Statistical analysis method for the worldvolume hybrid Monte Carlo algorithm

Masafumi Fukuma[1], Nobuyuki Matsumoto[2,⋆], and Yusuke Namekawa[3]

[1]*Department of Physics, Kyoto University, Kyoto 606-8502, Japan*
[2]*RIKEN/BNL Research Center, Brookhaven National Laboratory, Upton, NY 11973, USA*
[3]*Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan*
⋆E-mail: nobuyuki.matsumoto@riken.jp

....

We discuss the statistical analysis method for the worldvolume hybrid Monte Carlo (WV-HMC) algorithm [M. Fukuma and N. Matsumoto, Prog. Theor. Exp. Phys. **2021**, 023B08 (2021)], which was recently introduced to substantially reduce the computational cost of the tempered Lefschetz thimble method. In the WV-HMC algorithm, the configuration space is a continuous accumulation (worldvolume) of deformed integration surfaces, and sample averages are considered for various subregions in the worldvolume. We prove that, if a sample in the worldvolume is generated as a Markov chain, then the subsample in the subregion can also be regarded as a Markov chain. This ensures the application of the standard statistical techniques to the WV-HMC algorithm. We particularly investigate the autocorrelation times for the Markov chains in various subregions, and find that there is a linear relation between the probability of being in a subregion and the autocorrelation time for the corresponding subsample. We numerically confirm this scaling law for a chiral random matrix model.

....

Subject Index   A20, A22, A23, B01, B38

1.   Introduction

The numerical sign problem has prevented us from the first-principles analysis of various important systems, such as quantum chromodynamics (QCD) at finite density [1], quantum Monte Carlo calculations of quantum statistical systems [2], and the real-time dynamics of quantum fields.

Among various approaches to the sign problem, some utilize the complexification of dynamical variables. For example, in the complex Langevin method [3–6], one considers the Langevin equation in the complexified configuration space. In the path optimization method [7–10], with the aid of machine learning one looks for an optimized integration surface for which the average phase factor is maximal. In the Lefschetz thimble method [11–23], one deforms the integration surface according to the antiholomorphic gradient flow. The deformed surface asymptotes to a union of Lefschetz thimbles, each of which gives a constant value to the imaginary part of the action and thus is free from the sign problem. Although the ergodicity problem can appear due to the existence of infinitely high potential barriers between thimbles, this ergodicity problem can be diminished by tempering the system with the flow time [19]. This *tempered Lefschetz thimble method* (TLTM) thus solves the sign and ergodicity problems simultaneously. Moreover, the computational cost of TLTM has recently been reduced significantly by developing

the *worldvolume tempered Lefschetz thimble method* (WV-TLTM) [23], which we are going to review now.

Let $\mathbb{R}^N = \{x\}$ be the configuration space and $S(x)$ the action (allowed to be complex-valued). Our main interest is to numerically estimate the expectation values of observables $\mathcal{O}(x)$,

$$\langle \mathcal{O} \rangle \equiv \frac{\int_{\mathbb{R}^N} dx \, e^{-S(x)} \, \mathcal{O}(x)}{\int_{\mathbb{R}^N} dx \, e^{-S(x)}}. \tag{1}$$

Under the assumption that $e^{-S(z)}$ and $e^{-S(z)} \mathcal{O}(z)$ are entire functions of $z \in \mathbb{C}^N$, Cauchy's theorem allows us to continuously deform the integration surface without changing the value of the integral. By expressing the deformation as a flow $z_t(x)(t \geq 0)$ with $z_0(x) = x$, the deformed integration surface at flow time t can be written as $\Sigma_t = \{z_t(x) \,|\, x \in \mathbb{R}^N\}$, and thus we have the equality

$$\langle \mathcal{O} \rangle = \frac{\int_{\Sigma_t} dz_t \, e^{-S(z_t)} \, \mathcal{O}(z_t)}{\int_{\Sigma_t} dz_t \, e^{-S(z_t)}}. \tag{2}$$

Since the numerator and the denominator are both independent of t, we can integrate each of them over an arbitrary interval $[T_0, T_1]$ with an arbitrary function $W(t)$. This rewrites Eq. (2) to the integration over the *worldvolume* $\mathcal{R} \equiv \bigcup_{T_0 \leq t \leq T_1} \Sigma_t$:

$$\langle \mathcal{O} \rangle = \frac{\int_{T_0}^{T_1} dt \, e^{-W(t)} \int_{\Sigma_t} dz_t \, e^{-S(z_t)} \, \mathcal{O}(z_t)}{\int_{T_0}^{T_1} dt \, e^{-W(t)} \int_{\Sigma_t} dz_t \, e^{-S(z_t)}} = \frac{\int_{\mathcal{R}} Dz \, K(z) \, e^{-W(t(z))} \, e^{-S(z)} \, \mathcal{O}(z)}{\int_{\mathcal{R}} Dz \, K(z) \, e^{-W(t(z))} \, e^{-S(z)}}, \tag{3}$$

where Dz is the induced volume element on $\mathcal{R}$, and $K(z)$ is the Jacobian, $dt \, dz_t = K(z) \, Dz$. The weight factor $e^{-W(t)}$ is chosen so that the probability for a configuration to appear at time t is (almost) independent of t. This setting is especially necessary when the whole range of t is relevant to simulations, as in the WV-TLTM. We further rewrite Eq. (3) to the ratio of reweighted averages:

$$\langle \mathcal{O} \rangle = \frac{\int_{\mathcal{R}} Dz \, e^{-V(z)} \, A(z) \, \mathcal{O}(z)}{\int_{\mathcal{R}} Dz \, e^{-V(z)} \, A(z)} = \frac{\langle A(z) \, \mathcal{O}(z) \rangle_{\mathcal{R}}}{\langle A(z) \rangle_{\mathcal{R}}}. \tag{4}$$

Here, the reweighted average

$$\langle \mathcal{O}(z) \rangle_{\mathcal{R}} \equiv \frac{\int_{\mathcal{R}} Dz \, e^{-V(z)} \, \mathcal{O}(z)}{\int_{\mathcal{R}} Dz \, e^{-V(z)}} \tag{5}$$

is defined for the weight $e^{-V(z)} \equiv e^{-\mathrm{Re}\,S(z) - W(t(z))}$, and $A(z)$ is the reweighting factor,

$$A(z) \equiv \frac{e^{-S(z) - W(t(z))} dt \, dz_t}{e^{-V(z)} Dz} = e^{-i\,\mathrm{Im}\,S(z)} K(z), \tag{6}$$

whose explicit form can be found in Ref. [23].

The reweighted average in Eq. (5) is estimated by the average over a sample generated by the hybrid Monte Carlo (HMC) algorithm with the potential $V(z)$. We will generally call an HMC algorithm on an accumulation of integration surfaces the *worldvolume hybrid Monte Carlo algorithm* (the WV-HMC algorithm), which includes the WV-TLTM.

