PTEP

Prog. Theor. Exp. Phys. **2021**, 023B08 (33 pages)
DOI: 10.1093/ptep/ptab010

Worldvolume approach to the tempered Lefschetz thimble method

Masafumi Fukuma* and Nobuyuki Matsumoto*

Department of Physics, Kyoto University, Kyoto 606-8502, Japan
*E-mail: fukuma@gauge.scphys.kyoto-u.ac.jp, nobu.m@gauge.scphys.kyoto-u.ac.jp

Received December 22, 2020; Accepted January 20, 2021; Published February 10, 2021

As a solution towards the numerical sign problem, we propose a novel hybrid Monte Carlo algorithm, in which molecular dynamics is performed on a continuum set of integration surfaces foliated by the antiholomorphic gradient flow ("the worldvolume of an integration surface"). This is an extension of the tempered Lefschetz thimble method (TLTM) and solves the sign and multimodal problems simultaneously, as the original TLTM does. Furthermore, in this new algorithm, one no longer needs to compute the Jacobian of the gradient flow in generating a configuration, and only needs to evaluate its phase upon measurement. To demonstrate that this algorithm works correctly, we apply the algorithm to a chiral random matrix model, for which the complex Langevin method is known not to work.

Subject Index A20, A22, A23, B01, B38

1. Introduction

The sign problem is one of the major obstacles to numerical computation in various areas of physics, including finite density QCD [1], quantum Monte Carlo simulations of statistical systems [2], and the numerical simulations of real-time quantum field theories.

There have been proposed many Monte Carlo algorithms towards solving the sign problem, such as those based on the complex Langevin equation [3–9] (see, e.g., Ref. [10] for a review) and those on Lefschetz thimbles [11–21], each of which has its own advantages and disadvantages. The advantage of using the complex Langevin equation is its cheap computational cost, but such algorithms are known to suffer from a notorious problem called the "wrong convergence problem" (giving incorrect results with small statistical errors) for physically important ranges of parameters [6,7,9]. On the other hand, although computationally expensive, the algorithms based on Lefschetz thimbles are basically free from the wrong convergence. However, this is the case when and only when a single Lefschetz thimble is relevant for the estimation of observables, because otherwise there can appear another problem of multimodality due to infinitely high potential barriers between different thimbles.

The tempered Lefschetz thimble method (TLTM) [17] is a Lefschetz thimble method that avoids the sign and multimodal problems simultaneously, by introducing a discrete set of integration surfaces (replicas of integration surface) and exchanging configurations between replicas. The TLTM has proved effective and versatile when applied to various models, including the $(0+1)$-dimensional massive Thirring model [17] and the 2D Hubbard model away from half filling [20,21].[1] The disadvantage of the original TLTM is its computational cost, the cost coming from the computation of the Jacobian and from the additional cost due to the introduction of replicas.

[1] See Refs. [22,23] for another Lefschetz thimble approach to the sign problem in the Hubbard model.

Funded by SCOAP[3]

In this paper, as an extension of the TLTM, we propose a novel hybrid Monte Carlo (HMC) algorithm, where molecular dynamics is performed on a continuum set of replicas, not on each replica as was done in Ref. [21] (see also Refs. [14,24]). This algorithm no longer requires the computation of the Jacobian in generating a configuration, which is expensive for large systems. To overview this new algorithm, we first review the basics of the sign problem, and then introduce our algorithm.

Let us consider a system of an N-dimensional dynamical variable $x \in \mathbb{R}^N$ with an action $S(x)$. Our aim is to evaluate the expectation value of an operator $\mathcal{O}(x)$,

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

where $dx \equiv dx^1 \cdots dx^N$. When the action takes complex values, the Boltzmann weight $e^{-S(x)}/Z$ can no longer be interpreted as a probability distribution, which invalidates a direct application of the Markov-chain Monte Carlo (MCMC) method. The simplest workaround is the so-called reweighting method, where a positive weight is constructed from the real part of the action, $e^{-\mathrm{Re}\, S(x)} / \int_{\mathbb{R}^N} dx\, e^{-\mathrm{Re}\, S(x)}$, and $\langle \mathcal{O}(x) \rangle$ is estimated by a ratio of reweighted averages,

$$\langle \mathcal{O}(x) \rangle = \frac{\langle e^{-i\, \mathrm{Im}\, S(x)} \mathcal{O}(x) \rangle_{\mathrm{rewt}}}{\langle e^{-i\, \mathrm{Im}\, S(x)} \rangle_{\mathrm{rewt}}}, \tag{2}$$

where

$$\langle f(x) \rangle_{\mathrm{rewt}} \equiv \frac{\int_{\mathbb{R}^N} dx\, e^{-\mathrm{Re}\, S(x)} f(x)}{\int_{\mathbb{R}^N} dx\, e^{-\mathrm{Re}\, S(x)}}. \tag{3}$$

For large degrees of freedom (DOF), however, the phase factor $e^{-i\, \mathrm{Im}\, S(x)}$ in reweighted averages makes the integrals highly oscillatory, so that Eq. (2) becomes a ratio of exponentially small quantities of $e^{-O(N)}$ even when the ratio should give a quantity of $O(1)$. Since the reweighting is a mathematically equivalent rewriting, it should not give any problems if one can obtain the values of the reweighted averages precisely both in the numerator and the denominator. However, in the Monte Carlo calculations, they are evaluated separately as sample averages, which should be accompanied by statistical errors of $O(1/\sqrt{N_{\mathrm{conf}}})$, where N_{conf} is the size of the sample. Thus, for the naive reweighting method, the expectation value is estimated in the form

$$\langle \mathcal{O}(x) \rangle \approx \frac{e^{-O(N)} \pm O(1/\sqrt{N_{\mathrm{conf}}})}{e^{-O(N)} \pm O(1/\sqrt{N_{\mathrm{conf}}})}, \tag{4}$$

which means that we need an exponentially large sample size, $N_{\mathrm{conf}} = e^{O(N)}$, in order to make the statistical errors relatively small compared to the exponentially small mean values. This enormous computational time makes the MCMC computations impractical. This is the sign problem.

Lefschetz thimble methods are a class of algorithms towards solving the sign problem. In these methods, we complexify the integration variables, $x \in \mathbb{R}^N \to z \in \mathbb{C}^N$, with the assumption that $e^{-S(z)}$ and $e^{-S(z)}\mathcal{O}(z)$ are entire functions over $\mathbb{C}^N$. Then, from Cauchy's theorem, the integrals do not change under continuous deformations of the integration surface from $\mathbb{R}^N$ to $\Sigma \subset \mathbb{C}^N$, as long as the boundary at infinity ($|x| \to \infty$) is kept fixed under the deformations:

$$\langle \mathcal{O}(x) \rangle = \frac{\int_\Sigma dz\, e^{-S(z)} \mathcal{O}(z)}{\int_\Sigma dz\, e^{-S(z)}}, \tag{5}$$

where $dz \equiv dz^1 \cdots dz^N$. We expect that the sign problem is reduced if we can find an integration surface Σ on which $e^{-i\,\mathrm{Im}\,S(z)}$ is almost constant.

Such surfaces are obtained from the following antiholomorphic gradient flow $z_t(x)$ at large flow times:

$$\frac{dz_t^i}{dt} = [\partial_i S(z_t)]^*, \quad z_{t=0}^i = x^i. \tag{6}$$

In fact, this flow defines a map from the original integration surface $\Sigma_0 \equiv \mathbb{R}^N$ to a real N-dimensional submanifold $\Sigma_t \equiv \{z_t(x)\,|\,x \in \mathbb{R}^N\}$ in $\mathbb{C}^N = \mathbb{R}^{2N}$:

$$z_t \ : \ \Sigma_0 \ni x \mapsto z_t(x) \in \Sigma_t, \tag{7}$$

and the flowed surface Σ_t approaches in the limit $t \to \infty$ a union of Lefschetz thimbles, on each of which $\mathrm{Im}\,S(z)$ is constant.[2] We thus expect that the sign problem is substantially remedied on Σ_t for sufficiently large t. The expectation value is then expressed in the form[3]

$$\langle \mathcal{O}(x) \rangle = \frac{\langle e^{-i\,\mathrm{Im}\,S(z)+i\varphi(z)} \mathcal{O}(z) \rangle_{\Sigma_t}}{\langle e^{-i\,\mathrm{Im}\,S(z)+i\varphi(z)} \rangle_{\Sigma_t}} \tag{8}$$

with

$$\langle f(z) \rangle_{\Sigma_t} \equiv \frac{\int_{\Sigma_t} |dz_t|\, e^{-\mathrm{Re}\,S(z_t)} f(z_t)}{\int_{\Sigma_t} |dz_t|\, e^{-\mathrm{Re}\,S(z_t)}}. \tag{9}$$

Here, $|dz_t|$ is the invariant volume element of Σ_t, and can be expressed with the Jacobian of the flow, $J_t(x) \equiv \partial z_t(x)/\partial x$, as

$$dz_t = \det J_t\, dx, \quad |dz_t| = |\det J_t|\, dx. \tag{10}$$

The phase factor $e^{i\varphi(z)}$ in Eq. (8) is then given by

$$e^{i\varphi(z)} \equiv \frac{dz_t}{|dz_t|} = \frac{\det J_t}{|\det J_t|}. \tag{11}$$

[2] Since $(d/dt)\,S(z_t) = |S(z_t)|^2 \geq 0$, $\mathrm{Im}\,S(z_t)$ is constant along the flow, and $\mathrm{Re}\,S(z_t)$ increases except at critical points z_σ [at which $\partial_i S(z_\sigma) = 0$]. For each critical point z_σ, the Lefschetz thimble $\mathcal{J}_\sigma$ is defined as $\mathcal{J}_\sigma \equiv \{z \in \mathbb{C}^N \,|\, \lim_{t\to-\infty} z_t(z) = z_\sigma\}$, on which $\mathrm{Im}\,S(z)$ is constant [$= \mathrm{Im}\,S(z_\sigma)$].

[3] The original reweighting (2) and (3) corresponds to the $t = 0$ case. When only a single Lefschetz thimble is relevant, one can argue that the exponentially small part in the estimation (4) increases as $e^{-e^{-\lambda t}O(N)}$, where λ is the minimum singular value of $H(z) = (\partial_i \partial_j S(z))$ at the critical point. We thus expect that the sign problem is removed for the flow time $t \gtrsim O(\ln N)$.

The Jacobian $J_t(x)$ can be computed by solving the second flow equation:[4]

$$\frac{dJ_t}{dt} = [H(z_t)J_t]^*, \quad J_{t=0} = 1, \tag{12}$$

where $H(z) \equiv (\partial_i \partial_j S(z))$.

