finite number of loops, we can be more specific to particular slow observables.

e.g., change instanton potential to stimulate tunneling

cf. Smith, Teper hep-lat/9801008 DeGrand, Hasenfratz, Kovacs hep-lat/9801037 De Forcrand, Garcia Perez, Hetrick, Stamatescu hep-lat/9802017 Hasenfratz-Nieter hep-lat/9806026



 $\tau_{\rm int}(Q)$ increases exponentially with β

Summary and Outlook

Summary

- We proposed a way to design an approximate trivializing map with the Schwinger-Dyson equation Advantages of this method
 - the basis for the flow kernel can be chosen arbitrarily by hand
 - can be applied to the general action of interest
 - the coefficients in the kernel are determined by lattice estimates of the observables;
 no need for analytic calculation such as *t*-expansion
 - truncation effects and goodness of the flow can be measured by the force norm
- We showed that
 - With the SD method, we can have a better control of the effective action
 - We have positive effects in autocorrelation of long-ranged objects by adding rectangle and chair to the flow

Outlook

- Improve statistics for definite conclusions & confirm the scaling
- Develop more efficient strategies
 - E.g., Choose different trajectory of β'_t (exponential decrease, decrease $c_1, ...$)
 - Be specific to particular slow modes (change instanton potential, ...)

cf. Grid: <u>https://github.com/paboyle/Grid</u> GPT: <u>https://github.com/lehner/gpt</u>

Include fermion Do we need more extended Wilson loops in the flow?

Thank you.

Integrated autocorrelation time [DBW2 action] (1/1)

Boyle-Izubuchi-L. Jin-Jung-NM-Lehner-Tomiya, work in progress