In the case of the WV-TLTM, the interval $[T_0, T_1]$ should include the region with large t so as to solve the sign problem, and at the same time include the region with small t so as to reduce the ergodicity problem. However, the small-t region may be contaminated by the sign problem, and the large-t region by the ergodicity problem. Thus, in order to reduce the contributions from these potentially contaminated regions, it was proposed in Ref. [23] to estimate the observables from a subsample in a subregion with $[\tilde{T}_0, \tilde{T}_1] \, (\subset [T_0, T_1])$, see Fig. 1, where $\tilde{T}_0$ and $\tilde{T}_1$ are chosen
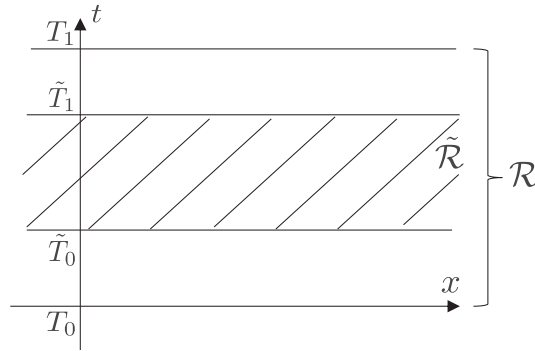
Fig. 1. The worldvolume $\mathcal{R}$ and a subregion $\tilde{\mathcal{R}}$.

such that the estimated values vary only slightly against the changes of $\tilde{T}_0$ and $\tilde{T}_1$.[1]

In Ref. [23] the standard error analysis was employed as if the subsample itself is generated as a Markov chain with ergodicity and detailed balance. However, this is not obvious and needs a justification. In this paper we prove that, if consecutive configurations in the worldvolume are generated as a Markov chain with ergodicity and detailed balance, then the subset consisting of the configurations belonging to a subregion can also be regarded as a Markov chain with ergodicity and detailed balance intact. We particularly investigate the integrated autocorrelation times for the Markov chains in various subregions, and find that there is a linear relation between the probability of being in a subregion and the integrated autocorrelation time for the corresponding subsample. We numerically confirm this scaling law for a chiral random matrix model (the Stephanov model [24,25]).

This paper is organized as follows. In Sect. 2 we first prove that the subset consisting of the configurations in a subregion is a Markov chain. We then argue that there should be a linear relation between the probability of bgine in a subregion and the integrated autocorrelation time for the corresponding subsample. Section 3 demonstrates this scaling by explicit numerical calculations for the Stephanov model. Section 4 is devoted to conclusion and outlook. In Appendix A we explain our Jackknife method for estimating the statistical errors of the integrated autocorrelation times.

2. Stochastic process in a subregion and the integrated autocorrelation time

Let $\mathcal{R} = \{z\}$ be the full configuration space (the worldvolume in the case of the WV-TLTM). Suppose that we are given a Markov chain $\{z_k\}(k = 0, 1, 2, \ldots)$ in $\mathcal{R}$ with the transition probability $P(z_{k+1}|z_k)$ that satisfies ergodicity as well as the detailed balance condition with respect to the unique equilibrium distribution $p_{eq}(z)$:

$$P(z' \mid z)\, p_{eq}(z) = P(z \mid z')\, p_{eq}(z'). \tag{7}$$

Only in this section, we write $\int_{\mathcal{R}} dz\, p_{eq}(z)\, \mathcal{O}(z)$ simply by $\langle \mathcal{O} \rangle$ (instead of $\langle \mathcal{O} \rangle_{\mathcal{R}}$).

2.1 *Stochastic process in a subregion*

We now look at a subregion $\tilde{\mathcal{R}}$ in $\mathcal{R}$, whose complement we denote by $\tilde{\mathcal{R}}^c \equiv \mathcal{R} \backslash \tilde{\mathcal{R}}$.[2] Then, from the Markov chain $\{z_k\}(k = 0, 1, 2, \ldots)$ in $\mathcal{R}$, we can extract a subsequence $\{\tilde{z}_\ell\}(\ell = 0, 1, 2, \ldots)$

---

[1] $\tilde{T}_{0,1}$ were written as $\hat{T}_{0,1}$ in Ref. [23].

[2] In the WV-TLTM, $\mathcal{R} = \bigcup_{t=T_0}^{T_1} \Sigma_t$ and $\tilde{\mathcal{R}} = \bigcup_{t=\tilde{T}_0}^{\tilde{T}_1} \Sigma_t$ with $T_0 \leq \tilde{T}_0 < \tilde{T}_1 \leq T_1$.

that consists of the configurations belonging to $\tilde{\mathcal{R}}$. We first notice that this sequence is a Markov chain with the following transition probability from $\tilde{z} \in \tilde{\mathcal{R}}$ to $\tilde{z}' \in \tilde{\mathcal{R}}$:

$$\tilde{P}(\tilde{z}' \mid \tilde{z}) = P(\tilde{z}' \mid \tilde{z}) + \int_{\tilde{\mathcal{R}}^c} dw \, P(\tilde{z}' \mid w) P(w \mid \tilde{z})$$

$$+ \int_{\tilde{\mathcal{R}}^c} dw_2 \, dw_1 \, P(\tilde{z}' \mid w_2) P(w_2 \mid w_1) P(w_1 \mid \tilde{z}) + \cdots . \tag{8}$$

Since $P(z'|z)$ is ergodic by assumption, and thus since a configuration which has left $\tilde{\mathcal{R}}$ will eventually reenter $\tilde{\mathcal{R}}$ at a finite number of steps, $\tilde{P}(\tilde{z}' \mid \tilde{z})$ is ergodic and satisfies probability conservation:

$$\int_{\tilde{\mathcal{R}}} d\tilde{z}' \, \tilde{P}(\tilde{z}' \mid \tilde{z}) = 1 \qquad (\forall \tilde{z} \in \tilde{\mathcal{R}}). \tag{9}$$

Furthermore, using the expression in Eq. (8), one can easily show that $\tilde{P}(\tilde{z}' \mid \tilde{z})$ satisfies the following equality:

$$\tilde{P}(\tilde{z}' \mid \tilde{z}) \, p_{\text{eq}}(\tilde{z}) = \tilde{P}(\tilde{z} \mid \tilde{z}') \, p_{\text{eq}}(\tilde{z}') \qquad (\tilde{z}', \tilde{z} \in \tilde{\mathcal{R}}), \tag{10}$$

from which we find that the unique equilibrium distribution $\tilde{p}_{\text{eq}}(\tilde{z})$ for $\tilde{P}(\tilde{z}' \mid \tilde{z})$ is given by

$$\tilde{p}_{\text{eq}}(\tilde{z}) = \frac{p_{\text{eq}}(\tilde{z})}{\int_{\tilde{\mathcal{R}}} d\tilde{z}' \, p_{\text{eq}}(\tilde{z}')}. \tag{11}$$

We have thus proved that the estimation of the expectation value $\int_{\tilde{\mathcal{R}}} d\tilde{z} \, \tilde{p}_{\text{eq}}(\tilde{z}) \, \mathcal{O}(\tilde{z})$ with the subsample from a subregion $\tilde{\mathcal{R}}$ can be statistically analyzed as if it is a Markov chain.