When Σ_t approaches more than one Lefschetz thimble, Σ_t gets decomposed into separate regions as t increases, each region being surrounded by infinitely high potential barriers. This causes a multimodal problem in MCMC calculations.[5] The tempered Lefschetz thimble method (TLTM) was proposed in Ref. [17] to solve this multimodality by implementing the parallel tempering algorithm [25–27] with the flow time t used as the tempering parameter. Namely, we prepare a finite set of flow times, $\{t_\alpha\}$, and introduce copies (replicas) of the corresponding configuration spaces, $\{\Sigma_{t_\alpha}\}$.[6] The set $\{t_\alpha\}$ is chosen so as to include large enough flow times to resolve the sign problem, as well as small enough flow times to resolve the multimodality.[7] Then, in addition to Monte Carlo updates on each Σ_t, we swap configurations between adjacent replicas, which enables easy communications between configurations around different modes, and thus accelerates the relaxation to global equilibrium. Thus, the TLTM is an algorithm that solves the sign and multimodal problems simultaneously, and has proved effective for various models [17,20,21], as mentioned before. However, in this original TLTM, we need to increase the number of replicas as we increase DOF in order to keep sufficient acceptance rates in the swapping process. Furthermore, we have to compute the Jacobian $J_t(x)$ every time in the swapping process, which is expensive because the second flow equation (12) involves a matrix multiplication, whose cost is of $O(N^3)$.

In this paper, we propose a novel hybrid Monte Carlo (HMC) algorithm, where molecular dynamics is performed on a continuum set of integration surfaces, $\bigcup_t \Sigma_t$. This algorithm solves the multimodal problem without preparing replicas. Furthermore, the Jacobian of the gradient flow no longer needs to be computed in generating a configuration, and only its phase needs to be evaluated upon measurement.

This algorithm is based on the observation that, since integrals on Σ_t do not depend on t due to Cauchy's theorem, the values do not change even when we average them over t in a range $[T_0, T_1]$ with an arbitrary weight $e^{-W(t)}$:

$$\langle \mathcal{O}(x) \rangle = \frac{\int_{T_0}^{T_1} dt\, e^{-W(t)} \int_{\Sigma_t} dz_t\, e^{-S(z_t)}\, \mathcal{O}(z_t)}{\int_{T_0}^{T_1} dt\, e^{-W(t)} \int_{\Sigma_t} dz_t\, e^{-S(z_t)}}. \tag{13}$$

We denote the new integration region by $\mathcal{R}$ (see Fig. 1):

$$\mathcal{R} \equiv \bigcup_{t=T_0}^{T_1} \Sigma_t = \left\{ z_t(x) \in \mathbb{C}^N \,\middle|\, t \in [T_0, T_1],\, x \in \mathbb{R}^N \right\}, \tag{14}$$

[4] This can be shown as $(d/dt)(J_t(x))^i{}_a = (\partial/\partial t)(\partial z_t^i(x)/\partial x^a) = \partial[\partial z_t^i(x)/\partial t]/\partial x^a = \partial[\partial_i S(z_t(x))]^*/\partial x^a = [\partial_i \partial_j S(z_t(x))\, (\partial z_t^j(x)/\partial x^a))]^* = [H_{ij}(z_t(x))\, (J_t(x))^j{}_a]^*$.

[5] In the following discussions, we assume that there is no multimodal problem on the original integration surface $\Sigma_{t=0} = \mathbb{R}^N$. If this is not the case, we implement an extra algorithm to resolve the multimodality (such as tempering with respect to the overall coefficient of the action) or make a shift of the starting integration surface from $\mathbb{R}^N$.

[6] Note that Σ_t is not necessarily homeomorphic to $\Sigma_0 = \mathbb{R}^N$ because we remove zeros of $e^{-S(z)}$ from Σ_t (see Ref. [21]).

[7] See Refs. [28,29] for a geometrical optimization of the values t_α based on the distance between configurations introduced in Ref. [30].

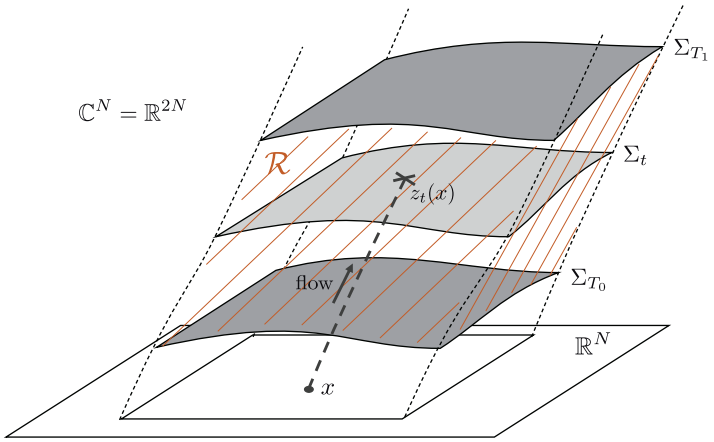

Fig. 1. The worldvolume $\mathcal{R}$ (shaded region) embedded in $\mathbb{C}^N = \mathbb{R}^{2N}$.

which we regard as the *worldvolume* of an integration surface moving in the "target space" $\mathbb{C}^N = \mathbb{R}^{2N}$. We abbreviate the TLTM based on Eq. (13) as the WV-TLTM.
Although the weight function $e^{-W(t)}$ can be chosen arbitrarily, a good choice in practice is one that gives an almost uniform distribution with respect to t (see Sect. 3.4 for details).

The expectation value is now expressed as a ratio of reweighted averages over $\mathcal{R}$:

$$\langle \mathcal{O}(x) \rangle = \frac{\int_{\mathcal{R}} Dz\, e^{-V(z)}\, A(z)\, \mathcal{O}(z)}{\int_{\mathcal{R}} Dz\, e^{-V(z)}\, A(z)} = \frac{\langle A(z)\, \mathcal{O}(z) \rangle_{\mathcal{R}}}{\langle A(z) \rangle_{\mathcal{R}}}. \tag{15}$$

Here, the reweighted average

$$\langle f(z) \rangle_{\mathcal{R}} \equiv \frac{\int_{\mathcal{R}} Dz\, e^{-V(z)} f(z)}{\int_{\mathcal{R}} Dz\, e^{-V(z)}} \tag{16}$$

is defined with respect to the (real-valued) invariant volume element Dz on the $(N+1)$-dimensional region $\mathcal{R}$ and to the new weight[8]

$$e^{-V(z)} \equiv e^{-\mathrm{Re}\, S(z) - W(t(z))}. \tag{17}$$

The accompanying reweighting factor $A(z)$ is then given by

$$A(z) \equiv \frac{e^{-S(z)-W(t(z))} dt\, dz_t}{e^{-V(z)} Dz} = e^{-i\, \mathrm{Im}\, S(z)} \frac{dt\, dz_t}{Dz}. \tag{18}$$

The aim of this paper is to establish an HMC algorithm for the reweighted average (16) on the worldvolume $\mathcal{R}$. To demonstrate that this algorithm works correctly, we apply the algorithm to a chiral random matrix model (the Stephanov model) [31,32], for which the complex Langevin method is known not to work [33].

This paper is organized as follows. In Sect. 2, we review the basics of the HMC algorithm on a general constrained space $\mathcal{R}$. In Sect. 3, we deepen the argument for the case where the constrained surface $\mathcal{R}$ is the worldvolume of an integration surface. We first study the geometry of the worldvolume $\mathcal{R}$ by using the Arnowitt–Deser–Misner parametrization of the metric [34]. We then construct

[8] The function $t = t(z)$ is given by t in $z = z_t(x)$. Later we will extend the defining region from $\mathcal{R}$ to the vicinity of $\mathcal{R}$ in order to define the gradient $\partial_i t(z)$ on $\mathcal{R}$ (see Sect. 3.2).

molecular dynamics on $\mathcal{R}$, with a prescription to determine the weight $e^{-W(t)}$. After summarizing the HMC algorithm on the worldvolume $\mathcal{R}$, we give an algorithm to estimate observables. In Sect. 4, we apply this algorithm to the Stephanov model, and show that our algorithm correctly reproduces the exact results, solving both the sign and multimodal problems. Section 5 is devoted to the conclusion and outlook.

2. HMC algorithm on a constrained space (review)

In this section, we briefly review the basics of the RATTLE algorithm [35,36], which is an HMC algorithm on a constrained space such as our worldvolume $\mathcal{R}$. A detailed discussion is given in Appendix A with more geometrical terms.

2.1. *Stochastic process on a constrained space*

Let $\mathcal{R}$ be an m-dimensional manifold embedded in the flat space $\mathbb{R}^M = \{z = (z^I)\}$ $(I = 1, \dots, M)$. We assume that $\mathcal{R}$ is characterized by $M - m$ independent constraint equations $\phi^r(z) = 0$ $(r = 1, \dots, M-m)$. When $\mathcal{R}$ is the worldvolume of an integration surface, we set $M = 2N$ and $m = N+1$, treating $\mathbb{C}^N$ as a real space $\mathbb{R}^{2N}$.

Denoting the coordinates on $\mathcal{R}$ by $\xi = (\xi^\mu)$ $(\mu = 1, \dots, m)$, the embedding is expressed by functions $z^I = z^I(\xi)$, and the induced metric on $\mathcal{R}$ is given by

$$ds^2 = (dz^I(\xi))^2 \equiv g_{\mu\nu}(\xi)\, d\xi^\mu d\xi^\nu \quad \text{with} \quad g_{\mu\nu}(\xi) = \partial_\mu z^I(\xi)\, \partial_\nu z^I(\xi), \tag{19}$$

which defines the invariant volume element as

$$Dz \equiv \sqrt{g(\xi)}\, d\xi, \tag{20}$$

where $d\xi = d\xi^1 \cdots \xi^m$.

The probability distribution $p(z)$ on $\mathcal{R}$ is defined with respect to Dz, and thus is normalized as $\int_{\mathcal{R}} Dz\, p(z) = 1$. The transition matrix is also defined for Dz, so that a transition from a probability distribution $p(z)$ to $p'(z)$ $(z \in \mathcal{R})$ is expressed with a transition matrix $P(z'|z)$ as

$$p'(z') = \int_{\mathcal{R}} Dz\, P(z'|z)\, p(z) \quad (z' \in \mathcal{R}). \tag{21}$$

For the equilibrium distribution on $\mathcal{R}$ with respect to a potential $V(z)$,

$$p_{\rm eq}(z) \equiv e^{-V(z)}/Z_{\mathcal{R}} \quad \left(Z_{\mathcal{R}} = \int_{\mathcal{R}} Dz\, e^{-V(z)}\right), \tag{22}$$

the detailed balance condition is given by

$$P(z'|z)\, e^{-V(z)} = P(z|z')\, e^{-V(z')} \quad (z, z' \in \mathcal{R}). \tag{23}$$

Throughout this paper, we denote a function on $\mathcal{R}$ by $f(z)$ and $f(\xi)$, interchangeably, with the understanding that $z = z(\xi)$. The transition matrix on $\mathcal{R}$ is also written as $P(z'|z)$ and $P(\xi'|\xi)$ for $z = z(\xi)$, $z' = z(\xi') \in \mathcal{R}$.

