PROCEEDINGS OF SCIENCE

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

Peter Boyle,[a,b] **Taku Izubuchi,**[b,c] **Luchang Jin,**[d] **Chulwoo Jung,**[b] **Christoph Lehner,**[e] **Nobuyuki Matsumoto**[c,*] **and Akio Tomiya**[f]

[a] *School of Physics and Astronomy, University of Edinburgh, Edinburgh EH9 3FD, United Kingdom*
[b] *Physics Department, Brookhaven National Laboratory, Upton, NY 11973, USA*
[c] *RIKEN/BNL Research center, Brookhaven National Laboratory, Upton, NY 11973, USA*
[d] *Physics Department, University of Connecticut, Storrs, CT 06269, USA*
[e] *Universit ät Regensburg, Fakultät für Physik, 93040 Regensburg, Germany*
[f] *Faculty of Technology and Science, International Professional University of Technology, Osaka, Japan*

E-mail: nobuyuki.matsumoto@riken.jp

We construct an approximate trivializing map by using a Schwinger-Dyson equation. The advantage of this method is that: (1) The basis for the flow kernel can be chosen arbitrarily by hand. (2) It can be applied to the general action of interest. (3) The coefficients in the kernel are determined by lattice estimates of the observables, which does not require analytic calculations beforehand. We perform the HMC with the effective action obtained by the Schwinger-Dyson method, and show that we can have better control of the effective action than the known t-expansion construction. However, the algorithmic overhead is still large and overwhelming the gain though faster decorrelation is observed for long-range observables in some cases. This contribution reports the preliminary results of this attempt.

The 39th International Symposium on Lattice Field Theory (Lattice2022),
8-13 August, 2022
Bonn, Germany

*Speaker

PoS(LATTICE2022)229

1. Introduction

In recent years, lattice QCD has been taking an important and crucial role in the precision test of the standard model, as represented by the calculation of QCD contributions to the muon g-2 (see, e.g., [1]). Empowered by the recent hardware and experimental developments, target precision is becoming in a notable order, for which calculation on fine lattices has become an urgent demand. However, as we reach the continuum limit, we face the infamous critical slowing down, often characterized by long autocorrelation of the topological charge. Such long autocorrelations make the calculation at fine lattices inefficient adding extreme computational cost to the simple lattice volume scaling.

There have been many attempts to overcome the critical slowing down [2–13], and recent studies are presented thoroughly at the conference [14–29]. We in this work concentrate on developing the idea of the *trivializing map* proposed by Lüscher [7] (cf. the Nicolai map [30]), which maps a finite β theory to the $\beta = 0$ theory. Evidently, if we manage to construct such a map, one can obtain configurations on a fine lattice from the configurations of purely random $SU(3)$ fields.

In [7], Lüscher proved the existence of the map in the form of a gradient flow, and gave a way to construct the flow as a t-expansion (t: flow time) as we review in section 2. To the leading order in t, this flow corresponds to the Wilson flow for the plaquette action. Following this work, Engel and Schaefer [31] tested the method by implementing the leading order gradient flow in a CP^{N-1} model, whose result asserted its performance to be rather negative observing the scaling towards the continuum limit.

The aim of this work is to give an alternative way to construct an approximate trivializing map. In our method, the coefficients in the flow kernel are determined from lattice estimates of Wilson loops by using a Schwinger-Dyson equation. This method is versatile in the sense that: (1) The basis for the flow kernel can be chosen by hand. (2) It can be applied to the general action of interest without an analytical calculation. Here, the truncation effects and goodness of the flow can be evaluated by a force norm. We perform the HMC [32] with the effective action obtained by the Schwinger-Dyson method. The configuration generation part of the algorithm is the same as [7]. We apply our method to Wilson and DBW2 actions and show that we can have better control of the effective action than the t-expansion. Furthermore, in some cases, faster decorrelation (in Monte Carlo steps) is observed for long-ranged observables by adding the extended shapes such as rectangle and chair to the flow; however, the algorithmic overhead is still large and overwhelming the gain. This contribution therefore reports the preliminary results in this direction.

The rest of this paper is organized as follows. Section 2 is a review of the trivializing map including the t-expansion. In section 3, we first introduce a Schwinger-Dyson equation that determines the coefficients of Wilson loops in the effective action. We then describe its use in constructing an approximate trivializing map. Section 4 describes the results focusing on the norm measuring the closeness of the effective action to the target action and the autocorrelation of the smeared energy density. Section 5 is devoted for discussion.

2. Trivializing map

2.1 Review of Lüscher's construction

Following Lüscher [7], we consider a field transformation: $\mathcal{F} : V \mapsto U$, by which the original action $S(U)$ of interest will be mapped to the effective action $S_{\text{eff}}(V)$ according to the formula:

$$\begin{aligned}\int (dU)\, e^{-S(U)} &= \int (dV)\, \det\mathcal{F}_*(V)\, e^{-S(\mathcal{F}(V))} \\ &\equiv \int (dV)\, e^{-S_{\text{eff}}(V)},\end{aligned} \tag{1}$$

i.e.,

$$S_{\text{eff}}(V) \equiv S(\mathcal{F}(V)) - \ln\det\mathcal{F}_*(V). \tag{2}$$

The Jacobian matrix $\mathcal{F}_*(V)$ can be defined by the local parameterization $\theta^a_{x,\mu}$ of the field space:

$$e^{\theta^a_{x,\mu}T^a}U_{x,\mu}, \tag{3}$$

where T^a are the $\mathfrak{su}(3)$ generators normalized as $\mathrm{tr}\, T^aT^b = -(1/2)\delta^{ab}$. Note that the Haar measure can then be written as:

$$(dU) = \text{const.} \prod_{x,\mu,a} d\theta^a_{x,\mu}. \tag{4}$$

The derivative $\partial^a_{x,\mu} \equiv \partial_{\theta^a_{x,\mu}}|_{\theta=0}$ acts as

$$\partial^a_{x,\mu}U_{y,\nu} = \delta_{x,y}\delta_{\mu,\nu}T^aU_{x,\mu}. \tag{5}$$

Writing the indices in a short way as $A \equiv (x,\mu,a)$, the Jacobian matrix $\mathcal{F}_*(V) \equiv (\mathcal{F}_*^{AB}(V))$ relates the one-forms of the U-space and the V-space:

$$d\theta^A_U \equiv \mathcal{F}_*^{AB}(V)d\theta^B_V, \tag{6}$$

where the subscript is supplied to distinguish the coordinates of the two field spaces.