2.2 *Integrated autocorrelation time for a subchain*

Let $\{z_k\}(k = 0, 1, 2, \ldots)$ again be a Markov chain in $\mathcal{R}$ with the transition probability $P(z'|z)$. Denoting $\mathcal{O}(z_k)$ by $\mathcal{O}_k$, we define the integrated autocorrelation time of $\mathcal{O}$ by[3]

$$\tau_{\text{int}}(\mathcal{O}) \equiv 1 + 2 \sum_{k=1}^{\infty} \frac{C_k(\mathcal{O})}{C_0(\mathcal{O})}, \tag{12}$$

where $C_k(\mathcal{O}) \equiv \langle \mathcal{O}_0 \mathcal{O}_k \rangle_c \equiv \langle (\mathcal{O}_0 - \langle \mathcal{O} \rangle)(\mathcal{O}_k - \langle \mathcal{O} \rangle) \rangle$ is the autocorrelation. Similarly, we define the integrated autocorrelation time for the subchain $\{\tilde{z}_\ell\}(\ell = 0, 1, 2, \ldots)$ in $\tilde{\mathcal{R}}$, and denote it by $\tilde{\tau}_{\text{int}}(\mathcal{O})$. Note that $\tilde{\tau}_{\text{int}}(\mathcal{O})$ is generically smaller than $\tau_{\text{int}}(\mathcal{O})$, because a one-step transition with $\tilde{P}$ can correspond to transitions of multiple steps with P.

The ratio $\tilde{\tau}_{\text{int}}(\mathcal{O})/\tau_{\text{int}}(\mathcal{O})$ can be evaluated as follows, when both the numerator and the denominator are not too small. We first write by ϵ and $\tilde{\epsilon}$, respectively, the average Monte Carlo times evolved in one-step transition with P and $\tilde{P}$.[4] We then note that $\tilde{\epsilon}$ can be written with ϵ by using the probability p for a configuration in $\tilde{\mathcal{R}}$ to stay in $\tilde{\mathcal{R}}$ at the next step and the probability q for a configuration in $\tilde{\mathcal{R}}^c$ to stay in $\tilde{\mathcal{R}}^c$ as well:

$$\tilde{\epsilon} = p \epsilon + (1 - p)(1 - q) 2\epsilon + (1 - p)q(1 - q) 3\epsilon + \cdots$$

$$= \frac{2 - p - q}{1 - q} \epsilon$$

$$\left(= \frac{(1 - p) + (1 - q)}{1 - q} \epsilon \geq \epsilon \right). \tag{13}$$

---

[3]This normalization gives the effective sample size as $N_{\text{conf}}^{\text{eff}} = N_{\text{conf}}/\tau_{\text{int}}$ (see, e.g., Ref. [26]).

[4]This ϵ corresponds to the Langevin time for Langevin algorithms and to the molecular dynamics time multiplied by the average acceptance rate for HMC algorithms.

Furthermore, since autocorrelations should be the same at large Monte Carlo time scales, we have the equality

$$\tau_{\rm int}(\mathcal{O})\,\epsilon = \tilde{\tau}_{\rm int}(\mathcal{O})\,\tilde{\epsilon}. \tag{14}$$

Note that this renormalization-group-like argument holds only when $\tau_{\rm int}(\mathcal{O})$ and $\tilde{\tau}_{\rm int}(\mathcal{O})$ are not too small. Combining Eqs. (13) and (14), we obtain the desired result:

$$\frac{\tilde{\tau}_{\rm int}(\mathcal{O})}{\tau_{\rm int}(\mathcal{O})} = \frac{\epsilon}{\tilde{\epsilon}} = \frac{1-q}{2-p-q} \le 1. \tag{15}$$

2.3 *Scaling law for the integrated autocorrelation times*

We now apply the preceding arguments to the case where the configuration space is the world-volume of the WV-TLTM. We argue that there must be a linear relation between the probability of being in a subregion $\tilde{\mathcal{R}}$ and the integrated autocorrelation time $\tilde{\tau}_{\rm int}$ for the corresponding subchain. This claim will be confirmed numerically in the next section.

To simplify discussions, we assume that the integrated autocorrelation time for the flow time t, i.e. $\tau_{\rm int}(\mathcal{O})$ with $\mathcal{O}(z) = t(z)$, is sufficiently small, $\tau_{\rm int}(t) \simeq 1$. This can be easily realized, if necessary, by removing consecutive configurations from the chain at appropriate intervals. The smallness of $\tau_{\rm int}(t)$ means that the probability p simply expresses the probability for a configuration to be in $\tilde{\mathcal{R}}$. We now recall that the distribution of t is uniform in equilibrium for the WV-TLTM—see the discussion below Eq. (3). Thus, p is given by

$$p = \frac{\tilde{T}_1 - \tilde{T}_0}{T_1 - T_0}. \tag{16}$$

A similar statement holds for the probability q, which now expresses the probability for a configuration to be in $\tilde{\mathcal{R}}^c$, and thus we obtain the relation $p + q = 1$. Then, Eq. (13) leads to the relation $\tilde{\epsilon} = \epsilon/p$, and thus, combined with Eq. (16), we obtain the following scaling law for the integrated autocorrelation times:

$$p = \frac{\tilde{T}_1 - \tilde{T}_0}{T_1 - T_0} = \frac{\tilde{\tau}_{\rm int}(\mathcal{O})}{\tau_{\rm int}(\mathcal{O})}. \tag{17}$$

We now consider the numerical estimation of $\langle\mathcal{O}\rangle$ using a subsample belonging to the interval $[\tilde{T}_0, \tilde{T}_1]$. Since Cauchy's theorem ensures that $\langle\mathcal{O}\rangle$ is independent of the choice of $[\tilde{T}_0, \tilde{T}_1]$, the estimate should not vary largely against the changes around an appropriately chosen pair $[\tilde{T}_0, \tilde{T}_1]$. Furthermore, the statistical error $\delta\langle\mathcal{O}\rangle$ also hardly depends on the choice of $[\tilde{T}_0, \tilde{T}_1]$. To see this, let us write the number of configurations in the interval $[\tilde{T}_0, \tilde{T}_1]$ as $N_{\rm conf}(\tilde{T}_0, \tilde{T}_1)$. Since the distribution of t is almost uniform, the ratio $N_{\rm conf}(\tilde{T}_0, \tilde{T}_1)/N_{\rm conf}(T_0, T_1)$ almost equals p, and thus we obtain the relation

$$\frac{N_{\rm conf}(\tilde{T}_0, \tilde{T}_1)}{N_{\rm conf}(T_0, T_1)} \simeq \frac{\tilde{\tau}_{\rm int}(\mathcal{O})}{\tau_{\rm int}(\mathcal{O})}. \tag{18}$$