2.2. *HMC on a constrained space*

Denoting by $\pi = (\pi^I)$ the conjugate momentum to $z = (z^I) \in \mathcal{R}$, we consider the Hamiltonian dynamics on $\mathcal{R}$ with the Hamiltonian

$$H(z, \pi) = \frac{1}{2}(\pi^I)^2 + V(z), \tag{24}$$

which can be expressed as a set of first-order differential equations in time s with Lagrange multipliers λ_r:

$$\partial_s z = \pi, \tag{25}$$

$$\partial_s \pi = -\partial V(z) - \lambda_r\, \partial\phi^r(z), \tag{26}$$

$$\phi^r(z) = 0, \tag{27}$$

$$\pi \cdot \partial\phi^r(z) = 0. \tag{28}$$

Here, $\partial \equiv (\partial_{z^I})$ is the gradient in $\mathbb{R}^M$.

Equations (25)–(28) can be discretized such that the symplecticity and the reversibility still hold after the discretization (below, Δs is the step size) [35,36]:

$$\pi_{1/2} = \pi - \frac{\Delta s}{2}\, \partial V(z) - \lambda_r\, \partial\phi^r(z), \tag{29}$$

$$z' = z + \Delta s\, \pi_{1/2}, \tag{30}$$

$$\pi' = \pi - \frac{\Delta s}{2}\, \partial V(z') - \lambda'_r\, \partial\phi^r(z'), \tag{31}$$

where λ_r and λ'_r are determined, respectively, so that the following constraints are satisfied:[9]

$$z' \in \mathcal{R} \quad (\text{i.e., } \phi^r(z') = 0), \tag{32}$$

$$\pi' \in T_{z'}\mathcal{R}. \tag{33}$$

One can easily show that the map $\Phi_{\Delta s} : (z, \pi) \to (z', \pi')$ actually satisfies the symplecticity and the reversibility (with λ_r and λ'_r interchanged):[10]

- $\omega(z', \pi') = \omega(z, \pi)$, (34)
- $(z', \pi') = \Phi_{\Delta s}(z, \pi) \Rightarrow (z, -\pi) = \Phi_{\Delta s}(z', -\pi')$. (35)

The Hamiltonian is conserved to the order of Δs^2, i.e., $H(z', \pi') - H(z, \pi) = O(\Delta s^3)$.

The HMC algorithm on $\mathcal{R}$ then consists of the following three steps for a given initial configuration $z \in \mathcal{R}$:

Step 1. Generate a vector $\tilde{\pi} = (\tilde{\pi}^I) \in T_z\mathbb{R}^M$ from the Gaussian distribution

$$\frac{1}{(2\pi)^{M/2}} e^{-\tilde{\pi}^2/2}, \tag{36}$$

and project it onto $T_z\mathcal{R}$ to obtain an initial momentum $\pi = (\pi^I) \in T_z\mathcal{R}$.

Step 2. Calculate $\Phi_{\Delta s}(z, \pi)$ from Eqs. (29)–(33). We repeat this step n times to obtain $(z', \pi') = \Phi^n_{\Delta s}(z, \pi)$.

Step 3. Update the configuration z to z' with a probability

$$\min\left(1, e^{-H(z',\pi')+H(z,\pi)}\right). \tag{37}$$

[9] We regard the tangent bundle $T\mathcal{R} = \bigcup_z T_z\mathcal{R}$ (*not* the cotangent bundle $T^*\mathcal{R}$) as the phase space for motions on $\mathcal{R}$ [35,36]. See Appendix A for details.

[10] $\omega(z, \pi) \equiv d\pi^I \wedge dz^I\big|_{T\mathcal{R}}$ is the induced symplectic form for the embedding of the phase space $T\mathcal{R}$ into $T\mathbb{R}^M$ (see Appendix A).

The above process defines a stochastic process on $\mathcal{R}$. One can show that its transition matrix $P(z'|z)$ satisfies the detailed balance condition (see Appendix A):

$$P(z'|z)\, e^{-V(z)} = P(z|z')\, e^{-V(z')}. \tag{38}$$

3. HMC on the worldvolume

In this section, we apply the general formalism in the previous section to the case where $\mathcal{R}$ is the worldvolume of an integration surface. We first clarify the geometry of the worldvolume $\mathcal{R}$ and then construct the HMC algorithm on $\mathcal{R}$.

3.1. Geometry of the worldvolume $\mathcal{R}$

Recall that our worldvolume $\mathcal{R}$ is an $N+1$-dimensional submanifold embedded in $\mathbb{C}^N = \mathbb{R}^{2N}$. As in the previous section, we again assume that $\mathcal{R} \subset \mathbb{R}^{2N}$ is characterized by a set of independent constraint equations: $\phi^r(z) = 0$ $(r = 1, \dots, N-1)$. We will often write a point $z = (z^i) \in \mathbb{C}^N$ $(i = 1, \dots, N)$ with real coordinates as[11]

$$z = (z^I) = \begin{pmatrix} \mathrm{Re}\, z^i \\ \mathrm{Im}\, z^i \end{pmatrix} \in \mathbb{R}^{2N} \quad (I = 1, \dots, 2N). \tag{39}$$

Since $\xi = (\xi^\mu) \equiv (t, x^a)$ specifies a point $z(\xi) = (z^I(\xi)) \equiv (z^I_t(x))$ in $\mathcal{R}$, we can use ξ as coordinates of $\mathcal{R}$. The flow equation (6) then takes the form

$$\frac{\partial z^I}{\partial t} = \partial_{z^I} \mathrm{Re}\, S(z), \quad (z^I)|_{t=0} = \begin{pmatrix} x^i \\ 0 \end{pmatrix}. \tag{40}$$

Similarly, we write an N-dimensional complex vector $v = (v^i) \in \mathbb{C}^N$ $(i = 1, \dots, N)$ as a real vector[12]

$$v = (v^I) = \begin{pmatrix} \mathrm{Re}\, v^i \\ \mathrm{Im}\, v^i \end{pmatrix} \in \mathbb{R}^{2N} \quad (I = 1, \dots, 2N). \tag{41}$$

The vectors $E_\mu = (E^I_\mu = \partial z^I / \partial \xi^\mu)$ form a basis of $T_z\mathcal{R}$ (see Fig. 2), from which the induced metric $g_{\mu\nu}$ on $\mathcal{R}$ is given by

$$g_{\mu\nu} = E_\mu \cdot E_\nu. \tag{42}$$

Since our worldvolume $\mathcal{R} = \bigcup_t \Sigma_t$ is foliated by the antiholomorphic gradient flow, its intrinsic geometry should be best described (at least for physicists) by the Arnowitt–Deser–Misner (ADM) parametrization [34], for which the metric is expressed in the form

$$ds^2 = g_{\mu\nu}(\xi)\, d\xi^\mu d\xi^\nu = \alpha^2(\xi)\, dt^2 + \gamma_{ab}(\xi)\, (dx^a + \beta^a(\xi) dt)(dx^b + \beta^b(\xi) dt). \tag{43}$$

[11] We use the same symbol for both complex and real coordinates to avoid a mess of many symbols. We will clarify when one needs to specify which coordinates are implied.

[12] We define the inner product of two real vectors $u = (u^I)$, $v = (v^I) \in \mathbb{R}^{2N}$ by $u \cdot v \equiv u^T v = u^I v^I$. In terms of complex vectors $u = (u^i)$, $v = (v^i) \in \mathbb{C}^N$, the inner product is expressed as $\mathrm{Re}\,(u^\dagger v)$. We do not distinguish the upper and lower indices for I.

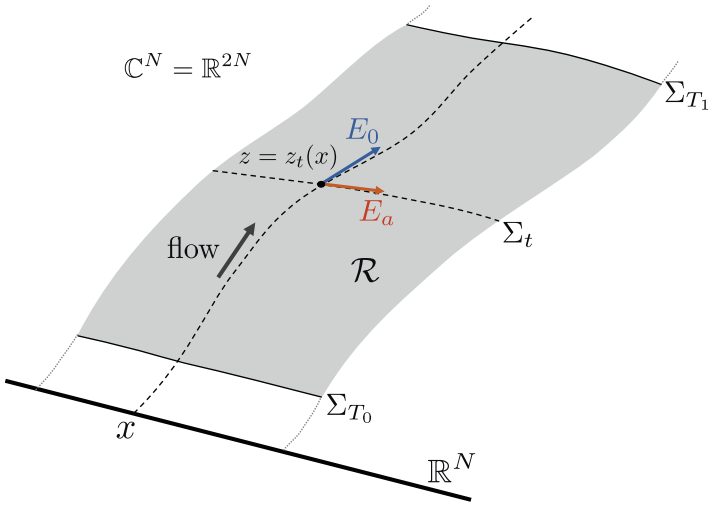

Fig. 2. A basis of $T_z\mathcal{R}$, $\{E_\mu\} = \{E_0, E_a\}$. $\mathcal{R}$ is drawn one dimension less than in Fig. 1.

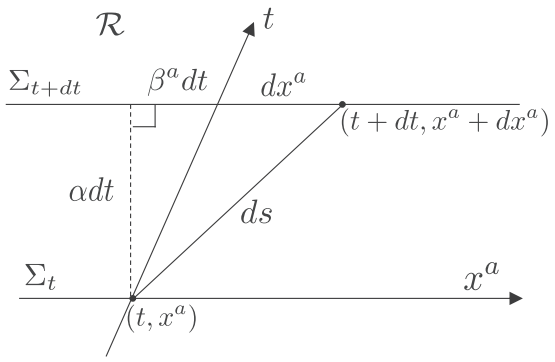

Fig. 3. Geometrical meaning of the ADM parametrization. $\alpha\, dt$ is the geodesic distance from a point $\xi = (t, x^a)$ on Σ_t to the surface Σ_{t+dt}. $\beta^a dt$ shows that the time axis t is tilted with respect to the normal of Σ_t by this amount in the x coordinates. γ_{ab} is the induced metric on Σ_t. The geodesic distance ds between two points $\xi = (t, x^a)$ and $\xi + d\xi = (t + dt, x^a + dx^a)$ is then obtained from the Pythagorean theorem as in Eq. (43).