Lüscher obtained the *trivializing map* by considering a flow $\mathcal{F}_t : V \mapsto U_t \equiv \mathcal{F}_t(V)$ and by demanding the effective action at flow time t to be:

$$S_{\text{eff},t}(V) \overset{*}{=} (1-t)S(\mathcal{F}_t(V)), \tag{7}$$

up to an irrelevant t-dependent constant which we ignore hereafter. The symbol $\overset{*}{=}$ means it is a relation we require so that the V-space action becomes trivial at $t = 1$. We divide the trivializing map $\mathcal{F}_{t=1}$ into infinitesimal steps:

$$\mathcal{F}_{t=1} = \mathcal{F}_{(n-1)\epsilon,\epsilon} \circ \mathcal{F}_{(n-2)\epsilon,\epsilon} \circ \cdots \circ \mathcal{F}_{0,\epsilon}, \tag{8}$$

where $1 = n\epsilon$, and use the gradient flow ansatz:

$$\mathcal{F}_{t,\epsilon}(U)_{x,\mu} = e^{-\epsilon T^a\partial^a_{x,\mu}\tilde{S}_t(U)}U_{x,\mu}. \tag{9}$$

The derivative acts on the direct argument of the function.[1] Using that, for $U' \equiv \mathcal{F}_{t,\epsilon}(U)$, $d = d\theta^A_{U'}\partial^A_{U'} = d\theta^A_U \partial^A_U$ and

$$d\theta_{U'}{}^a_{x,\mu} = -2\mathrm{tr}[T^a dU'_{x,\mu} U'^{-1}_{x,\mu}], \tag{10}$$

one finds for an infinitesimal ε,

$$\mathcal{F}_{t,\epsilon,*}{}^a_{x,\mu}|^b_{y,\nu} = \delta_{x,y}\delta_{\mu,\nu}\delta^{ab} - \epsilon\partial^a_{x,\mu}\partial^b_{y,\nu}\tilde{S}_t - \epsilon\delta_{x,y}\delta_{\mu,\nu}f^{abc}\partial^c_{x,\mu}\tilde{S}_t, \tag{11}$$

where $f^{abc} \equiv -2\mathrm{tr}\big(T^a[T^b, T^c]\big)$ is the totally antisymmetric tensor. Therefore,

$$\ln\det \mathcal{F}_{t,\epsilon,*} = -\epsilon(\partial^A)^2\tilde{S}_t. \tag{12}$$

We note that the ansatz (9) is equivalent to:

$$\dot{U}_{t,x,\mu} = \frac{d}{dt}\mathcal{F}_t(V)_{x,\mu} = -T^a\partial^a_{x,\mu}\tilde{S}_t(U_t)\,U_{t,x,\mu}. \tag{13}$$

Rewriting eq. (7) as

$$\ln\det \mathcal{F}_{t,*}(V) \overset{*}{=} -tS(\mathcal{F}_t(V)), \tag{14}$$

and by taking the t-derivative, we obtain the functional equation:

$$[-(\partial^A)^2 + t(\partial^A S)\partial^A]\tilde{S}_t \overset{*}{=} S. \tag{15}$$

Since the differential operator $\mathcal{L}_t \equiv -(\partial^A)^2 + t(\partial^A S)\partial^A$ is elliptic and there is an inner product that makes $\mathcal{L}_t$ a symmetric operator, it is invertible for finite lattice up to the constant function, which is the zero-mode [7]. This fact suggests that the trivializing map formally exists.

Lüscher further gave a way to construct $\mathcal{F}_{t=1}$ as a t-expansion, which was explicitly demonstrated for the Wilson plaquette action:

$$S_W = -\frac{\beta}{6}W_0, \tag{16}$$

where W_0 is the sum of plaquettes (see figure 1 for a graphical representation). The construction begins with expanding $\tilde{S}_t$ as

$$\tilde{S}_t = \sum_{k\geq 0} t^k \tilde{S}^{(k)}, \tag{17}$$

whose convergence radius is proven to be finite for finite lattice [7]. Plugging the expansion into eq. (15), we obtain the recurrence relation:

$$-(\partial^A)^2\tilde{S}^{(0)} \overset{*}{=} S_W, \tag{18}$$

$$-(\partial^A)^2\tilde{S}^{(k)} \overset{*}{=} -(\partial^A S_W)(\partial^A\tilde{S}^{(k-1)}) \quad (k \geq 1). \tag{19}$$

[1] Therefore, the expression $\partial^a_{x,\mu}\tilde{S}_t(U)$ can be thought of as a vector function of U: $\partial^a_{x,\mu}\tilde{S}_t(U) = (\partial^a_{x,\mu}\tilde{S}_t)(U)$. We use this convention throughout this contribution.

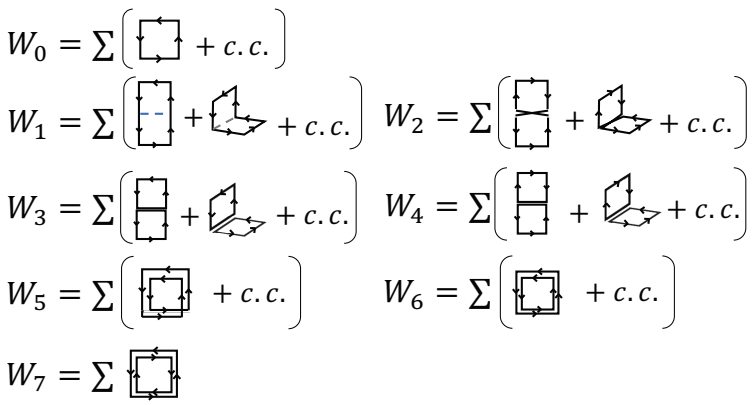

Figure 1: Wilson loops that are relevant in the argument.

These relations will be solved in the space of Wilson loops. Note that the derivative operator $\partial^a_{x,\mu}$ inserts T^a before $U_{x,\mu}$ or inserts $-T^a$ after $U^\dagger_{x,\mu}$. The contraction will then be calculated by the completeness relation; for complex 3×3 matrices A and B,

$$\mathrm{tr}[T^a A T^a B] = -\frac{1}{2}\Big(\mathrm{tr}A\mathrm{tr}B - \frac{1}{3}\mathrm{tr}AB\Big), \tag{20}$$

$$\mathrm{tr}[T^a A]\mathrm{tr}[T^a B] = -\frac{1}{2}\Big(\mathrm{tr}AB - \frac{1}{3}\mathrm{tr}A\mathrm{tr}B\Big). \tag{21}$$

See figure 2 for examples shown graphically. Since W_0 is an eigenfunction of $-(\partial^A)^2$, we have for

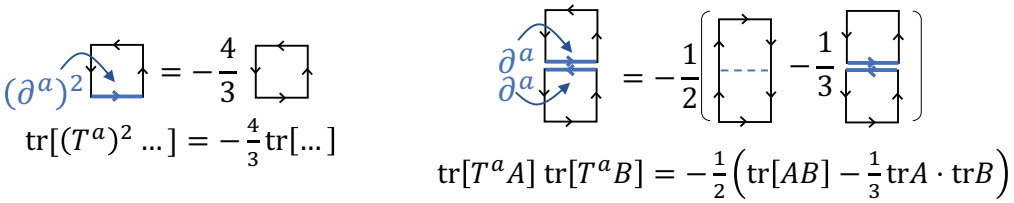

Figure 2: Two representative actions of the differential operators: (Left) The plaquette is an eigenfunction of the operator $(\partial^a)^2$. (Right) $\partial^a\cdot\partial^a\cdot$ glues Wilson loops with a trace subtraction.

the leading order:

$$\tilde{S}^{(0)} = -\frac{\beta}{32}W_0. \tag{22}$$

To obtain the next-to-leading order, we need the Wilson loops $W_0,\cdots,W_7$ shown in figure 1. The operator $-(\partial^A)^2$ can be represented in this subspace, and after inverting the matrix we get

$$\tilde{S}^{(1)} = \frac{\beta^2}{192}\Big(-\frac{4}{33}W_1 + \frac{12}{119}W_2 + \frac{1}{33}W_3 - \frac{5}{119}W_4 + \frac{3}{10}W_5 - \frac{1}{5}W_6 + \frac{1}{9}W_7\Big). \tag{23}$$

Since the operation $\partial^A\cdot\partial^A\cdot$ glues the Wilson loops in all possible ways, the number of relevant Wilson loops increases for the higher orders in a combinatorial manner.