This in turn means that the effective number of configurations in a subsample does not depend on the choice of the interval $[\tilde{T}_0, \tilde{T}_1]$, because

$$N_{\rm conf}^{\rm eff}(\mathcal{O}; \tilde{T}_0, \tilde{T}_1) \equiv \frac{N_{\rm conf}(\tilde{T}_0, \tilde{T}_1)}{\tilde{\tau}_{\rm int}(\mathcal{O})} \simeq \frac{N_{\rm conf}(T_0, T_1)}{\tau_{\rm int}(\mathcal{O})} \qquad (\forall\, \tilde{T}_0, \tilde{T}_1). \tag{19}$$

When $N_{\rm conf}$ is sufficiently large (as we always assume), the statistical error is given by the formula $\delta\langle\mathcal{O}\rangle = \sigma_{\mathcal{O}}/\{N_{\rm conf}^{\rm eff}(\mathcal{O}; \tilde{T}_0, \tilde{T}_1)\}^{1/2}$ with a constant $\sigma_{\mathcal{O}}\ (= \{\langle\mathcal{O}^2\rangle_c\}^{1/2})$. Since $N_{\rm conf}^{\rm eff}(\mathcal{O}; \tilde{T}_0, \tilde{T}_1)$ is almost independent of the choice of $[\tilde{T}_0, \tilde{T}_1]$, so is $\delta\langle\mathcal{O}\rangle$.

3. Numerical confirmation of the scaling law

We numerically confirm the scaling law in Eq. (17) for the WV-TLTM [23] applied to a chiral random matrix model (the Stephanov model [24,25]).

3.1 Setup

The partition function of the Stephanov model for N_f quarks with equal mass m is given by

$$Z_n^{N_f} = e^{n\mu^2} \int d^2X \, e^{-S(X, X^\dagger)} \equiv e^{n\mu^2} \int d^2X \, e^{-n\,\mathrm{tr}\,X^\dagger X} \det^{N_f}(D + m), \tag{20}$$

where $X = (X_{ij} = x_{ij} + i\,y_{ij})$ is an $n \times n$ complex matrix. The $2n \times 2n$ matrix D expresses the Dirac operator in the chiral representation,

$$D \equiv \begin{pmatrix} 0 & i(X + C) \\ i(X^\dagger + C) & 0 \end{pmatrix}, \tag{21}$$

$$iC \equiv \begin{pmatrix} (\mu + i\tau)\,1_{n/2} & 0 \\ 0 & (\mu - i\tau)\,1_{n/2} \end{pmatrix}, \tag{22}$$

where μ and τ represent the chemical potential and the temperature, respectively. The chiral condensate and the number density are defined, respectively, by

$$\langle \bar{\psi}\psi \rangle \equiv \frac{1}{2n} \frac{\partial}{\partial m} \ln Z_n^{N_f}, \qquad \langle \psi^\dagger \psi \rangle \equiv \frac{1}{2n} \frac{\partial}{\partial \mu} \ln Z_n^{N_f}. \tag{23}$$

We will set the parameters to $N_f = 1$, $n = 2$, $\mu = 0.6$, $\tau = 0$, and $m = 0.004$.

We generate a sample $\{z_j\}(j = 1, \ldots, N_{\mathrm{conf}})$ with the HMC algorithm using the potential $V(z)$, and estimate the reweighted average $\langle \mathcal{O}(z) \rangle_{\mathcal{R}}$, see Eq. (5), by the sample average

$$\overline{\mathcal{O}} \equiv \frac{1}{N_{\mathrm{conf}}} \sum_{j=1}^{N_{\mathrm{conf}}} \mathcal{O}(z_j). \tag{24}$$

The estimator of the integrated autocorrelation time $\tau_{\mathrm{int}}(\mathcal{O})$ is given by

$$\overline{\tau}_{\mathrm{int}}(\mathcal{O}; k_{\max}) \equiv 1 + 2 \sum_{k=1}^{k_{\max}} \frac{\overline{C}_k(\mathcal{O})}{\overline{C}_0(\mathcal{O})}. \tag{25}$$

Here, $\overline{C}_k(\mathcal{O})$ is an estimator of the autocorrelation $C_k(\mathcal{O}) = \langle \mathcal{O}_0 \mathcal{O}_k \rangle_{\mathcal{R}, c}$, whose explicit form is given in Appendix A. We have truncated the summation at $k_{\max}$ to avoid summing statistical fluctuations around zero (see, e.g., Ref. [28]). The statistical error $\delta\overline{\tau}_{\mathrm{int}}(\mathcal{O}; k_{\max})$ is estimated by a jackknife method that is described in Appendix A. Values of $k_{\max}$ and bin sizes used in the estimations of $\overline{\tau}_{\mathrm{int}}(\mathcal{O}; k_{\max})$ are summarized in Table 1.

In the numerical simulation with the HMC, we set $T_0 = 0$ and $T_1 = 0.02$. The HMC updates are performed with the molecular dynamics time increment $\Delta s = 0.001$ and the step number $n_{\mathrm{HMC}} = 100$, with the average acceptance rate more than 0.99. We employ 20 independent sets of configurations, each set consisting of 3×10^6 configurations in $[T_0, T_1]$. The observables are measured every six iterations of the HMC algorithm, so that we have 20 independent samples of size $N_{\mathrm{conf}} = 5 \times 10^5$. We fix $\tilde{T}_1 = T_1$ and vary $\tilde{T}_0$ as $\tilde{T}_0^{(i)} \equiv (i/20)T_1$ ($i = 0, \ldots, 19$), for each of which an independent set of configurations is used.

3.2 Results

We now demonstrate the scaling law in Eq. (17) from explicit numerical calculations. Recall that the argument for the scaling is based on the smallness of the integrated autocorrela-

Table 1. Values of k_{max} and bin sizes.

$\mathcal{O}$	i	0	1	2	3	4	5	6	7	8	9
t	k_{max}	75	60	60	60	60	60	50	50	50	50
	bin size	140	140	140	140	140	120	120	100	100	100
$\mathrm{Re}\,(\bar{\psi}\psi A)$	k_{max}	900	250	250	250	250	300	300	300	400	500
	bin size	1100	800	900	500	900	500	600	500	500	400
$\mathrm{Re}\,(\psi^{\dagger}\psi A)$	k_{max}	500	500	500	500	500	500	500	500	500	500
	bin size	1000	1300	1200	1300	1100	1100	1100	1100	1100	1300
$\mathrm{Re}\,A$	k_{max}	500	500	500	500	500	500	500	500	500	500
	bin size	1000	1000	900	700	700	600	600	500	500	500
$\mathcal{O}$	i	10	11	12	13	14	15	16	17	18	19
t	k_{max}	50	50	50	50	50	40	30	25	25	15
	bin size	100	100	100	100	120	80	100	120	100	100
$\mathrm{Re}\,(\bar{\psi}\psi A)$	k_{max}	500	400	500	500	300	250	400	250	400	300
	bin size	400	400	300	300	300	300	400	400	400	400
$\mathrm{Re}\,(\psi^{\dagger}\psi A)$	k_{max}	500	500	500	500	500	500	500	500	250	400
	bin size	1100	1100	1300	1300	1100	1100	1100	1100	1100	1100
$\mathrm{Re}\,A$	k_{max}	500	500	500	500	500	500	500	250	250	250
	bin size	500	400	500	400	400	400	400	500	400	500
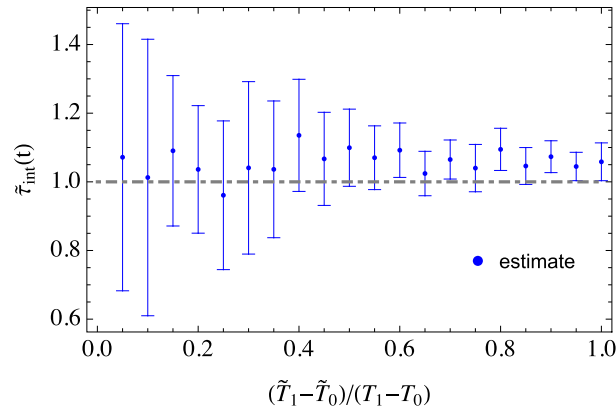
Fig. 2. Results of $\tilde{\tau}_{\mathrm{int}}(t)$.