Here, γ_{ab} is the induced metric on Σ_t (with its inverse matrix γ^{ab}), β^a is the shift vector, and α is the lapse function:

$$\gamma_{ab} \equiv E_a \cdot E_b, \tag{44}$$

$$\beta^a \equiv \gamma^{ab}\, \beta_b \equiv \gamma^{ab}\, E_0 \cdot E_b, \tag{45}$$

$$\alpha^2 \equiv E_0 \cdot E_0 - \gamma^{ab}\, \beta_a \beta_b = E_0 \cdot (1 - \gamma^{ab}\, E_a E_b^T)\, E_0. \tag{46}$$

The inverse matrix of $(g_{\mu\nu})$ can be easily calculated to be

$$(g^{\mu\nu}) = \begin{pmatrix} 1/\alpha^2 & -\beta^b/\alpha^2 \\ -\beta^a/\alpha^2 & \beta^a\beta^b/\alpha^2 + \gamma^{ab} \end{pmatrix}. \tag{47}$$

The geometrical meaning of the ADM parametrization is explained in Fig. 3.

Since the flowed surfaces Σ_t are determined by the flow equation, we can write down the explicit form of the basis $E_\mu = (E^I_\mu = \partial z^I/\partial\xi^\mu)$ of $T_z\mathcal{R}$ as

$$E^I_0 = \partial_{z^I} \mathrm{Re}\, S(z), \quad E^I_a = \begin{pmatrix} \mathrm{Re}\, J^i_a \\ \mathrm{Im}\, J^i_a \end{pmatrix}, \tag{48}$$

where we have defined $J(\xi) = J(t,x) \equiv J_t(x)$. Thus, the induced metric γ_{ab} can be directly expressed in terms of the Jacobian as[13]

$$\gamma_{ab} = \mathrm{Re}\,(J^\dagger J)_{ab} = (J^\dagger J)_{ab}. \tag{49}$$

The lapse function α can be expressed as the length of the normal component of E_0:

$$\alpha^2 = E^\perp_0 \cdot E^\perp_0 = (E^\perp_0)^2. \tag{50}$$

Here, the decomposition of E_0 to the tangential and normal components is given by

$$E_0 = E^\parallel_0 + E^\perp_0, \tag{51}$$

$$E^\parallel_0 \equiv \Pi_{\Sigma_t} E_0, \quad E^\perp_0 \equiv (1 - \Pi_{\Sigma_t}) E_0, \tag{52}$$

where Π_{Σ_t} is the projector onto $T_z\Sigma_t$:[14]

$$\Pi_{\Sigma_t} \equiv \gamma^{ab} E_a E_b^T. \tag{53}$$

Note that the following relation holds:

$$\beta^a E_a = E^\parallel_0. \tag{54}$$

Then, by using Eqs. (47), (50), (53), and (54), we see that the projector from $T_z\mathbb{R}^{2N}$ onto $T_z\mathcal{R}$ is given by

$$\Pi_{\mathcal{R}} \equiv g^{\mu\nu} E_\mu E_\nu^T = \frac{1}{(E^\perp_0)^2} E^\perp_0 (E^\perp_0)^T + \Pi_{\Sigma_t}. \tag{55}$$

In the ADM parametrization, the volume element of $\mathcal{R}$ is given by (see Fig. 4)

$$Dz = \sqrt{g}\, d\xi = \alpha\, |\det J|\, dt\, dx. \tag{56}$$

Since the complex measure on Σ_t is given by $dz_t = \det J\, dx$, we find that the reweighting factor (18) takes the form

$$A(z) = \frac{dt\, dz_t}{Dz} e^{-i\,\mathrm{Im}\, S(z)} = \alpha^{-1}(z)\, e^{i\varphi(z) - i\,\mathrm{Im}\, S(z)} \quad \left(e^{i\varphi(z)} \equiv \frac{\det J}{|\det J|}\right). \tag{57}$$

Note that the inverse lapse, $\alpha^{-1}(z)$, plays the role of the radius of $A(z)$.

[13] The second equality is a direct consequence of the identity $\mathrm{Im}\,(J^\dagger J) = 0$, which can be proved by a differential equation $(d/dt)\,\mathrm{Im}\,(J^\dagger J) = 0$ [as can be shown from Eq. (12)] with the initial condition $\mathrm{Im}\,(J^\dagger J)|_{t=0} = 0$.

[14] In actual calculations, we do not need the explicit form of Π_{Σ_t} in projecting vectors in $T_z\mathbb{R}^{2N}$ onto $T_z\mathcal{R}$ (see Sect. 3.2), and thus do not have to calculate the Jacobian J. A similar statement can be applied to the expressions below.

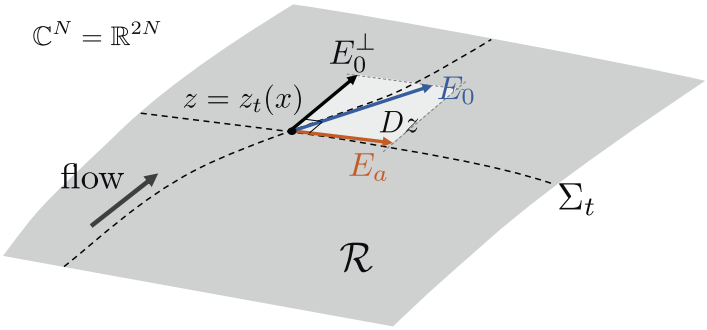

Fig. 4. $E_0^\perp$ and the volume element.

In Appendix B, in order to show the typical behaviors of various geometrical quantities near critical points at large flow times, we give explicit expressions of these quantities for the Gaussian case with the action

$$S(x) = \frac{\beta}{2}\sum_{k=1}^{N}(x^k - i)^2. \tag{58}$$

There, we find that $\alpha^{-1}(z)$ increases exponentially in flow time t as $z^k = z^k(t,x)$ approaches the Lefschetz thimble at $\mathrm{Im}\, z^k = 1$.

3.2. *Molecular dynamics on the worldvolume*

We first rewrite the Lagrange multiplier term $\lambda_r\, \partial\phi^r$ in Eq. (26) in a more convenient form. Note that $\lambda_r\, \partial\phi^r$ is normal to $\mathcal{R}$, and thus it satisfies

$$(\lambda_r\, \partial\phi^r) \cdot E_0 = 0, \quad (\lambda_r\, \partial\phi^r) \cdot E_a = 0. \tag{59}$$

Since the vectors

$$F_a \equiv \begin{pmatrix} \mathrm{Im} J^i{}_a \\ -\mathrm{Re} J^i{}_a \end{pmatrix} \quad (a = 1, \ldots, N) \tag{60}$$

span the normal vector space $N_z\Sigma_t$ at $z \in \Sigma_t$,[15] the second equation in Eq. (59) means that $\lambda_r\, \partial\phi^r$ can be written as a linear combination of F_a with new Lagrange multipliers $\lambda^a \in \mathbb{R}$ $(a = 1, \ldots, N)$:[16]

$$\sum_{r=1}^{N-1} \lambda_r\, \partial\phi^r = \sum_{a=1}^{N} \lambda^a F_a. \tag{61}$$

The first equation in Eq. (59) is then treated as a constraint on $\lambda^a F_a$:[17]

$$\lambda^a F_a \cdot E_0 = 0. \tag{62}$$

[15] F_a form a basis of the normal space $N_z\Sigma_t$ because $E_a \cdot F_b = -\mathrm{Im}\,(J^\dagger J)_{ab} = 0$ (see footnote 13). They can be written as $F_a = iE_a$ as complex vectors.

[16] We put the summation symbols here to stress the summation ranges for r and a.

[17] It is possible to solve the constraint (62) as follows. We first take a subset $\{F_r\}$ $(r = 1, \ldots, N-1)$ of $\{F_a\}$, whose elements are not parallel to E_0. Then, we construct a basis of $N_z\mathcal{R}$ by $\tilde{F}_r \equiv F_r - (F_r \cdot E_0/(E_0)^2)E_0$, and replace $\lambda^a F_a$ in Eq. (62) by $\mu^r \tilde{F}_r$ with new Lagrange multipliers $(\mu^r) \in \mathbb{R}^{N-1}$.

From the above argument, we find that the RATTLE algorithm (29)–(33) can be written as

$$\pi_{1/2} = \pi - \frac{\Delta s}{2}\,\partial V(z) - \lambda^a F_a(z), \tag{63}$$

$$z' = z + \Delta s\,\pi_{1/2}, \tag{64}$$

$$\pi' = \pi - \frac{\Delta s}{2}\,\partial V(z') - \lambda'^a F_a(z'), \tag{65}$$

where λ^a and λ'^a are determined, respectively, so that the following constraints are satisfied:

$$z' \in \mathcal{R} \quad \text{and} \quad \lambda^a F_a(z) \cdot E_0(z) = 0, \tag{66}$$

$$\pi' \in T_{z'}\mathcal{R} \quad \text{and} \quad \lambda'^a F_a(z') \cdot E_0(z') = 0. \tag{67}$$

The gradient of the potential, $\partial V(z)$, now takes the form

$$\partial V(z) = \partial \mathrm{Re}\, S(z) + W'(t(z))\,\partial t(z). \tag{68}$$

In order to define the gradient $\partial t(z)$ at $z \in \mathcal{R}$, we regard (ϕ^r) as coordinates in the extra dimensions and construct $2N$ coordinates (ξ^μ, ϕ^r) in the vicinity of $\mathcal{R}$ in $\mathbb{R}^{2N}$. Then, one can show (see Appendix C) that the gradient $\partial t(z)$ is given by

$$\partial t(z) = \frac{1}{(E_0^\perp)^2}\,E_0^\perp - \frac{E_r \cdot E_0^\perp}{(E_0^\perp)^2}\,\partial\phi^r(z) \quad (z \in \mathcal{R}). \tag{69}$$

Since the last term is a linear combination of the gradients $\partial\phi^r(z)$ and thus can be absorbed into the Lagrange multiplier terms in Eqs. (63) and (65), we can (and will) set $\partial t(z)$ to the form

$$\partial t(z) = \frac{1}{(E_0^\perp)^2}\,E_0^\perp \quad (z \in \mathcal{R}). \tag{70}$$

3.3. Solving the constraints in molecular dynamics

In this subsection, we present numerical algorithms to solve the constraints (66) and (67).