2.2 Another look at the map from the space of effective actions

The reason for the rapid increase of the involved Wilson loop can be understood as follows. First note that, as can be seen in eq. (13), the flow runs from the trivial theory to a nontrivial theory. Correspondingly, the t-expansion around $t = 0$ is an expansion around the trivial theory. At small flow times, we only need to add the plaquettes in the effective action, and thus the expansion begins with W_0. However, as we evolve the flow, the effective action goes through the non-trivial theories, at which the wave-like particle picture should become relevant. To decrease these modes directly, we expect that the extended Wilson loops (presumably summed in certain linear combinations) become relevant. Therefore, obtaining the exact kernel for the trivializing map is extremely difficult, and thus we are forced for practical reasons to choose a finite basis and construct the efficient flow kernel that decreases the autocorrelations within the chosen subspace.

For this purpose, we reexamine the trivializing map from a different point of view, namely, from the effective action. As we demonstrate momentarily, while the flow in configuration space, eq. (13), evolves from the trivial theory to the finite β theory, in the action perspective the flow time runs in the opposite direction, from the finite β theory to the trivial theory. We note that the difference is just the way to see the map, and, if solved exactly, the two maps coincide. The difference, however, can occur when we approximate the map, and such a functional space point of view allows us to construct the map without using the t-expansion, whose example is the Schwinger-Dyson method we give in section 3. To distinguish the maps in the two different viewpoints, we add the superscript (c) to the expressions derived in the previous configuration space viewpoint.

We define the effective action at time t, S_t (we drop the subscript eff for notational simplicity), by the following recurrence relation:

$$S_{t=0}(V) \equiv S_W(V), \tag{24}$$

$$S_{t+\epsilon}(V) \equiv S_t(\mathcal{F}_{t,\epsilon}(V)) - \ln\det \mathcal{F}_{t,\epsilon,*}(V), \tag{25}$$

where $\mathcal{F}_{t,\epsilon}$ is again the map that increases t by an amount ϵ, and $\mathcal{F}_{t,\epsilon,*}(V)$ is its Jacobian matrix. We here require that $\mathcal{F}_{t,\epsilon}$ satisfies the relation:

$$S_{t+\epsilon}(V) \overset{*}{=} S_t(V) - \epsilon S_W(V), \tag{26}$$

up to an irrelevant constant which we again ignore. Using eqs. (24) and (26), we obtain:

$$S_t(V) \overset{*}{=} (1-t) S_W(V), \tag{27}$$

and thus we have the trivializing map at $t = 1$.

Note that the composition ordering of $\mathcal{F}_{t,\epsilon}$ will be the opposite from eq. (8). In fact, from the

recurrence formula, we have

$$\begin{aligned} S_{t=n\epsilon}(V) &= S_{(n-1)\epsilon}(\mathcal{F}_{(n-1)\epsilon,\epsilon}(V)) - \ln\det \mathcal{F}_{(n-1)\epsilon,\epsilon,*}(V) \\ &= S_{(n-2)\epsilon}(\mathcal{F}_{(n-2)\epsilon,\epsilon} \circ \mathcal{F}_{(n-1)\epsilon,\epsilon}(V)) \\ &\quad - \ln\det \mathcal{F}_{(n-2)\epsilon,\epsilon,*}(\mathcal{F}_{(n-1)\epsilon,\epsilon}(V)) - \ln\det \mathcal{F}_{(n-1)\epsilon,\epsilon,*}(V) \\ &= \cdots \\ &= S_0(\mathcal{F}_{0,\epsilon} \circ \cdots \circ \mathcal{F}_{(n-1)\epsilon,\epsilon}(V)) \\ &\quad - \sum_{\ell=0}^{n-1} \ln\det \mathcal{F}_{\ell\epsilon,\epsilon,*}(\mathcal{F}_{(\ell+1)\epsilon,\epsilon} \circ \cdots \circ \mathcal{F}_{(n-1)\epsilon,\epsilon}(V)). \end{aligned} \tag{28}$$

We identify $\mathcal{F}_t$ with the composite function:

$$\mathcal{F}_t \equiv \mathcal{F}_{0,\epsilon} \circ \cdots \circ \mathcal{F}_{(n-1)\epsilon,\epsilon}, \tag{29}$$

where the Jacobian matrix is given by

$$\mathcal{F}_{t,*}(V) \equiv \prod_{\ell=0}^{n-1} \mathcal{F}_{\ell\epsilon,\epsilon,*}(\mathcal{F}_{(\ell+1)\epsilon,\epsilon} \circ \cdots \circ \mathcal{F}_{(n-1)\epsilon,\epsilon}(V)). \tag{30}$$

The matrix product is taken in descending order from right to left. It is easy to see that S_t is indeed the effective action for the map $\mathcal{F}_t$:

$$S_t(V) = S_W(\mathcal{F}_t(V)) - \ln\det \mathcal{F}_{t,*}(V). \tag{31}$$

To further rewrite the expression, we again use the gradient flow ansatz:

$$\mathcal{F}_{t,\epsilon}(V)_{x,\mu} = e^{-\epsilon T^a \partial^a_{x,\mu} \tilde{S}_t(V)} V_{x,\mu}. \tag{32}$$

Then the requirement (26) gives the functional equation for $\tilde{S}_t$:

$$[-(\partial^A)^2 + (1-t)(\partial^A S_W)\partial^A]\tilde{S}_t \stackrel{*}{=} S_W. \tag{33}$$

Comparing eq. (33) with (15), we notice that

$$1 - t = t^{(c)}, \tag{34}$$

and thus

$$\tilde{S}_t = \tilde{S}^{(c)}_{1-t} \tag{35}$$

and

$$\mathcal{F}_{t,\epsilon} = \mathcal{F}^{(c)}_{1-t,\epsilon}. \tag{36}$$

From eqs. (29) and (36), the map $\mathcal{F}_t$ in the functional space viewpoint can be expressed with $\mathcal{F}^{(c)}_{t,\epsilon}$ and $1 = n\epsilon$ as:

$$\mathcal{F}_{t=m\epsilon} = \mathcal{F}^{(c)}_{n\epsilon,\epsilon} \circ \cdots \circ \mathcal{F}^{(c)}_{(n-m+1)\epsilon,\epsilon}. \tag{37}$$

In particular,

$$\mathcal{F}_{1=n\epsilon} = \mathcal{F}^{(c)}_{n\epsilon,\epsilon} \circ \cdots \circ \mathcal{F}^{(c)}_{\epsilon,\epsilon}, \tag{38}$$

$$\mathcal{F}^{(c)}_{1=n\epsilon} = \mathcal{F}^{(c)}_{(n-1)\epsilon,\epsilon} \circ \cdots \circ \mathcal{F}^{(c)}_{0,\epsilon}, \tag{39}$$

which should agree in $\epsilon \to 0$.