tion time of t ($\tau_{\mathrm{int}}(t) \simeq 1$) and the uniformity of the distribution of t. Figure 2 shows that the condition $\tilde{\tau}_{\mathrm{int}}(t) \simeq 1$ is satisfied for all $p = (\tilde{T}_1 - \tilde{T}_0)/(T_1 - T_0)$. Figure 3 is a histogram of $p = (\tilde{T}_1 - \tilde{T}_0)/(T_1 - T_0)$, which is almost flat as required. This is realized by tuning the functional form of $W(t)$ as in Ref. [23].

Figure 4 exhibits the scaling law in Eq. (17) for three operators, $\mathcal{O} = \mathrm{Re}\,(\bar{\psi}\psi A)$, $\mathcal{O} = \mathrm{Re}\,(\psi^{\dagger}\psi A)$, and $\mathcal{O} = \mathrm{Re}\,A$. We see that the scaling is satisfied for $\mathcal{O} = \mathrm{Re}\,(\bar{\psi}\psi A)$ in the region $(\tilde{T}_1 - \tilde{T}_0)/(T_1 - T_0) \geq 0.35$, and for $\mathcal{O} = \mathrm{Re}\,(\psi^{\dagger}\psi A)$ and $\mathcal{O} = \mathrm{Re}\,A$ in the region $(\tilde{T}_1 - \tilde{T}_0)/(T_1 - T_0) \geq 0.20$. Deviations from the scaling at small $(\tilde{T}_1 - \tilde{T}_0)/(T_1 - T_0)$ should be due to $\tilde{\tau}_{\mathrm{int}}(\mathcal{O}) = O(1)$—see the comment below Eq. (14). We perform the χ^2-fit to these data points with

$$\chi^2(\mathcal{O}) \equiv \sum_{i=0}^{i_{\mathrm{max}}} \frac{\left[\tilde{\tau}_{\mathrm{int}}^{(i)}(\mathcal{O}) - c\left(\tilde{T}_1 - \tilde{T}_0^{(i)}\right)/(T_1 - T_0)\right]^2}{\left[\delta\tilde{\tau}_{\mathrm{int}}^{(i)}(\mathcal{O})\right]^2}, \tag{26}$$
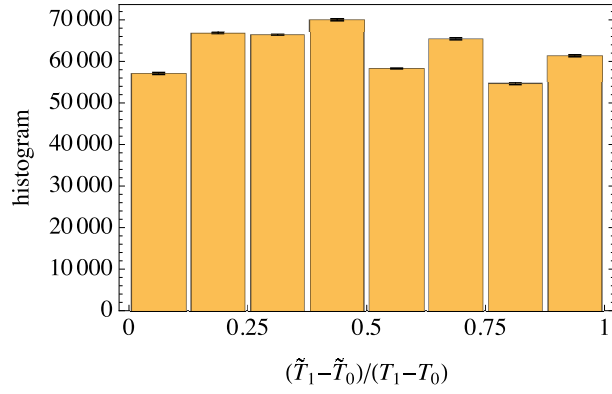
Fig. 3. Histogram of t. The number of configurations in a bin is the average over 20 independent samples, each consisting of $N_{\mathrm{conf}} = 5 \times 10^5$. The function $W(t)$ is tuned such that the histogram becomes almost flat for these eight bins.
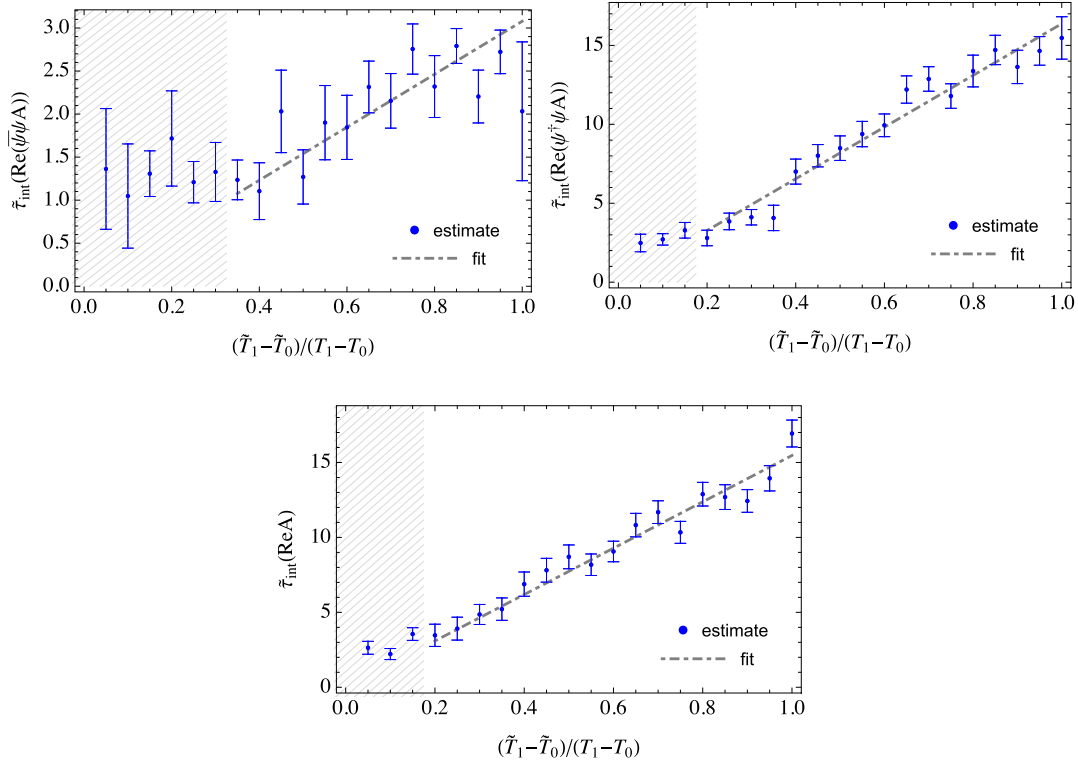