3.3.1. Solving Eq. (66)

The condition $z' \in \mathcal{R}$ for a given $z = z(\xi) \in \mathcal{R}$ is equivalent to the existence of an $N+1$-dimensional vector $\varepsilon = (\varepsilon^\mu) = (h, u^a)$ such that $z' = z(\xi + \varepsilon) = z_{t+h}(x + u)$. Thus, Eq. (66) can be solved by finding a solution to $2N + 1$ equations

$$f^P(w) = 0 \quad (P = 0, 1, \ldots, 2N) \tag{71}$$

for $2N + 1$ unknowns $w = (w^P) = (\varepsilon^\mu, \lambda^a) \in \mathbb{R}^{2N+1}$ (see Fig. 5), where

$$f^0(w) \equiv \lambda^a\, F_a(\xi) \cdot E_0(\xi) = -\lambda^a\, \mathrm{Im}\big[\partial_{z^i} S(z(\xi))\, J^i{}_a(\xi)\big], \tag{72}$$

$$f^I(w) \equiv z^I(\xi + \varepsilon) - z^I(\xi) - \Delta z^I + \lambda^a F_a^I(\xi) \quad (I = 1, \ldots, 2N) \tag{73}$$

with

$$\Delta z \equiv \Delta s\,\pi - \frac{\Delta s^2}{2}\,\partial V(z). \tag{74}$$

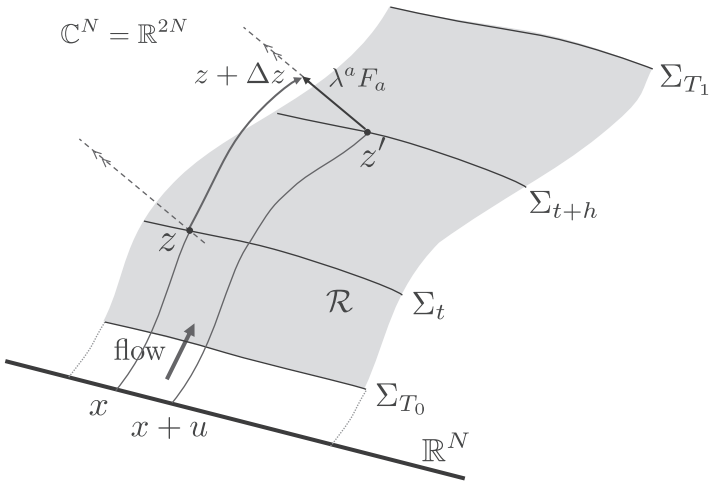

Fig. 5. Finding $z' = z(\xi + \varepsilon) = z_{t+h}(x + u)$ on $\mathcal{R}$ from $z = z(\xi) = z_t(x) \in \mathcal{R}$.

Adopting Newton's method to find a solution, we iteratively update the vector $w = (w^P) = (\varepsilon^\mu, \lambda^a)$ as $w \to w + \Delta w$, where Δw is a solution of the linear equation

$$\frac{\partial f^P(w)}{\partial w^Q} \Delta w^Q = -f^P(w). \tag{75}$$

The matrix elements $\partial f^P / \partial w^Q$ are easily found to be

$$\begin{aligned}\left(\frac{\partial f^P(w)}{\partial w^Q}\right) &= \begin{bmatrix} 0 & 0 & F_a(\xi) \cdot E_0(\xi) \\ E_0^I(\xi + \varepsilon) & E_a^I(\xi + \varepsilon) & F_a^I(\xi) \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & -\mathrm{Im}\left[\partial_{z^i} S(z(\xi))\, J^i{}_a(\xi)\right] \\ \mathrm{Re}\left[\partial_{z^i} S(z(\xi + \varepsilon))\right] & \mathrm{Re}\, J^i{}_a(\xi + \varepsilon) & -\mathrm{Im}\, J^i{}_a(\xi) \\ -\mathrm{Im}\left[\partial_{z^i} S(z(\xi + \varepsilon))\right] & \mathrm{Im}\, J^i{}_a(\xi + \varepsilon) & \mathrm{Re}\, J^i{}_a(\xi) \end{bmatrix}. \end{aligned} \tag{76}$$

When the DOF ($= N$) is small, Eq. (75) can be solved by a direct method such as the LU decomposition of Eq. (76), for which we integrate the second flow equation (12) to know the value of J. When the DOF is large, the computation of J becomes expensive, and we can instead use an iterative method such as GMRES [37] or BiCGStab [38], for which we do not need to compute the matrix elements of J as in Ref. [19]. To see this, we first note that the right-hand side of Eq. (75) can be written in terms of complex vectors as

$$f^0(w) = -\mathrm{Im}\left[\partial_{z^i} S(z(\xi))\, J^i{}_a(\xi) \lambda^a\right], \tag{77}$$

$$f^i(w) = z^i(\xi + \varepsilon) - z^i(\xi) - \Delta z^i + i J^i{}_a(\xi) \lambda^a. \tag{78}$$

The left-hand side of Eq. (75) can also be written as

$$\frac{\partial f^0(w)}{\partial w^N} \Delta w^N = -\mathrm{Im}\left[\partial_{z^i} S(z(\xi))\, J^i{}_a(\xi) \Delta\lambda^a\right], \tag{79}$$

$$\frac{\partial f^i(w)}{\partial w^N} \Delta w^N = [\partial_{z^i} S(z(\xi + \varepsilon))]^* + J^i{}_a(\xi + \varepsilon) \Delta u^a + i J^i{}_a(\xi) \Delta\lambda^a. \tag{80}$$

We thus see that, in the above equations, J appears only in the form $J^i{}_a(\xi) v^a$ or $J^i{}_a(\xi + \varepsilon) v^a$ with a real vector $v = (v^a) \in \mathbb{R}^N$. The former can be evaluated from the solution to the flow equation (6)

and the following equation [see Eq. (12)]:

$$\dot{v}_t^i = [H_{ij}(z_t)\, v_t^j]^*, \quad v_{t=0}^i = v^i, \tag{81}$$

by setting $J^i{}_a(\xi)v^a = v_t^i$. The latter is obtained in a similar way, by replacing $t \to t+h$ and $x^a \to x^a + u^a$. We thus find that both sides of Eq. (75) can be calculated without computing the matrix elements of J.

3.3.2. *Solving Eq. (67)*

We first note that solving the constraint (67) is equivalent to projecting the vector

$$\tilde{\pi}' \equiv \pi - \frac{\Delta s}{2}\, \partial V(z') \tag{82}$$

onto $T_{z'}\mathcal{R}$ [see Eq. (55)]:

$$\pi' \equiv \Pi_{\mathcal{R}}(z')\, \tilde{\pi}' = \frac{E_0^\perp(z') \cdot \tilde{\pi}'}{(E_0^\perp(z'))^2}\, E_0^\perp(z') + \Pi_{\Sigma_t}(z')\, \tilde{\pi}'. \tag{83}$$

Here, the second term can be computed as a complex vector as in Ref. [14] (see also Refs. [19,21]):

$$\Pi_{\Sigma_t}(z')\, \tilde{\pi}' = J(z')\, \mathrm{Re}\left(J^{-1}(z')\, \tilde{\pi}'\right). \tag{84}$$

One can also show that the normal vector $E_0^\perp(z')$ can be computed as a complex vector to be

$$E_0^\perp(z') = iJ(z')\, \mathrm{Im}\left(J^{-1}(z')\, [\partial S(z')]^*\right). \tag{85}$$

The expressions (84) and (85) can again be evaluated either by a direct method with the computation of $J(z')$, or by an iterative method without computing $J(z')$ as in Ref. [19]. When the iterative method is used, the inversion $J^{-1}(z')\, c$ is obtained for a given complex vector $c = (c^i) \in \mathbb{C}^N$ by looking for vectors $a = (a^a)$, $b = (b^a) \in \mathbb{R}^N$ iteratively such that

$$c = J(z')\, a + iJ(z')\, b, \tag{86}$$

where $J(z')\, a$ and $J(z')\, b$ are evaluated by integrating the flow equation (81) with the initial conditions $v_{t=0} = a$ and $v_{t=0} = b$, respectively. The multiplication $J(z')\, v$ is again calculated by integrating Eq. (81). Therefore, the projection (83) can be performed without computing $J(z')$.

Note that, every time we evaluate $\partial V(z)$ [Eq. (68)], we need $\partial t(z)$ and thus $E_0^\perp(z)$ [see Eq. (70)]. $E_0^\perp(z)$ can be calculated from Eq. (85) by replacing z' with z.

3.4. *Construction of $W(t)$*

In this subsection, we present a prescription to construct a weight function $e^{-W(t)}$ in Eq. (13) [or in Eq. (17)] so that it gives an almost uniform distribution with respect to t. The key is that, for a given weight $e^{-W(t)}$, the probability of finding a configuration at t is proportional to

$$Z(t; W) \equiv e^{-W(t)} \int_{\mathbb{R}^N} dx\, \alpha(t,x) |\det J(t,x)| e^{-\mathrm{Re}\, S(z(t,x))}. \tag{87}$$

Thus, when a weight $e^{-W(t)}$ does not give a uniform distribution of t, the desired weight can be obtained by (see, e.g., Ref. [39])

$$W^{(\mathrm{new})}(t) = W(t) + \ln Z(t; W) + \mathrm{const}, \tag{88}$$

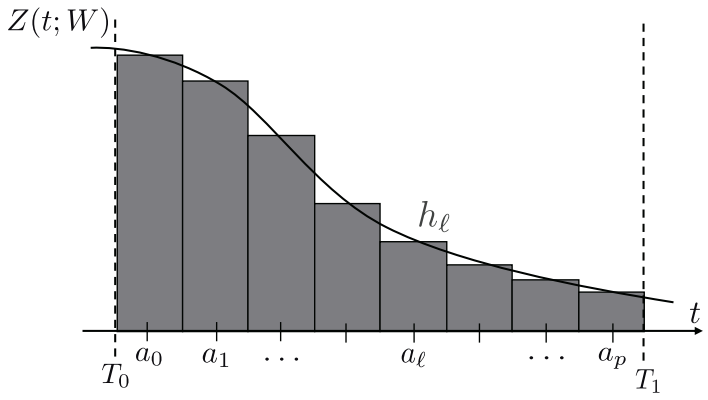

Fig. 6. Estimating $Z(t;W)$ from $\{h_\ell\}$.

because then $Z(t;W^{(\mathrm{new})})$ becomes constant in t:

$$\begin{aligned} Z(t;W^{(\mathrm{new})}) &= e^{-W^{(\mathrm{new})}(t)} \int_{\mathbb{R}^N} dx\, \alpha(t,x)\, |\det J(t,x)|\, e^{-\mathrm{Re}\, S(z(t,x))} \\ &= \mathrm{const}\, \frac{e^{-W(t)}}{Z(t;W)} \cdot e^{W(t)}\, Z(t;W) = \mathrm{const}. \end{aligned} \tag{89}$$

Of course, the above procedure is possible only when we know the values of $Z(t;W)$ explicitly, which is usually not the case. However, we can estimate $Z(t;W)$ from the histogram of flow times $\{t\}$. To be more specific, we first divide the interval $[T_0,T_1]$ into $p+1$ bins, $I_\ell \equiv [\ell h,(\ell+1)h]$ $(\ell=0,\ldots,p)$ with $h \equiv (T_1-T_0)/(p+1)$, and generate a certain number ($\equiv n_{\mathrm{tune}}$) of configurations by using $V(z) = \mathrm{Re}\, S(z) + W(t(z))$ as the potential. The numbers h_ℓ of configurations inside the bins I_ℓ give a rough estimate of the functional form of $Z(t;W)$ up to a normalization factor (see Fig. 6). Then, from the histogram h_ℓ $(\ell=0,\ldots,p)$, we calculate

$$W_\ell^{(\mathrm{new})} = W(a_\ell) + \ln h_\ell, \tag{90}$$

and construct a function $W^{(\mathrm{new})}(t)$ to be approximated by a polynomial satisfying $W^{(\mathrm{new})}(a_\ell) = W_\ell^{(\mathrm{new})}$ with $a_\ell \equiv (\ell+1/2)h$.