We therefore have the same trivializing map, but the direction of the construction is from finite β theory to the trivial theory (see figure 3). The complication is that it is not straightforward to

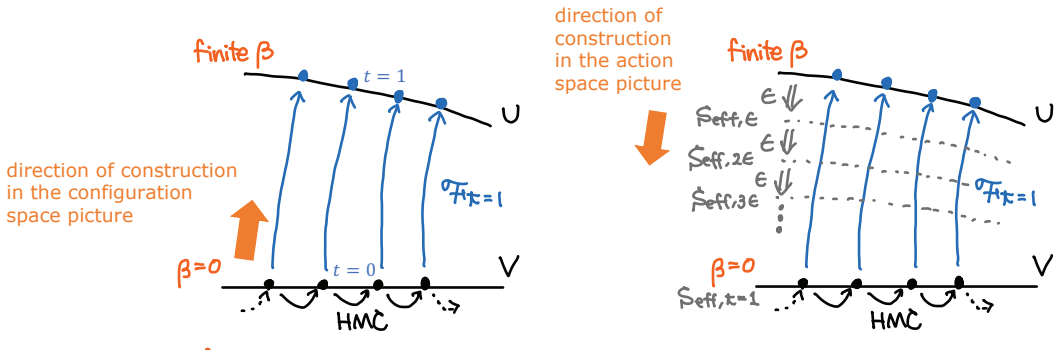

Figure 3: Direction of constructing trivializing maps.

expand in terms of t, as expected from the argument at the beginning of this section. We will solve eq. (33) by picking up a finite basis, for which we need a way to project the entire function space to the truncated space. For this purpose, we use a Schwinger-Dyson equation.

3. Schwinger-Dyson equation

In the Schwinger-Dyson method described below, we sequentially determine the flow from the finite β theory as explained in section 2.2. At each intermediate step, we determine the effective couplings by a Schwinger-Dyson equation. The flow will be designed to decrease the couplings from the lattice expectation values of Wilson loops.

3.1 Schwinger-Dyson equation for determining coupling constants

In this subsection, we review the use of a Schwinger-Dyson equation to determine the couplings in effective actions, based on refs [33, 34] (see also [35, 36]). Suppose that the effective action S_{eff} is expressed as a sum of Wilson loops (and their products):

$$S_{\text{eff}} = \sum_j \beta_j W_j. \tag{40}$$

Then, the coefficients β_j satisfy the linear equation:

$$\sum_j \beta_j \langle \partial^A W_j \partial^A W_i \rangle_{S_{\text{eff}}} = \langle (\partial^A)^2 W_i \rangle_{S_{\text{eff}}}, \tag{41}$$

where the expectation value $\langle\cdot\rangle_{S_{\rm eff}}$ is taken with respect to the action $S_{\rm eff}$. In fact, under the variation δV using W_i in the gradient flow kernel:

$$\delta V_{x,\mu} \equiv -\epsilon \sum_{x,\mu} T^a (\partial^a_{x,\mu} W_i)\, V_{x,\mu}, \tag{42}$$

the path integral is invariant (we vary all the links (x,μ) at the same time):

$$0 = \delta \int (dV) e^{-S_{\rm eff}(V)} = \epsilon \int (dV) e^{-S_{\rm eff}(V)} [-(\partial^A)^2 W_i + (\partial^A S_{\rm eff}) \partial^A W_i]. \tag{43}$$

Combining eqs. (40) and (43), we obtain eq. (41).

The linear equation (41) allows us to obtain the couplings β_j from the lattice expectation values. However, we generically need an infinite number of basis functions to fully parameterize the action, and thus $\langle \partial^A W_j \partial^A W_i \rangle_{S_{\rm eff}}$ in eq. (41) becomes an infinite-dimensional matrix. Practically, we cannot calculate an infinite number of expectation values, and we need to introduce a truncation. By choosing a finite basis $\{W_{j'}\}$, where now j' only runs a finite range, we approximate the action as

$$S_{\rm eff} \approx {\sum_{j'}}' \beta'_{j'} W_{j'} \equiv S'_{\rm eff}. \tag{44}$$

The prime symbols indicate the truncation. It then turns out that β'_j determined by the finite-dimensional counterpart of eq. (41):

$${\sum_{j'}}' \beta'_{j'} \langle \partial^A W_{j'} \partial^A W_{i'} \rangle_{S_{\rm eff}} = \langle (\partial^A)^2 W_{i'} \rangle_{S_{\rm eff}}, \tag{45}$$

gives the best approximation of $S_{\rm eff}$ in the sense that they minimize the norm $\|S_{\rm eff} - S'_{\rm eff}\|_{S_{\rm eff}}$, where $\|\cdot\|_{S_{\rm eff}}$ is the force norm:

$$\|S\|^2_{S_{\rm eff}} \equiv \langle (\partial^A S)^2 \rangle_{S_{\rm eff}}. \tag{46}$$

In fact, by subtracting eqs. (41) and (45) for the range of the index i':

$${\sum_{j'}}' (\beta_{j'} - \beta'_{j'}) \langle \partial^A W_{j'} \partial^A W_{i'} \rangle_{S_{\rm eff}} = 0. \tag{47}$$

This is equivalent to the stationary condition:

$$\begin{aligned} 0 &= \frac{\partial}{\partial \beta'_{i'}} \|S_{\rm eff} - S'_{\rm eff}\|^2_{\rm eff} \\ &= \frac{\partial}{\partial \beta'_{i'}} \langle [\partial^A (S_{\rm eff} - S'_{\rm eff})]^2 \rangle_{\rm eff} \\ &= -2 {\sum_{j'}}' (\beta_{j'} - \beta'_{j'}) \langle \partial^A W_{j'} \partial^A W_{i'} \rangle_{S_{\rm eff}}. \end{aligned} \tag{48}$$

Since the norm (46) is bounded from below but not from above, we can generically expect the stationary point to be the minimum. Therefore, this Schwinger-Dyson method gives us a systematic way to project effective actions onto a subspace of actions, where we can measure its goodness from the force fields in numerical calculations.

3.2 Designing the kernel with the Schwinger-Dyson equation

We now apply the Schwinger-Dyson equation in the determination of the approximate trivializing map. The idea is to reduce the action with t as in eq. (26), replacing the effective action with the approximate one, eq. (44), so that we can work completely in the functional space spanned by the finite basis. Below, we write down equations in this finite basis, and we drop the prime symbols on j for notational simplicity. We write as $\mathcal{S}$ the subspace spanned by the chosen finite set $\{W_j\}$ of Wilson loops and their products.

For the effective action $S_{\text{eff},t}$ at a given time t, we apply the Schwinger-Dyson method to approximate it in $\mathcal{S}$:

$$S_{\text{eff},t} \approx S'_{\text{eff},t} = \sum_j{}' \beta'_{j,t} W_j. \tag{49}$$

Here the effective couplings $\beta_{j,t}$ are determined by the linear equation:

$$\Big\langle -(\partial^A)^2 W_i + \sum_j{}' \beta'_{j,t} \partial^A W_j \partial^A W_i \Big\rangle_{S_{\text{eff},t}} = 0, \tag{50}$$

corresponding to eq. (45). We again consider the infinitesimal map of the form:

$$\mathcal{F}_{t,\epsilon}(V)_{x,\mu} = e^{-\epsilon T^a \partial^a_{x,\mu} \tilde{S}_t(V)} V_{x,\mu}, \tag{51}$$

where this time we construct $\tilde{S}_t$ in the truncated space $\mathcal{S}$:

$$\tilde{S}_t = \sum_k{}' \gamma_{k,t} W_k. \tag{52}$$

By taking d/dt in eq. (50), we obtain the linear equation:

$$\sum_k{}' \gamma_{k,t} \Big\langle \partial^B W_k \partial^B \Big[-(\partial^A)^2 W_i + \sum_j{}' \beta'_{j,t} \partial^A W_j \partial^A W_i \Big] \Big\rangle_{S_{\text{eff},t}} = -\sum_j{}' \dot{\beta}'_{j,t} \Big\langle \partial^A W_j \partial^A W_i \Big\rangle_{S_{\text{eff},t}}. \tag{53}$$