Fig. 4. $\tilde{\tau}_{\mathrm{int}}(\mathcal{O})$ against $(\tilde{T}_1 - \tilde{T}_0)/(T_1 - T_0)$. Top left: $\mathcal{O} = \mathrm{Re}\,(\bar{\psi}\psi A)$. Top right: $\mathcal{O} = \mathrm{Re}\,(\psi^{\dagger}\psi A)$. Bottom: $\mathcal{O} = \mathrm{Re}\,A$. Data points in the shaded region give $\tilde{\tau}_{\mathrm{int}}(\mathcal{O}) = O(1)$ and thus are excluded from the linear fitting—see the comment below Eq. (14).

where $\tilde{\tau}_{\mathrm{int}}^{(i)}(\mathcal{O})$ is the integrated autocorrelation time for the subsample with the interval $[\tilde{T}_0, \tilde{T}_1] = [\tilde{T}_0^{(i)}, T_1]$. The fit results are as follows: for $\mathcal{O} = \mathrm{Re}\,(\bar{\psi}\psi A)$ with $i_{\max} = 13$, $c = 3.1$ and $\chi^2/\mathrm{DOF} = 1.0$; for $\mathcal{O} = \mathrm{Re}\,(\psi^{\dagger}\psi A)$ with $i_{\max} = 16$, $c = 16.4$ and $\chi^2/\mathrm{DOF} = 1.2$; for $\mathcal{O} = \mathrm{Re}\,A$ with $i_{\max} = 16$, $c = 15.5$ and $\chi^2/\mathrm{DOF} = 1.1$. These data support the scaling law in Eq. (17).

We plot in Fig. 5 the values of $N_{\mathrm{conf}}^{\mathrm{eff}}(\mathcal{O}; \tilde{T}_0, \tilde{T}_1)$ for various operators $\mathcal{O}$—see Eq. (19). The statistical errors are estimated with the jackknife method. We observe, as expected, that $N_{\mathrm{conf}}^{\mathrm{eff}}(\mathcal{O}; \tilde{T}_0, \tilde{T}_1)$ takes almost the same values for each $\mathcal{O}$ in the range where we observe the scaling law.
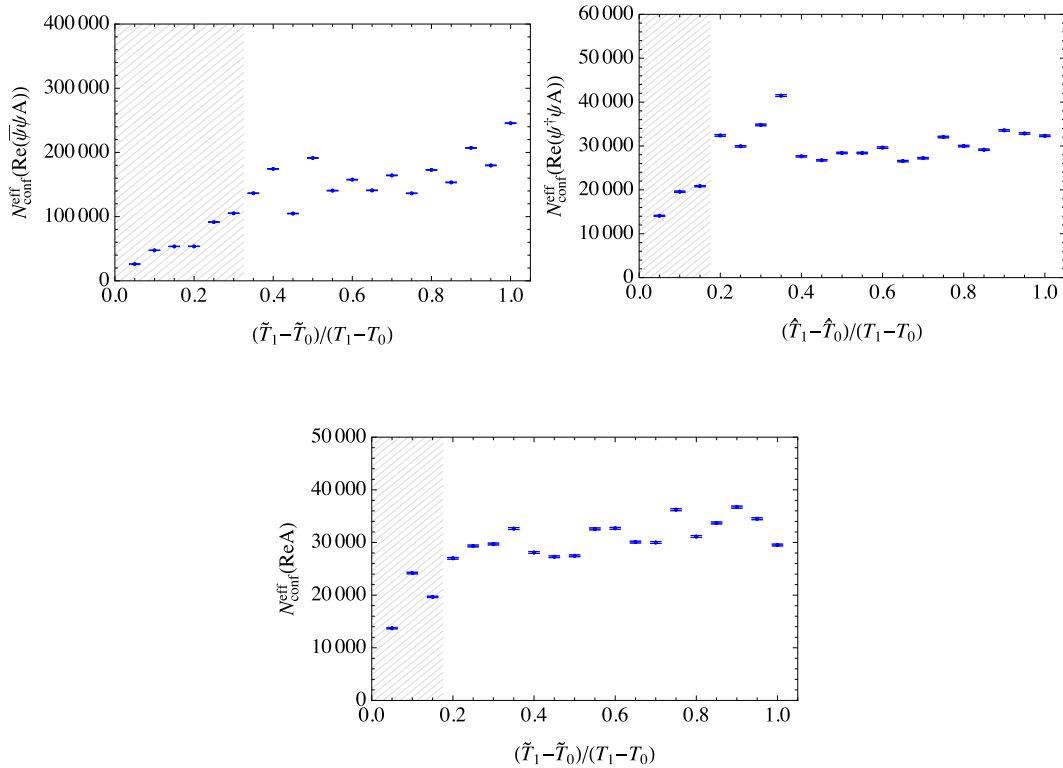
Fig. 5. The values of $N_{\mathrm{conf}}^{\mathrm{eff}}(\mathcal{O}; \tilde{T}_0, \tilde{T}_1)$—see Eq. (19). Top left: $\mathcal{O} = \mathrm{Re}\,(\bar{\psi}\psi A)$. Top right: $\mathcal{O} = \mathrm{Re}\,(\psi^\dagger\psi A)$. Bottom: $\mathcal{O} = \mathrm{Re}\,A$. The regions where the scaling law in Eq. (17) is broken are shaded.

Finally, in Fig. 6 we plot the expectation values of the chiral condensate and the number density, together with their statistical errors. The statistical errors are estimated with the jackknife method with the bin sizes fixed to 500. We see that both the means and the statistical errors take almost constant values in a region where $(\tilde{T}_1 - \tilde{T}_0)/(T_1 - T_0)$ is not small. The deviation of the means should be attributed to the complex geometry at large flow times, which requires larger statistics and better control of systematic errors (such as those from numerical integrations of the flow equation and of the Hamilton equations accompanied by the projection from $\mathbb{C}^N$ to $\mathcal{R}$). The deviation of the statistical errors is due to the violation of the scaling law in Eq. (17) for either the numerator or the denominator (or both) in the ratio of the reweighted averages.

4.   Summary and outlook

We have established a statistical analysis method for the WV-HMC algorithm, whose major use is intended for the WV-TLTM [23]. We proved that, if consecutive configurations in the world-volume are generated as a Markov chain with ergodicity and detailed balance, then the subset consisting of the configurations belonging to a subregion can also be regarded as a Markov chain with ergodicity and detailed balance intact. We particularly investigated the integrated autocorrelation times for the Markov chains in various subregions, and found that there is a linear relation between the probability of being in a subregion and the integrated autocorrelation time for the corresponding subsample. We numerically confirmed this scaling law for the Stephanov model.