In general, the minimum-order polynomial that has values b_ℓ at $a_\ell = (\ell+1/2)h$ is given by a Lagrange interpolation of the form

$$L(t;\{b_\ell\}) \equiv \sum_{\ell=0}^{p} \frac{\prod_{m\neq\ell}(t-a_m)}{\prod_{m\neq\ell}(a_\ell-a_m)}\, b_\ell = \sum_{\ell=0}^{p} \Big[\prod_{m<\ell}(t-a_\ell)\Big] \frac{\Delta^\ell b_0}{\ell!\, h^\ell}, \tag{91}$$

where, for an array $\{v_0,v_1,v_2,\ldots\}$, we define $\Delta v_\ell \equiv v_{\ell+1} - v_\ell$, so that

$$\Delta^\ell v_0 = \sum_{k=0}^{\ell} (-1)^k \binom{\ell}{k} v_{\ell-k}. \tag{92}$$

Using this polynomial, we define[18]

$$W^{(\mathrm{new})}(t) \equiv L(t; \{W_\ell^{(\mathrm{new})}\}). \tag{93}$$

Since the estimate of $Z(t; W)$ from the histogram $\{h_\ell\}$ includes statistical errors, we use an iterative algorithm to update $\{W_\ell\}$ until an almost uniform distribution is obtained:

- Initialize $\{W_\ell\}$ with appropriate values (e.g., $W_\ell^{(0)} = 0$).
- From an array $\{W_\ell^{(k)}\}$, construct an order-$(p+1)$ polynomial $L(t; \{W_\ell^{(k)}\})$, and set $W^{(k)}(t) \equiv L(t; \{W_\ell^{(k)}\})$.
- Generate n_{tune} configurations with the potential $V(z) = \mathrm{Re}\, S(z) + W^{(k)}(t(z))$, and record the numbers $h_\ell^{(k)}$ of configurations in the intervals I_ℓ.
- Update $\{W_\ell\}$ as $W_\ell^{(k+1)} \equiv W_\ell^{(k)} + \ln(h_\ell^{(k)} + \epsilon_c)$. Here, ϵ_c is a cutoff to avoid the divergence arising when $h_\ell^{(k)} = 0$.
- Terminate the iteration when the histogram becomes almost flat. We use the following stopping condition:

$$\frac{1}{p}\sum_{l=0}^{p-1}\left[\frac{h_{l+1} - h_l}{(h_{l+1} + h_l)/2}\right]^2 < \delta^2. \tag{94}$$

In the calculation below, we set $p = 7$, $n_{\mathrm{tune}} = (p+1) \times 200 = 1,600$, $\epsilon_c = 0.01$, and $\delta^2 = 0.2$.

3.5. *Summary of the HMC algorithm on the worldvolume*

We summarize the HMC algorithm for a given initial configuration $z \in \mathcal{R}$.

Step 1. Generate $\tilde{\pi} = (\tilde{\pi}^I)$ from the Gaussian distribution, and project it onto $T_z\mathcal{R}$ to obtain an initial momentum $\pi = (\pi^I)$: $\pi = \Pi_{\mathcal{R}}(z)\,\tilde{\pi}$.

Step 2. Calculate $(z, \pi) \to \Phi_{\Delta s}(z, \pi)$ with Eqs. (63)–(67). The gradient $\partial V(z)$ takes the form (68), where $\partial t(z)$ is given by Eq. (70) and $W(t)$ is determined from test runs by using the iterative algorithm given in Sect. 3.4. The first constraint (66) is solved by finding a root of the functions (72) and (73), and the second constraint (67) is solved by calculating Eqs. (82) and (83). We repeat this process n times to obtain $(z', \pi') = \Phi_{\Delta s}^n(z, \pi)$.

Step 3. Update the configuration z to z' with a probability

$$\min\left(1, e^{-H(z',\pi')+H(z,\pi)}\right). \tag{95}$$

Upon measurement, we further compute the reweighting factor $A(z)$ [see Eq. (57)], which requires the phase $e^{i\varphi(z)} = \det J(z)/|\det J(z)|$, which is evaluated by first solving Eq. (12) to get $J(z)$ and then computing its determinant. The lapse function $\alpha(z) = |E_0^\perp(z)|$ is already obtained in the preceding molecular dynamics step (Step 2).

[18] In the calculation below, we put two additional terms in Eq. (91) to prevent the function from changing drastically near boundaries. The above $L(t; \{b_\ell\})$ is then replaced by

$$L(t; \{b_\ell\}) \equiv (\mathrm{const}\, t + \mathrm{const}) \prod_{\ell=0}^{p}(t - a_\ell) + \sum_{\ell=0}^{p}\left[\prod_{m<\ell}(t - a_\ell)\right]\frac{\Delta^\ell b_0}{\ell!\, h^\ell}.$$

The constants are determined by the conditions $L'(T_0; \{W_\ell\}) = c_0$ and $L'(T_1; \{W_\ell\}) = c_1$. We set $c_0 = 1.2 \times \min_\ell(0, \Delta W_\ell/h)$ and $c_1 = 0.01$ in the calculation below.

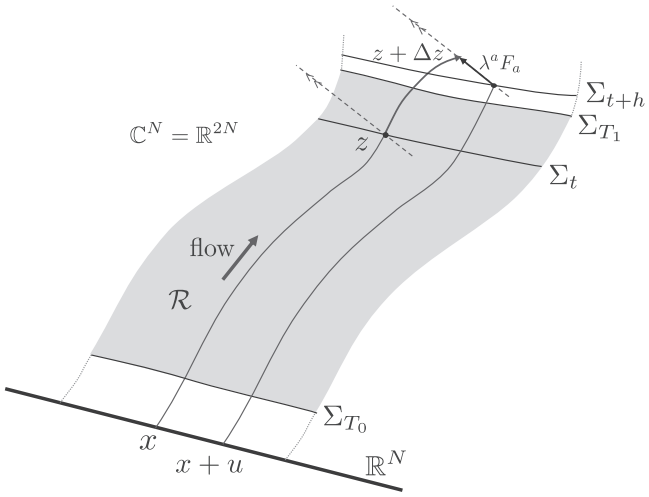

Fig. 7. A transition passing over the boundary at T_1.

In the course of molecular dynamics (Step 2), it sometimes happens that the equation $(z', \pi') = \Phi_{\Delta s}(z, \pi)$ does not have a solution because $z + \Delta z$ passes over the boundary of $\mathcal{R}$ (see Fig. 7). Here, the boundary in the t direction is given by T_0, T_1, while that in the x direction is given by zeros of $e^{-S(z)}$. When a solution is not found near a boundary, we replace the operator $\Phi_{\Delta s}$ by the *momentum reflection* $\mathcal{T}$:

$$\mathcal{T}(z, \pi) \equiv (z, \pi'), \tag{96}$$

where for π, which is expanded with $E^\mu \equiv g^{\mu\nu} E_\nu$ as $\pi = \eta_0 E^0 + \eta_a E^a$, the reflected momentum π' is defined by $\pi' \equiv -\eta_0 E^0 + \eta_a E^a = \pi - 2\eta_0 E^0$, i.e.,

$$\pi' \equiv \pi - 2 \frac{E_0 \cdot \pi}{E_0^{\perp 2}} E_0^\perp. \tag{97}$$

This preserves the reversibility and the phase-space volume element because the induced symplectic form is given by $\omega = d\eta_\mu \wedge d\xi^\mu = d\eta_0 \wedge d\xi^0 + d\eta_a \wedge d\xi^a$ (see Appendix A). However, this can change the value of the Hamiltonian. The change comes only from the difference between the norms of momenta π and π', and its effect is absorbed in the probability at the Metropolis test in Step 3 above, so that the detailed balance condition (38) still holds. If the change is larger than a prescribed value (e.g., if $e^{-|\Delta H|} = e^{-|\pi'^2 - \pi^2|/2} < 0.8$), we instead use the *momentum flip* Ψ [21]:

$$\Psi(z, \pi) \equiv (z, -\pi). \tag{98}$$

Since the replacement of $\Phi_{\Delta s}$ by $\mathcal{T}$ or Ψ preserves the phase-space volume element and the reversibility, the detailed balance condition (38) still holds.[19]

[19] In practice, we check the reversibility at every step of molecular dynamics, $(z, \pi) \to (z', \pi') = \Phi_{\Delta s}(z, \pi)$, by monitoring that the time-reversed process $(\tilde{z}, -\tilde{\pi}) \equiv \Phi_{\Delta s}(z', -\pi')$ $(= \Phi_{\Delta s} \circ \Psi \circ \Phi_{\Delta s}(z, \pi))$ correctly gives $(z, -\pi)$. In the calculation below, we require that $|\tilde{z} - z|/\sqrt{N} < 10^{-5}$. If this condition is not met, we replace $\Phi_{\Delta s}$ by $\mathcal{T}$ or Ψ as in the case where a transition passes over a boundary.

3.6. Estimation of observables

We first recall that the boundary flow times (T_0 and T_1) can be chosen arbitrarily due to Cauchy's theorem. In practice, T_0 must be set sufficiently small in order to keep in $\mathcal{R}$ a region that is free from multimodality (to be set to $T_0 = 0$ when the multimodal problem is absent there). On the other hand, T_1 must be taken sufficiently large in order to keep a region where the sign problem is resolved, but, at the same time, T_1 should not be set too large in order to avoid introducing an unnecessarily large computational time.

When estimating observables, we take a subinterval in $[T_0, T_1]$ (to be denoted by $[\hat{T}_0, \hat{T}_1]$). Namely, for a sample of configurations that are generated in the range $[T_0, T_1]$ (with sample size $N_{\rm conf}$), we construct a subsample $\{z^{(k)}\}$ ($k = 1, \dots, \hat{N}_{\rm conf}$) taking configurations from the interval $[\hat{T}_0, \hat{T}_1]$ ($T_0 \le \hat{T}_0 < \hat{T}_1 \le T_1$), and take a ratio of sample averages over this subsample:[20]

$$\bar{\mathcal{O}}(\hat{T}_0, \hat{T}_1) \equiv \frac{(1/\hat{N}_{\rm conf}) \sum_k A(z^{(k)})\, \mathcal{O}(z^{(k)})}{(1/\hat{N}_{\rm conf}) \sum_k A(z^{(k)})}. \tag{99}$$

$\hat{T}_0$ should now be set sufficiently large in order to exclude a region that is contaminated by the sign problem. Note that $\hat{T}_0$ and $\hat{T}_1$ should be set far enough apart in order to maintain a sufficient size for the subsample. Then, if the original range $[T_0, T_1]$ is properly chosen as above, and if the system is close enough to global equilibrium, there must be a region in two-parameter space $(\hat{T}_0, \hat{T}_1)$ such that the estimations $\bar{\mathcal{O}}(\hat{T}_0, \hat{T}_1)$ are stable against the variation of the estimation ranges (i.e., the estimates change only within statistical errors).