The left hand side comes from the t-dependence of the Boltzmann weight, and the right hand side from the explicit t-dependence of $\beta'_{j,t}$. Equation (53) gives us the coefficients $\gamma_{k,t}$ for a given $\dot{\beta}'_{j,t}$, i.e., for a given trajectory of $S'_{\text{eff},t}$ in $\mathcal{S}$. We here choose $\dot{\beta}'_{j,t}$ to be:

$$\dot{\beta}'_{j,t} = -\frac{\beta'_{j,t}}{1-t}, \tag{54}$$

so that $\beta'_{j,t} = (1-t)\beta'_{j,t=0} = \beta_{j,t=0}$.[2] After obtaining the coefficients $\gamma_{k,t}$, and thus the map (51) at each time step $t = k\epsilon$, $(k = 0, \cdots, n-1)$, we compose them in the ordering:

$$\mathcal{F}_t \equiv \mathcal{F}_{0,\epsilon} \circ \cdots \circ \mathcal{F}_{(n-1)\epsilon,\epsilon}, \tag{55}$$

[2] Another choice of $\dot{\beta}'_{j,t}$ leading to the same functional form of $\beta_{j,t}$ is $\dot{\beta}'_{j,t} = -\beta_{j,t=0}$. The difference between the two appears in practical calculation where we have inexactness due to $O(\epsilon^2)$ terms and the statistical error. We choose the form (54) as it can take into account of these effects from the observed $\beta'_{j,t}$. It is, however, uncertain if this is the optimal choice though the difference is small compared to the truncation effect in choosing $\mathcal{S}$.

One of the advantages of designing the trivializing map with the Schwinger-Dyson equation is that the basis for the flow kernel can be chosen arbitrarily by hand. Furthermore, it can be applied to the general action of interest, and the coefficients in the kernel, $\gamma_{j,t}$, are determined by the lattice estimates of the observables. In this sense, it contains the non-perturbative information of the theory and there is no need to make analytic calculations beforehand. Note also that the truncation effects and goodness of the map can be measured by the force norm (46).

There is a practical note in the choice of $\mathcal{S}$. Since there are linear relations among Wilson loops and their products called Mandelstam constraints [37] (see also [38–40]), we need to choose the basis functions carefully so that the inversions (45) and (53) are possible. The simplest and relevant example is (see Fig 4 for a graphical representation):

$$W_6 = W_5 + 2W_0. \tag{56}$$

This relation follows from a simple $SU(3)$ identity: For $U \in \mathrm{SU}(3)$,

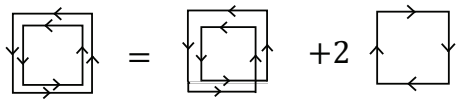

Figure 4: Simplest Mandelstam constraint.

$$(\mathrm{tr}\, U)^2 = \mathrm{tr}(U^2) + 2\mathrm{tr}\, U^\dagger. \tag{57}$$

Further relation can be obtained by the Cayley–Hamilton equation:

$$U^3 = (\mathrm{tr}U)U^2 - \frac{1}{2}\left[(\mathrm{tr}U)^2 - \mathrm{tr}(U^2)\right]U + 1. \tag{58}$$

4. Results

In this section, we show the evolution of the effective action and autocorrelation times for the Schwinger-Dyson method. We use the HMC algorithm with the exact transformed action, whose detail was given by Lüscher in [7]. At each step of the approximate trivializing map, we have the inversion, eq. (53), for which we use the numerical derivative with the five-point formula to calculate the matrix from the flows with $\epsilon = 0.0004$.

4.1 Evolution of the effective action

To compare the effective actions determined by the Schwinger-Dyson method to those by the t-expansion, we use the Wilson action. We use 8^4 lattice with $\beta = 6.13$, which corresponds to $a^{-1} = 2.56$GeV [41]. In figure 5, we show one of the determined coefficients, $\gamma_{0,t}$, which is the coefficient of W_0. We take the step size $\epsilon = 0.1$ and consider $t = 0.1, \cdots, 0.4$. We see from the figure that the coefficients $\gamma_{j,t}$ determined by the Schwinger-Dyson method differ significantly from the t-expansion. It is also notable that $\gamma_{j,t}$ largely depends on the choice of $\mathcal{S}$. In figure 6, we plot the norm (46) between $S_{\mathrm{eff},t}$ and the target action $(1-t)S$ at each time t. We see that, with the Schwinger-Dyson method, we can have better control of the effective action. Note that, the norm

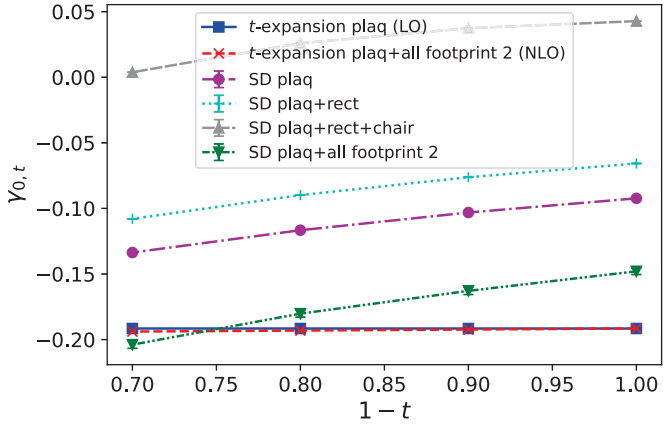

Figure 5: Comparison of $\gamma_{0,t}$ from the t-expansion method in the leading order (LO) and the next-to-leading order (NLO) to the Schwinger-Dyson (SD) method with various choices of $\mathcal{S}$.

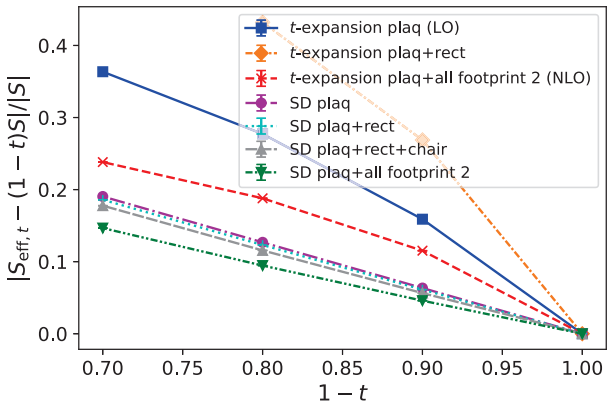

Figure 6: The difference between the effective action $S_{\mathrm{eff},t}$ and the target action $(1-t)S$.

monotonically decreases when we enlarge the truncated space $\mathcal{S}$ though the gain is small compared to the deviation from the exact trivializing trajectory. We lastly comment that, as shown in figure 6, naively adding the rectangle term in the lowest order t-expansion formula (orange dash-dotted line) does not improve the map, but rather makes it worse; we need to add all the terms in the next order expansion to improve the map in the t-expansion method.