Now, with this statistical analysis method at hand, we can safely apply the WV-TLTM to large-scale simulations of the systems that have serious sign problems, such as finite-density
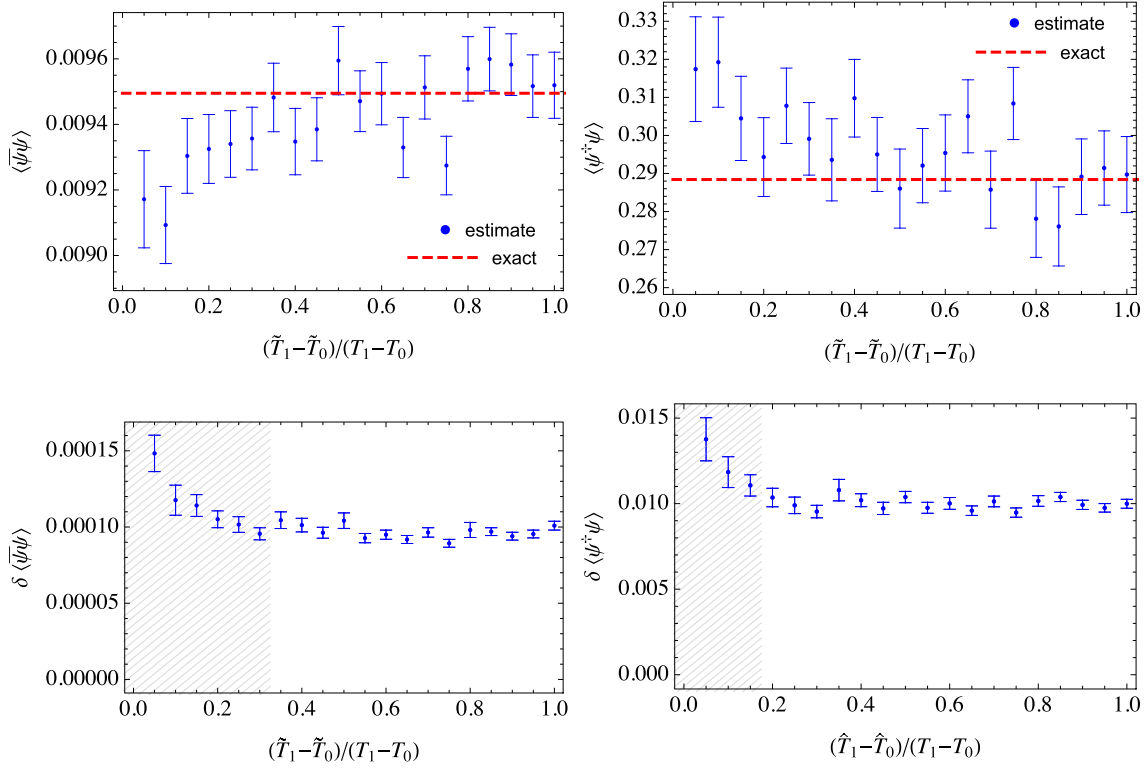
Fig. 6. The estimate (top) and its statistical error (bottom) of the expectation value $\langle \mathcal{O} \rangle$ against $(\tilde{T}_1 - \tilde{T}_0)/(T_1 - T_0)$. Left: $\mathcal{O} = \bar{\psi}\psi$ (the chiral condensate). Right: $\mathcal{O} = \psi^\dagger\psi$ (the number density). The regions where the scaling law in Eq. (17) is broken in the numerator and/or the denominator are shaded for the statistical errors.

QCD, strongly correlated electron systems, frustrated spin systems, and the real-time dynamics of quantum fields. A study along this line is now in progress and will be reported elsewhere.

Appendix A. Jackknife method for the integrated autocorrelation times

In this appendix we give a jackknife method to estimate the integrated autocorrelation times $\tau_{\mathrm{int}}(\mathcal{O})$.

Let $\{x_j\}(j = 1, \ldots, N_{\mathrm{conf}})$ be a set of consecutive configurations generated as a Markov chain. We estimate the expectation value $\langle \mathcal{O} \rangle$ by the sample average

$$\overline{\mathcal{O}} \equiv \frac{1}{N_{\mathrm{conf}}} \sum_{j=1}^{N_{\mathrm{conf}}} \mathcal{O}(x_j). \tag{A1}$$

The estimator of the integrated autocorrelation time $\tau_{\text{int}}(\mathcal{O}) \equiv 1 + 2\sum_{k=1}^{\infty} C_k(\mathcal{O})/C_0(\mathcal{O})$ is given by

$$\bar{\tau}_{\text{int}}(\mathcal{O}; k_{\max}) \equiv 1 + 2\sum_{k=1}^{k_{\max}} \frac{\overline{C}_k(\mathcal{O})}{\overline{C}_0(\mathcal{O})}, \tag{A2}$$

where $\overline{C}_k(\mathcal{O})$ is the estimator of the autocorrelation $C_k(\mathcal{O}) \equiv \langle \mathcal{O}_0 \mathcal{O}_k \rangle_c$. The summation is truncated at $k_{\max}$ to avoid summing statistical fluctuations around zero (see, e.g., Refs. [27,28]). The value of $k_{\max}$ should not be set very large compared to $\tau_{\text{int}}(\mathcal{O})$, otherwise contributions from statistical fluctuations around zero may dominate the error. An explicit formula for the statistical error $\delta\tau_{\text{int}}(\mathcal{O})$ is known when $1 \ll k_{\max} \ll N_{\text{conf}}$ (more precisely, as $N_{\text{conf}} \to \infty$, $k_{\max} \to \infty$, and $k_{\max}/N_{\text{conf}} \to 0$) [26] (see also Refs. [27,28]),[5] but this may not be applicable to the case when $\tau_{\text{int}}(\mathcal{O}) = O(1)$, for which the condition $k_{\max} \gg 1$ cannot be met. Therefore, in this paper we adopt the jackknife method for the estimation of $\delta\tau_{\text{int}}(\mathcal{O})$.

In order to apply a resampling jackknife method, we introduce a sample of multidimensional observables $\{X_r = (X_{r,k}^{\mathcal{O}\mathcal{O}}, X_r^{\mathcal{O}})\}$ $(r = 1, \ldots, N_{\text{conf}} - k_{\max})$ with

$$X_{r,k}^{\mathcal{O}\mathcal{O}} \equiv \frac{1}{k_{\max} - k + 1} \sum_{i=0}^{k_{\max}-k} \mathcal{O}(x_{r+i})\mathcal{O}(x_{r+k+i}) \qquad (k = 0, \ldots, k_{\max}), \tag{A3}$$

$$X_r^{\mathcal{O}} \equiv \frac{1}{k_{\max} + 1} \sum_{i=0}^{k_{\max}} \mathcal{O}(x_{r+i}). \tag{A4}$$