The whole process of the WV-TLTM thus proceeds as follows:

(1) Choose a sufficiently small T_0 and a sufficiently large T_1 to tame both sign and multimodal problems.
(2) Construct a weight function $e^{-W(t)}$ such that the distribution of t becomes almost uniform (see Sect. 3.4 for more details).
(3) Use the HMC algorithm in Sect. 3.5 to generate configurations in the range $[T_0, T_1]$ from the distribution $\propto e^{-V(z)}$ with $V(z) = \mathrm{Re}\, S(z) + W(t(z))$.
(4) For the obtained full sample, vary the estimation range $[\hat{T}_0, \hat{T}_1]$, looking for a stable region (plateau) in the two-parameter space $(\hat{T}_0, \hat{T}_1)$ that gives the same estimate $\bar{\mathcal{O}}(\hat{T}_0, \hat{T}_1)$ within statistical errors.
(5) Choose a point $(\hat{T}_0, \hat{T}_1)$ from the plateau and take the corresponding $\bar{\mathcal{O}}(\hat{T}_0, \hat{T}_1)$ as the estimate of $\langle \mathcal{O}(x) \rangle$. The error of estimation is read from the statistical error for the chosen subsample.

4. Application to a chiral random matrix model

In this section, to confirm that the WV-TLTM works correctly, we apply the WV-TLTM to a chiral random matrix model, the Stephanov model [31,32]. We show that the algorithm correctly reproduces the exact results, solving both the sign and multimodal problems.

4.1. The model

The Stephanov model is a large-N matrix model that approximates QCD at finite density. For N_f quarks with equal mass m, the partition function is given by the following integral over $n \times n$ complex

[20] $\hat{N}_{\rm conf} = N_{\rm conf}(\hat{T}_0, \hat{T}_1)$ is the number of configurations in $[\hat{T}_0, \hat{T}_1]$. The total number of configurations corresponds to $N_{\rm conf} = N_{\rm conf}(T_0, T_1)$.

matrices $X = (X_{ij} = x_{ij} + i\,y_{ij})$:

$$Z_n^{N_f} = e^{n\mu^2} \int d^2X\, e^{-S(X,X^\dagger)} \equiv e^{n\mu^2} \int d^2X\, e^{-n\,\mathrm{tr}\,X^\dagger X}\, \det{}^{N_f}(D+m). \tag{100}$$

Here, $D+m$ represents the Dirac operator in the chiral representation and takes the form

$$D + m \equiv \begin{pmatrix} m\,1_n & i\,(X+C) \\ i\,(X^\dagger + C) & m\,1_n \end{pmatrix}, \tag{101}$$

where

$$i\,C \equiv \begin{pmatrix} (\mu + i\tau)\,1_{n/2} & 0 \\ 0 & (\mu - i\tau)\,1_{n/2} \end{pmatrix}. \tag{102}$$

The parameters μ and τ correspond to the chemical potential and the temperature, respectively [31,32]. The number of DOF is $N = 2n^2$, which may be compared with the DOF of the $SU(N_c)$ gauge field on the lattice of linear size L as $N = 4(N_c^2 - 1)L^4$.

For the case $N_f = 1$, the partition function at finite n can be written as an integral over a single variable:

$$Z_n^{N_f=1} = n\, e^{n(\mu^2 - m^2)} \int_0^\infty d\rho\, e^{-n\rho}\, I_0(2nm\sqrt{\rho}) \left[(\rho - \mu^2 + \tau^2)^2 + (2\mu\tau)^2\right]^{n/2}, \tag{103}$$

where $I_k(x)$ $(k = 0, 1, 2, \ldots)$ are modified Bessel functions of the first kind. Then, the chiral condensate is expressed as

$$\begin{aligned} \langle \bar\psi \psi \rangle &\equiv \frac{1}{2n} \frac{\partial}{\partial m} \ln Z_n^{N_f=1} \\ &= -m + \frac{\int_0^\infty d\rho\, e^{-n\rho}\, I_1(2nm\sqrt{\rho})\, \sqrt{\rho} \left[(\rho - \mu^2 + \tau^2)^2 + (2\mu\tau)^2\right]^{n/2}}{\int_0^\infty d\rho\, e^{-n\rho}\, I_0(2nm\sqrt{\rho}) \left[(\rho - \mu^2 + \tau^2)^2 + (2\mu\tau)^2\right]^{n/2}}. \end{aligned} \tag{104}$$

Similarly, the number density is expressed as

$$\begin{aligned} \langle \psi^\dagger \psi \rangle &\equiv \frac{1}{2n} \frac{\partial}{\partial \mu} \ln Z_n^{N_f=1} \\ &= \mu - \mu \frac{\int_0^\infty d\rho\, e^{-n\rho}\, I_0(2nm\sqrt{\rho}) \left[(\rho - \mu^2 + \tau^2)^2 + (2\mu\tau)^2\right]^{n/2-1} (\rho - \mu^2 - \tau^2)}{\int_0^\infty d\rho\, e^{-n\rho}\, I_0(2nm\sqrt{\rho}) \left[(\rho - \mu^2 + \tau^2)^2 + (2\mu\tau)^2\right]^{n/2}}. \end{aligned} \tag{105}$$

We apply the WV-TLTM to this model, by complexifying the real and imaginary parts (x_{ij} and y_{ij}) separately, and by considering the antiholomorphic gradient flow with respect to the action given in Eq. (100). We estimate the chiral condensate and the number density using the formulas

$$\langle \bar\psi \psi \rangle = \frac{1}{2n} \langle \mathrm{tr}\,(D+m)^{-1} \rangle, \tag{106}$$

$$\langle \psi^\dagger \psi \rangle = \mu + \frac{1}{2n} \left\langle \mathrm{tr} \left[(D+m)^{-1} \begin{pmatrix} 0 & 1_n \\ 1_n & 0 \end{pmatrix} \right] \right\rangle. \tag{107}$$

It is convenient to introduce the matrices

$$A \equiv X + C, \quad B \equiv X^\dagger + C, \quad K \equiv (BA + m^2)^{-1}, \tag{108}$$

with which $D+m$ and $(D+m)^{-1}$ are expressed as[21]

$$D+m=\begin{pmatrix} m\,1_n & i\,A \\ i\,B & m\,1_n \end{pmatrix}, \tag{109}$$

$$\begin{aligned}(D+m)^{-1} &= \begin{pmatrix} m\,(AB+m^2)^{-1} & -i\,A(BA+m^2)^{-1} \\ -i\,(BA+m^2)^{-1}B & m\,(BA+m^2)^{-1} \end{pmatrix} \\ &= \begin{pmatrix} m\,AKA^{-1} & -i\,AK \\ -i\,KB & m\,K \end{pmatrix}. \end{aligned} \tag{110}$$

The flow equation is then written only with A, B, K, and the expectation values (106) and (107) are estimated from the expressions

$$\langle\bar\psi\psi\rangle = \frac{m}{n}\,\langle\mathrm{tr}\,K\rangle, \tag{111}$$

$$\langle\psi^\dagger\psi\rangle = \mu - \frac{i}{2n}\,\langle\mathrm{tr}\,K(A+B)\rangle. \tag{112}$$

4.2. *Setup in the simulation*

We summarize the setup in the simulation. We set $n=10$, $m=0.004$, $\tau=0$, and estimate the chiral condensate $\langle\bar\psi\psi\rangle$ and the number density $\langle\psi^\dagger\psi\rangle$ as functions of $\mu=0.4,\dots,0.8$.[22]

We set $T_0=0$ while we choose T_1 depending on μ as in Table 1. Other simulation parameters are also given in Table 1.

There, $N_{\rm init}$ is the number of initial configurations, and $N_{\rm conf}$ is the number of configurations in the simulation range $[T_0,T_1]$, while $N_{\rm conf}(\hat T_0,\hat T_1)$ is that in the estimation range $[\hat T_0,\hat T_1]$, corresponding to the point $(\hat T_0,\hat T_1)$ chosen from a plateau. Note that $\hat T_0$ and T_1 depend on the choice of observables. We set the initial configuration to the final configuration in the test run determining $W(t)$. The tuning of $W(t)$ turns out to take two iterations to realize the condition (94). In Fig. 8, we show the final form of $W(t)$ at $\mu=0.575$ and the resulting histogram of t.

It sometimes happens that $\mathcal{R}=\bigcup_t\Sigma_t$ is not well explored for large t because of the complicated geometrical structure there. To facilitate transitions, at every start of the HMC algorithm we change the step size Δs and the step number n by randomly taking them from a set $\mathcal{C}=\{(\Delta_c,n_c\}$ $(c=0,\dots,c_{\max})$.[23] In the calculation below, we set $c_{\max}=3$ and choose $\mathcal{C}$ as in Table 2.

We comment that, if we use the original TLTM based on parallel tempering, we need about 70 replicas for $n=10$. We list in Table 3 the numbers of replicas at $\mu=0.6$ for various n. Here, we first determine the maximum flow time T so that the sign problem is well resolved there, and then determine the number of replicas so that the acceptance rate of the swapping is in the range 0.2–0.5.

4.3. *Results and analysis*

Figure 9 shows the average reweighting factors from the naive reweighting method (blue) and from the WV-TLTM (orange), the former exhibiting the existence of the sign problem around $\mu=0.6$. Figure 10 gives the estimates of $\langle\bar\psi\psi\rangle$ from the WV-TLTM at $\mu=0.575$ with various estimation

[21] Note that $(AB+m^2)^{-1}=AKA^{-1}=B^{-1}KB=(1/m^2)\,(1-AKB)$.

[22] Since the lattice size is small, we adopt the direct method in the HMC algorithm; we compute J by integrating the flow equation (12) and use the LU decomposition in the inversion processes. The computation of J is not necessary if we adopt the iterative method (see Sect. 3.3).

[23] This prescription is justified by noticing that this gives a Markov chain on an extended space $\mathcal{R}\times\mathcal{C}$.

Table 1. Simulation parameters.