4.2 Autocorrelation

To study autocorrelation and the efficiency of the algorithm, we switch to the DBW2 action [42]. We take $8^3 \times 16$ lattice with $\beta = 0.89$, which corresponds to $a^{-1} = 1.49$GeV [43]. In figure 7, we plot the difference from the target action, which is the counterpart of figure 6 in the DBW2 case. We choose the step size to be $\epsilon = 0.2$ and take $t = 0, 0.2, \cdots, 1.0$. We see that, also for the DBW2 action, we have better control of the map by enlarging the space $\mathcal{S}$.

In figure 8, we show the history of the smeared energy density and the normalized autocorre-

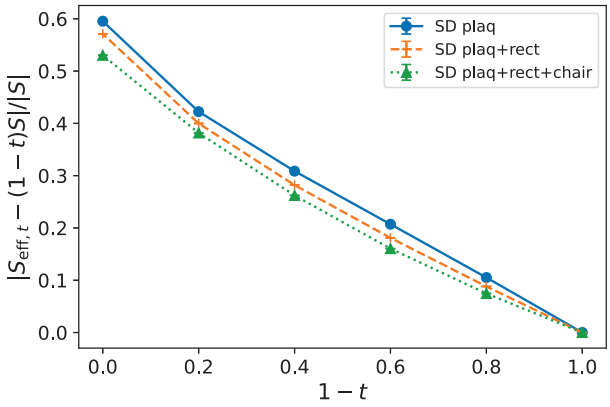

Figure 7: The difference between $S_{\mathrm{eff},t}$ and the target $(1-t)S$ for the DBW2 case.

lation function ρ_n calculated from the data:

$$\rho(n) \equiv C(n)/C(0), \quad C(n) \equiv \langle (E_m - \langle E_m \rangle)(E_{m+n} - \langle E_{m+n} \rangle) \rangle. \tag{59}$$

The smearing is performed with the Wilson flow with the flow time $t_w = 30t_0$, where t_0 is the time

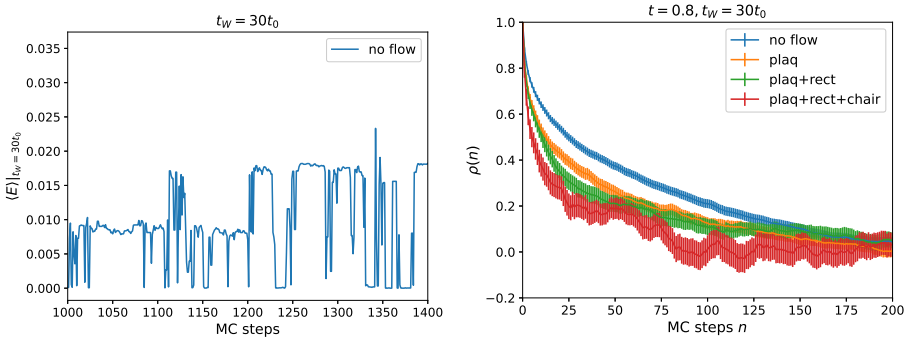

Figure 8: Autocorrelation of the smeared energy density with the Wilson flow time $t_w = 30t_0$. (Left) History of the observable generated by the ordinary (not field-transformed) HMC. (Right) The normalized autocorrelation function $\rho(n)$ with and without applying the approximated trivializing map determined by the Schwinger-Dyson method.

scale at which [44]

$$t_w^2 \langle E \rangle = 0.3. \tag{60}$$

After the smearing time $t_w = 30t_0$, the smeared energy density reflects the instantons (see the left panel of figure 8). The right panel of figure 8 shows that the configurations decorrelate faster (in Monte Carlo steps) by including the extended loops.

However, looking at the observable at the other scales, we notice that the autocorrelation is not controlled completely though we still observe the tendency of improvement by the extended loops. In figure 9, we show the autocorrelation of the energy densities at smearing times $t_w = t_0, 10t_0, 30t_0$. As we can see from, e.g., the $t = 0.6, t_w = 10t_0$ case, adding more Wilson loops in $\tilde{S}_t$ does not always reduce the autocorrelation.

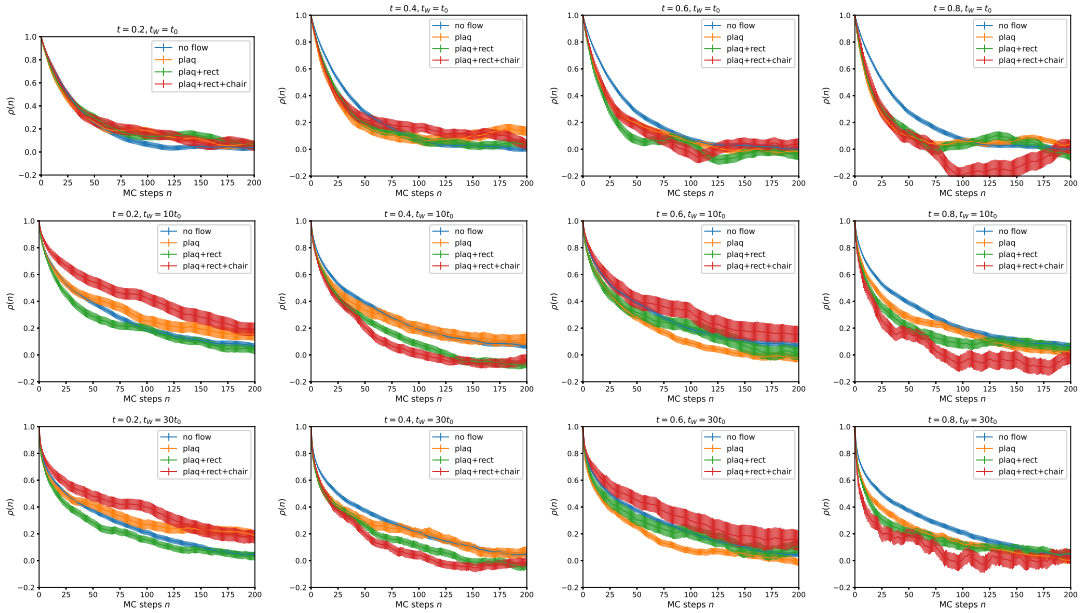

Figure 9: The autocorrelation function with various flow time scales

4.3 Algorithmic overhead

At this point, we also mention that the algorithmic overhead is not negligible when adding the extended Wilson loops. Figure 10 shows the computational cost for generating a single configuration with one-step flows with various choices of $\mathcal{S}$.[3] We see that the cost increases quite rapidly by

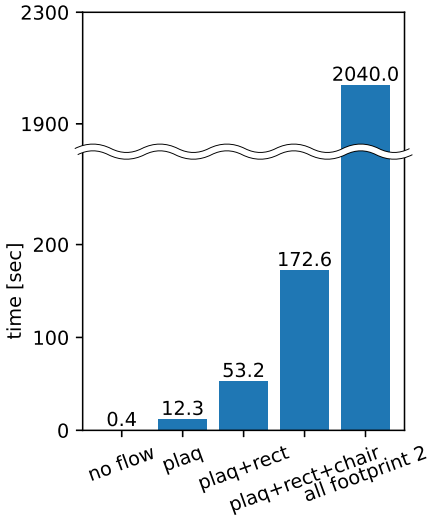

Figure 10: Computation time to generate configuration.