Since $\langle X_{r,k}^{\mathcal{O}\mathcal{O}} \rangle = \langle \mathcal{O}(x_0)\,\mathcal{O}(x_k) \rangle$ and $\langle X_r^{\mathcal{O}} \rangle = \langle \mathcal{O}(x) \rangle$, the autocorrelations $C_k(\mathcal{O}) = \langle \mathcal{O}(x_0)\mathcal{O}(x_k) \rangle_c$ can be estimated by

$$\overline{C}_k^{(X)}(\mathcal{O}) \equiv \frac{1}{N_{\text{conf}} - k_{\max}} \sum_{r=1}^{N_{\text{conf}}-k_{\max}} X_{r,k}^{\mathcal{O}\mathcal{O}} - \left(\frac{1}{N_{\text{conf}} - k_{\max}} \sum_{r=1}^{N_{\text{conf}}-k_{\max}} X_r^{\mathcal{O}} \right)^2. \tag{A5}$$

Since $\tau_{\text{int}}(\mathcal{O})$ is a function of $C_k(\mathcal{O})$, it can be estimated with $\overline{C}_k^{(X)}(\mathcal{O})$ as

$$\tau_{\text{int}}(\mathcal{O}) \approx 1 + 2\sum_{k=1}^{k_{\max}} \frac{\overline{C}_k^{(X)}(\mathcal{O})}{\overline{C}_0^{(X)}(\mathcal{O})}. \tag{A6}$$

The statistical error $\delta\tau_{\text{int}}(\mathcal{O})$ can then be estimated with the standard jackknife method.

References

[1] G. Aarts, J. Phys. Conf. Ser. **706**, 022004 (2016) [arXiv:1512.05145 [hep-lat]] [Search inSPIRE].
[2] L. Pollet, Rep. Prog. Phys. **75**, 094501 (2012) [arXiv:1206.0781 [cond-mat.quant-gas]] [Search inSPIRE].
[3] G. Parisi, Phys. Lett. B **131**, 393 (1983).
[4] J. R. Klauder, Phys. Rev. A **29**, 2036 (1984).
[5] G. Aarts, F. A. James, E. Seiler, and I. O. Stamatescu, Phys. Lett. B **687**, 154 (2010) [arXiv:0912.0617 [hep-lat]] [Search inSPIRE].
[6] J. Nishimura and S. Shimasaki, Phys. Rev. D **92**, 011501 (2015) [arXiv:1504.08359 [hep-lat]] [Search inSPIRE].
[7] Y. Mori, K. Kashiwa, and A. Ohnishi, Phys. Rev. D **96**, 111501 (2017) [arXiv:1705.05605 [hep-lat]] [Search inSPIRE].
[8] Y. Mori, K. Kashiwa, and A. Ohnishi, Prog. Theor. Exp. Phys. **2018**, 023B04 (2018) [arXiv:1709.03208 [hep-lat]] [Search inSPIRE].

---

[5] This is given by $\delta\tau_{\text{int}}(\mathcal{O}) = \tau_{\text{int}}(\mathcal{O})\sqrt{2(2k_{\max}+1)/N_{\text{conf}}}$ for the estimator of the autocorrelation, $\overline{C}_k(\mathcal{O}) = (N_{\text{conf}} - k)^{-1} \sum_{n=1}^{N_{\text{conf}}-k} (\mathcal{O}(x_n) - \overline{\mathcal{O}})(\mathcal{O}(x_{n+k}) - \overline{\mathcal{O}})$ [26].

[9]   A. Alexandru, P. F. Bedaque, H. Lamm, and S. Lawrence, Phys. Rev. D **97**, 094510 (2018) [arXiv:1804.00697 [hep-lat]] [Search inSPIRE].

[10] F. Bursa and M. Kroyter, J. High Energy Phys. **12**, 054 (2018) [arXiv:1805.04941 [hep-lat]] [Search inSPIRE].

[11] E. Witten, AMS/IP Stud. Adv. Math. **50**, 347 (2011) [arXiv:1001.2933 [hep-th]] [Search inSPIRE].

[12] M. Cristoforetti, F. Di Renzo, and L. Scorzato, Phys. Rev. D **86**, 074506 (2012) [arXiv:1205.3996 [hep-lat]] [Search inSPIRE].

[13] M. Cristoforetti, F. Di Renzo, A. Mukherjee, and L. Scorzato, Phys. Rev. D **88**, 051501(R) (2013) [arXiv:1303.7204 [hep-lat]] [Search inSPIRE].

[14] H. Fujii, D. Honda, M. Kato, Y. Kikukawa, S. Komatsu, and T. Sano, J. High Energy Phys. **1310**, 147 (2013) [arXiv:1309.4371 [hep-lat]] [Search inSPIRE].

[15] H. Fujii, S. Kamata, and Y. Kikukawa, J. High Energy Phys. **11**, 078 (2015); 02, 036 (2016) [erratum][arXiv:1509.08176 [hep-lat]] [Search inSPIRE].

[16] H. Fujii, S. Kamata, and Y. Kikukawa, J. High Energy Phys. **12**, 125 (2015); 09, 172 (2016) [erratum][arXiv:1509.09141 [hep-lat]] [Search inSPIRE].

[17] A. Alexandru, G. Başar, and P. Bedaque, Phys. Rev. D **93**, 014504 (2016) [arXiv:1510.03258 [hep-lat]] [Search inSPIRE].

[18] A. Alexandru, G. Başar, P. F. Bedaque, G. W. Ridgway, and N. C. Warrington, J. High Energy Phys. **1605**, 053 (2016) [arXiv:1512.08764 [hep-lat]] [Search inSPIRE].

[19] M. Fukuma and N. Umeda, Prog. Theor. Exp. Phys. **2017**, 073B01 (2017) [arXiv:1703.00861 [hep-lat]] [Search inSPIRE].

[20] A. Alexandru, G. Başar, P. F. Bedaque, and N. C. Warrington, Phys. Rev. D **96**, 034513 (2017) [arXiv:1703.02414 [hep-lat]] [Search inSPIRE].

[21] M. Fukuma, N. Matsumoto, and N. Umeda, Phys. Rev. D **100**, 114510 (2019) [arXiv:1906.04243 [cond-mat.str-el]] [Search inSPIRE].

[22] M. Fukuma, N. Matsumoto, and N. Umeda, arXiv:1912.13303 [hep-lat] [Search inSPIRE].

[23] M. Fukuma and N. Matsumoto, Prog. Theor. Exp. Phys. **2021**, 023B08 (2021) [arXiv:2012.08468 [hep-lat]] [Search inSPIRE].

[24] M. A. Stephanov, Phys. Rev. Lett. **76**, 4472 (1996) [arXiv:hep-lat/9604003] [Search inSPIRE].

[25] M. A. Halasz, A. D. Jackson, R. E. Shrock, M. A. Stephanov, and J. J. M. Verbaarschot, Phys. Rev. D **58**, 096007 (1998) [arXiv:hep-ph/9804290] [Search inSPIRE].

[26] N. Madras and A. D. Sokal, J. Stat. Phys. **50**, 109 (1988).

[27] U. Grenander and M. Rosenblatt, Ann. Math. Statist. **24**, 537 (1953).

[28] M. B. Priestley, Spectral Analysis and Time Series (Academic Press, London, 1981). 10.1002/for.3980010411.