adding extended Wilson loops. The increase can be understood by the cost of calculating the hessian $\partial^A \partial^B \tilde{S}_t$, which is used in the calculation of the Jacobian $\mathcal{F}_{t,*}$. Figure 11 pictorially shows the cost

[3]This is a one-node calculation fully parallelized with MPI and OpenMP.

difference of calculating $\partial^A \partial^B W$ for $W = W_0$ (plaquette) and $W = W_1$ (rectangle). Here note that,

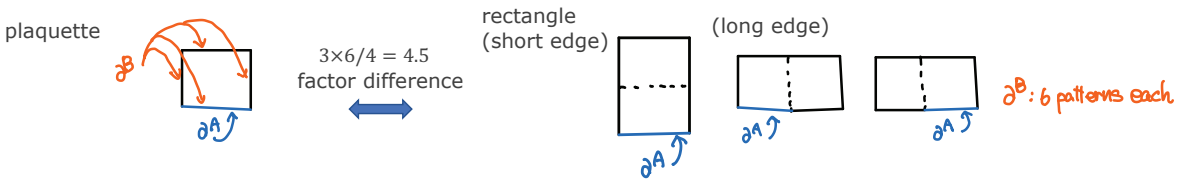

Figure 11: Numerical cost of calculating the Hessian $\partial^A \partial^B \tilde{S}_t$ for plaquette and rectangle.

since acting derivatives inserts the generator T^a, each pattern of acting the derivatives requires evaluation of Wilson loops with different insertions. By counting the number of ways to choose the links on which ∂^A and ∂^B act, we have a multiplicative factor of 4.5 in the cost for W_1 compared to W_0, which mostly agrees with the cost increase shown in figure 10. We also note that, in the above calculations, the flow is arranged to make the Jacobian calculation run in parallel for each link. Though for the extended shapes we need complicated schemes, it can be done by dividing the directions of the flowed links and by appropriately coloring the lattice for each type of loops [7, 45]. Figure 12 shows the examples of the coloring schemes.

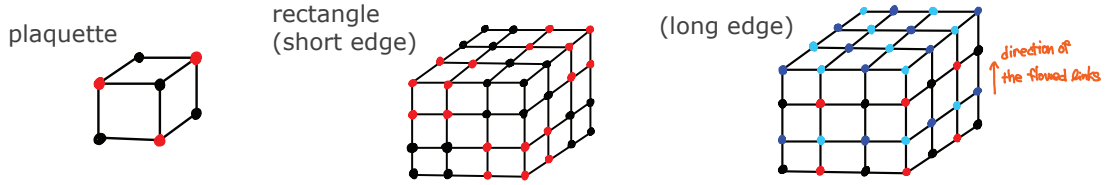

Figure 12: Coloring schemes.

5. Discussion

In this work, we proposed a way to design an approximate trivializing map using a Schwinger-Dyson equation. The algorithmic advantages of this method are that: (1) The basis for the flow kernel can be chosen arbitrarily by hand. (2) It can be applied to the general action of interest. (3) The coefficients in the kernel are determined by lattice estimates of the observables, which does not require analytic calculations such as those in the t-expansion. It is also notable that the truncation effects and goodness of the flow can be measured by the force norm. We showed that with the Schwinger-Dyson method, we can have better control of the effective action, and observed a tendency to decrease the autocorrelation of long-range observables by adding extended Wilson loops. However, the decorrelation is not sufficient compared to the induced overhead. This indicates that, though the effective action is getting closer to the target action, the large-size Wilson loops that are neglected in the construction still leave significant contributions.

Indeed, to obtain the exact trivializing map, in which case we can decrease the autocorrelation of all the modes of the system, we need an infinite number of Wilson loops as reviewed in section 2.1. However, adding more and more Wilson loop basis is not realistic because of the increasing

algorithmic overhead shown in subsection 4.3. Since it is likely that we need to restrict ourselves to a few small Wilson loops, one possible strategy may be to be more specific to a particular slow observable, e.g., the topological charge. For example, as is well known, the instanton potential with respect to its radius can be different for the same lattice spacing but with different gauge action. We thus may be able to change the instanton potential to stimulate tunneling with the small number of basis functions in $\mathcal{S}$ (cf. [46–49]). Studies along these lines are in progress and will be reported elsewhere.

6. Acknowledgments

The authors thank Norman Christ, William Detmold, Sam Foreman, Xiao-Yong Jin, Gurtej Kanwar, James Osborn, and Phiala Shanahan for valuable discussions. Computation was performed in RIKEN HOKUSAI, Univ of Tokyo Oakforest-PACS. We have also used a USQCD facility at BNL (KNL), which is funded by the Office of Science of the U.S. Department of Energy. AT is supported by JSPS KAKENHI Grant Numbers JP20K14479, JP22H05112, JP22H05111, and JP22K03539. NM is supported by JP22H01222 and the Special Postdoctoral Researchers Program of RIKEN.

References

[1] C. Lehner, *RBC/UKQCD update of the HVP contribution to the muon g-2*, *PoS* **LATTICE2022** (2022) 301.

[2] G. Parisi, *Prolegomena to Any Future Computer Evaluation of the QCD Mass Spectrum*, Progress in Gauge Field Theory, edited by 't Hooft et al. (Plenum, New York, 1984), p.351.

[3] G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, B. Svetitsky and K. G. Wilson, *Langevin Simulations of Lattice Field Theories*, Phys. Rev. D **32**, 2736 (1985).

[4] C. T. H. Davies, G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, K. G. Wilson, P. Rossi and B. Svetitsky, *Fourier Acceleration in Lattice Gauge Theories. 1. Landau Gauge Fixing*, Phys. Rev. D **37**, 1581 (1988).

[5] G. Katz, G. Batrouni, C. Davies, A. S. Kronfeld, P. Lepage, P. Rossi, B. Svetitsky and K. Wilson, *Fourier Acceleration. 2. Matrix Inversion and the Quark Propagator*, Phys. Rev. D **37**, 1589 (1988).

[6] C. T. H. Davies, G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, P. Rossi, B. Svetitsky and K. G. Wilson, *Fourier Acceleration in Lattice Gauge Theories. 3. Updating Field Configurations*, Phys. Rev. D **41**, 1953 (1990).

[7] M. Luscher, *Trivializing maps, the Wilson flow and the HMC algorithm*, Commun. Math. Phys. **293**, 899-919 (2010) [arXiv:0907.5491 [hep-lat]].

[8] M. Luscher and S. Schaefer, *Lattice QCD without topology barriers*, JHEP **07**, 036 (2011) [arXiv:1105.4749 [hep-lat]].

[9] M. S. Albergo, G. Kanwar and P. E. Shanahan, *Flow-based generative models for Markov chain Monte Carlo in lattice field theory*, Phys. Rev. D **100**, no.3, 034515 (2019) [arXiv:1904.12072 [hep-lat]].

[10] S. Foreman, X. Y. Jin and J. C. Osborn, *Deep Learning Hamiltonian Monte Carlo*, [arXiv:2105.03418 [hep-lat]].

[11] D. Albandea, P. Hernández, A. Ramos and F. Romero-López, *Topological sampling through windings*, Eur. Phys. J. C **81**, no.10, 873 (2021) [arXiv:2106.14234 [hep-lat]].

[12] S. Foreman, T. Izubuchi, L. Jin, X. Y. Jin, J. C. Osborn and A. Tomiya, *HMC with Normalizing Flows*, PoS **LATTICE2021**, 073 (2022) [arXiv:2112.01586 [cs.LG]].

[13] T. Nguyen, P. Boyle, N. H. Christ, Y. C. Jang and C. Jung, *Riemannian Manifold Hybrid Monte Carlo in Lattice QCD*, PoS **LATTICE2021**, 582 (2022) [arXiv:2112.04556 [hep-lat]].

[14] J. Urban, *Flow-based density of states for complex actions*, *PoS* **LATTICE2022** (2022) 041.

[15] M. Caselle, E. Cellini, A. Nada and M. Panero, *Stochastic normalizing flows for lattice field theory*, *PoS* **LATTICE2022** (2022) 005 [arXiv:2210.03139 [hep-lat]].

[16] F. Romeo-Lopez, *Gauge-equivariant flow models for sampling in lattice field theories with pseudofermions*, *PoS* **LATTICE2022** (2022) 033.

[17] P. Shanahan, *Status update on flow models for gauge field generation*, *PoS* **LATTICE2022** (2022) 036.

[18] J. M. Rossney, *Machine Learning Trivializing Maps*, *PoS* **LATTICE2022** (2022) 024.

[19] D. Albandea, L. Del Debbio, P. Hernández, R. Kenway, J. M. Rossney and A. Ramos, *Learning trivializing flows*, *PoS* **LATTICE2022** (2022) 001 [arXiv:2211.12806 [hep-lat]].

[20] J. Komijani, *Generative models for scalar field theories: how to deal with poor scaling?*, *PoS* **LATTICE2022** (2022) 019.

[21] A. Francis, *Translating topological benefits in very cold master-field simulations*, *PoS* **LATTICE2022** (2022) 368.

[22] T. Eichhorn, C. Hoelbling, P. Rouenhoff and L. Varnhorst, *Topology changing update algorithms for SU(3) gauge theory*, *PoS* **LATTICE2022** (2022) 009 [arXiv:2210.11453 [hep-lat]].

[23] P. Rouenhoff, *Metadynamics Surfing on Topology Barriers in the Schwinger Model*, *PoS* **LATTICE2022** (2022) 253.

[24] J. Swaim, *Fourier acceleration in strongly-interacting linear sigma models*, *PoS* **LATTICE2022** (2022) 254.

[25] C. Bonanno, M. D'Elia, B. Lucini and D. Vadacchino, *Towards glueball masses of large-N SU(N) Yang-Mills theories without topological freezing via parallel tempering on boundary conditions*, *PoS* **LATTICE2022** (2022) 392 [arXiv:2210.07622 [hep-lat]].

[26] R. Kara, *Parallel tempering algorithm applied for the deconfinement transition of quenched QCD*, *PoS* **LATTICE2022** (2022) 178.

[27] F. D'Angelo, *Topological susceptibility in high temperature full QCD via staggered spectral projectors*, *PoS* **LATTICE2022** (2022) 393.

[28] J. C. Pinto Barros, *Towards the Application of Skewed Detailed Balance in Lattice Gauge Theories*, *PoS* **LATTICE2022** (2022) 028.

[29] J. Finkenrath, *Review on Algorithms for dynamical fermions*, *PoS* **LATTICE2022** (2022) 227.

[30] H. Nicolai, *Supersymmetry and Functional Integration Measures*, Nucl. Phys. B **176**, 419-428 (1980).

[31] G. P. Engel and S. Schaefer, *Testing trivializing maps in the Hybrid Monte Carlo algorithm*, Comput. Phys. Commun. **182**, 2107-2114 (2011) [arXiv:1102.1852 [hep-lat]].

[32] S. Duane, A. D. Kennedy, B. J. Pendleton and D. Roweth, *Hybrid Monte Carlo*, Phys. Lett. B **195**, 216-222 (1987).

[33] A. Gonzalez-Arroyo and M. Okawa, *Renormalized Coupling Constants by Monte Carlo Methods*, Phys. Rev. D **35**, 672 (1987).

[34] A. Gonzalez-Arroyo and M. Okawa, *Universality of Deconfining Phase Transition in Finite Temperature Lattice Gauge Theories*, Phys. Rev. Lett. **58**, 2165 (1987).

[35] A. Gonzalez-Arroyo, M. Okawa and Y. Shimizu, *Monte Carlo Renormalization Group Study of the Four-dimensional Z(2) Gauge Theory*, Phys. Rev. Lett. **60**, 487 (1988).

[36] P. de Forcrand *et al.* [QCD-TARO], *Renormalization group flow of SU(3) gauge theory*, [arXiv:hep-lat/9806008 [hep-lat]].

[37] S. Mandelstam, *Charge - Monopole Duality and the Phases of Nonabelian Gauge Theories*, Phys. Rev. D **19**, 2391 (1979).

[38] R. Giles, *The Reconstruction of Gauge Potentials From Wilson Loops*, Phys. Rev. D **24**, 2160 (1981).

[39] R. Loll, *Yang-Mills theory without Mandelstam constraints*, Nucl. Phys. B **400**, 126-144 (1993).

[40] N. J. Watson, *Solution of the SU(2) Mandelstam constraints*, Phys. Lett. B **323**, 385-392 (1994) [arXiv:hep-th/9311126 [hep-th]].

[41] M. Cè, C. Consonni, G. P. Engel and L. Giusti, *Non-Gaussianities in the topological charge distribution of the SU(3) Yang–Mills theory*, Phys. Rev. D **92**, no.7, 074502 (2015) [arXiv:1506.06052 [hep-lat]].

[42] P. de Forcrand *et al.* [QCD-TARO], *Search for effective lattice action of pure QCD*, Nucl. Phys. B Proc. Suppl. **53**, 938-941 (1997) [arXiv:hep-lat/9608094 [hep-lat]].

[43] S. Necco, *Universality and scaling behavior of RG gauge actions*, Nucl. Phys. B **683**, 137-167 (2004) [arXiv:hep-lat/0309017 [hep-lat]].

[44] M. Lüscher, *Properties and uses of the Wilson flow in lattice QCD*, JHEP **08**, 071 (2010) [erratum: JHEP **03**, 092 (2014)] [arXiv:1006.4518 [hep-lat]].

[45] D. Boyda, G. Kanwar, S. Racanière, D. J. Rezende, M. S. Albergo, K. Cranmer, D. C. Hackett and P. E. Shanahan, *Sampling using $SU(N)$ gauge equivariant flows*, Phys. Rev. D **103**, no.7, 074504 (2021) [arXiv:2008.05456 [hep-lat]].

[46] D. A. Smith *et al.* [UKQCD], *Topological structure of the SU(3) vacuum*, Phys. Rev. D **58**, 014505 (1998) [arXiv:hep-lat/9801008 [hep-lat]].

[47] T. A. DeGrand, A. Hasenfratz and T. Kovacs, *Structure of the QCD vacuum as seen by lattice simulations*, Prog. Theor. Phys. Suppl. **131**, 573-584 (1998) [arXiv:hep-lat/9801037 [hep-lat]].

[48] P. de Forcrand, M. Garcia Perez, J. E. Hetrick and I. O. Stamatescu, *Topological properties of the QCD vacuum at T = O and T similar to T(c)*, [arXiv:hep-lat/9802017 [hep-lat]].

[49] A. Hasenfratz and C. Nieter, *Instanton content of the SU(3) vacuum*, Phys. Lett. B **439**, 366-372 (1998) [arXiv:hep-lat/9806026 [hep-lat]].