μ	0.4	0.45	0.5	0.55	0.575	0.6
T_0	0	0	0	0	0	0
T_1	0.025	0.048	0.056	0.068	0.068	0.068
N_{init}	80	60	50	50	50	40
N_{conf}	4000	4000	9000	14 000	17 000	12 000
$\hat{T}_0$ ($\langle\bar{\psi}\psi\rangle$)	0.01	0.027 84	0.004 48	0.0272	0.016 32	0.024 48
$\hat{T}_1$ ($\langle\bar{\psi}\psi\rangle$)	0.025	0.048	0.056	0.068	0.068	0.065 28
$N_{\text{conf}}(\hat{T}_0, \hat{T}_1)$ ($\langle\bar{\psi}\psi\rangle$)	2250	1600	8500	9000	13 500	7400
$\hat{T}_0$ ($\langle\psi^\dagger\psi\rangle$)	0.0105	0.021 12	0.0336	0.0272	0.029 92	0.016 32
$\hat{T}_1$ ($\langle\psi^\dagger\psi\rangle$)	0.025	0.048	0.056	0.068	0.068	0.065 28
$N_{\text{conf}}(\hat{T}_0, \hat{T}_1)$ ($\langle\psi^\dagger\psi\rangle$)	2300	2400	3600	8800	10 000	8800
μ	0.625	0.65	0.7	0.75	0.8	
T_0	0	0	0	0	0	
T_1	0.068	0.064	0.06	0.052	0.04	
N_{init}	50	75	50	40	40	
N_{conf}	9000	8000	4000	4000	4000	
$\hat{T}_0$ ($\langle\bar{\psi}\psi\rangle$)	0.035 36	0.010 24	0.0204	0.009 36	0.0056	
$\hat{T}_1$ ($\langle\bar{\psi}\psi\rangle$)	0.068	0.064	0.06	0.052	0.04	
$N_{\text{conf}}(\hat{T}_0, \hat{T}_1)$ ($\langle\bar{\psi}\psi\rangle$)	4800	6800	2600	3250	3400	
$\hat{T}_0$ ($\langle\psi^\dagger\psi\rangle$)	0.019 04	0.014 08	0.0048	0.004 16	0.0176	
$\hat{T}_1$ ($\langle\psi^\dagger\psi\rangle$)	0.068	0.064	0.06	0.052	0.04	
$N_{\text{conf}}(\hat{T}_0, \hat{T}_1)$ ($\langle\psi^\dagger\psi\rangle$)	6800	6800	3750	3500	2100	

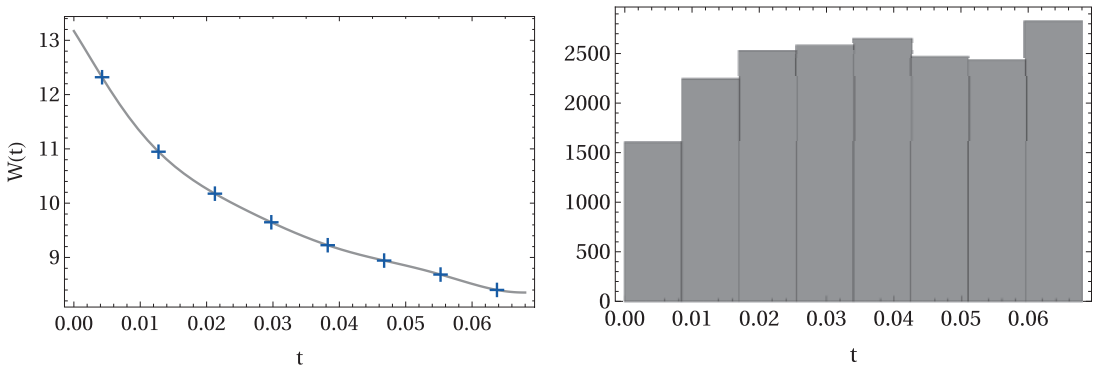

Fig. 8. (Left) Final $\{W_t\}$ and its polynomial fit $W(t)$. (Right) Final histogram of t. Both figures are at $\mu = 0.575$.

Table 2. HMC parameters.

index c	0	1	2	3
Δs_c	0.01	0.005	0.001	0.000 25
n_c	25	50	50	100

ranges $[\hat{T}_0, \hat{T}_1]$. We see a plateau with a value 0.043(15) close to the exact one 0.041. Figure 11 exhibits the estimated values thus obtained for the chiral condensate $\langle\bar{\psi}\psi\rangle$ and the number density $\langle\psi^\dagger\psi\rangle$. As a comparison, we also display in the same figure the results from the naive reweighting method and the complex Langevin method, both with a sample size of $N_{\text{conf}} = 10\,000$. We see that

Table 3. Maximum flow time T and the number of replicas at $\mu = 0.6$.

n	4	6	8	10
T	0.	0.02	0.06	0.068
#replicas	1	4	~ 33	~ 70

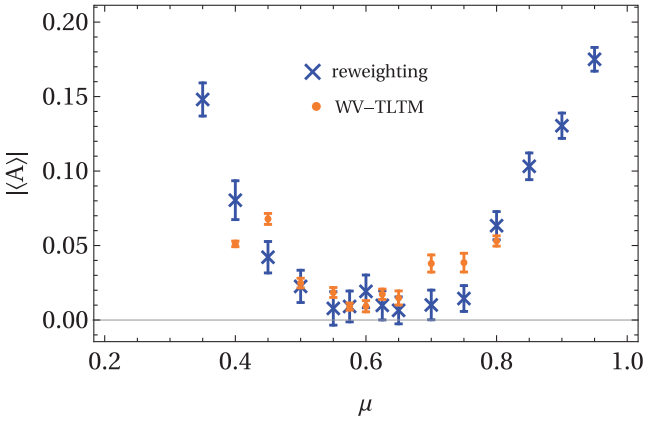

Fig. 9. Average phase factors $\langle A \rangle = \langle e^{-i \operatorname{Im} S(x)} \rangle_{\text{rewt}}$ from the naive reweighting method (blue) and the average reweighting factors $\langle A \rangle = \langle A(z) \rangle_{\mathcal{R}}$ from the WV-TLTM (orange). The estimation range $[\hat{T}_0, \hat{T}_1]$ in the WV-TLTM is set to that for the chiral condensate (see Table 1).

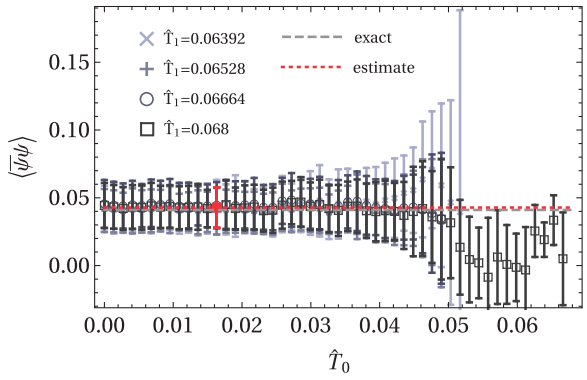

Fig. 10. Estimates of $\langle \bar{\psi}\psi \rangle$ at $\mu = 0.575$ with various $\hat{T}_0$ and $\hat{T}_1$. The red filled box is the point taken from a plateau, giving the estimate 0.043(15) (shown by the dotted line). The exact value ($= 0.041$) is shown by the dashed line.

the WV-TLTM correctly reproduces exact values, while the complex Langevin method suffers from the wrong convergence even for a parameter region free from the sign problem.

One may find it strange that correct estimates are still obtained from the WV-TLTM even for such parameters that give small average reweighting factors $\langle A(z) \rangle_{\mathcal{R}}$ (see Fig. 9). To understand this, let us see Figs. 12 and 13, which show the histogram of $A(z) = \alpha^{-1}(z)\, e^{-i \operatorname{Im} S(z) + i\varphi(z)}$, and those of its modulus and phase, in the estimation range $[\hat{T}_0, \hat{T}_1]$ at $\mu = 0.575$.

We observe that the distribution of $\alpha^{-1}(z)$ has a peak around $\alpha^{-1}(z) = 0.12$. The point is that this small value reduces not only the mean value of A, but also the statistical errors. This is in sharp contrast to the situation in the naive reweighting (see the right panel of Fig. 12). In fact, in the latter (the naive reweighting), the reweighting factor is actually a phase factor, and is distributed uniformly

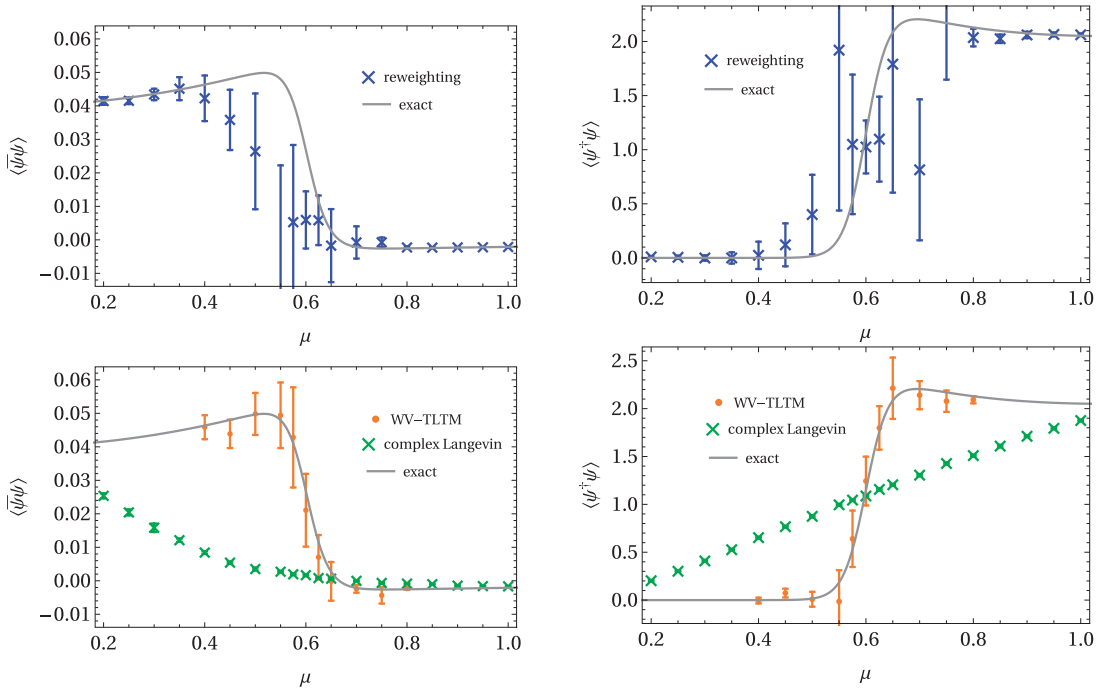

Fig. 11. (Left) The chiral condensate. (Right) The number density. The top panels are the results from the reweighting method, and the bottom panels are from the WV-TLTM and the complex Langevin method.

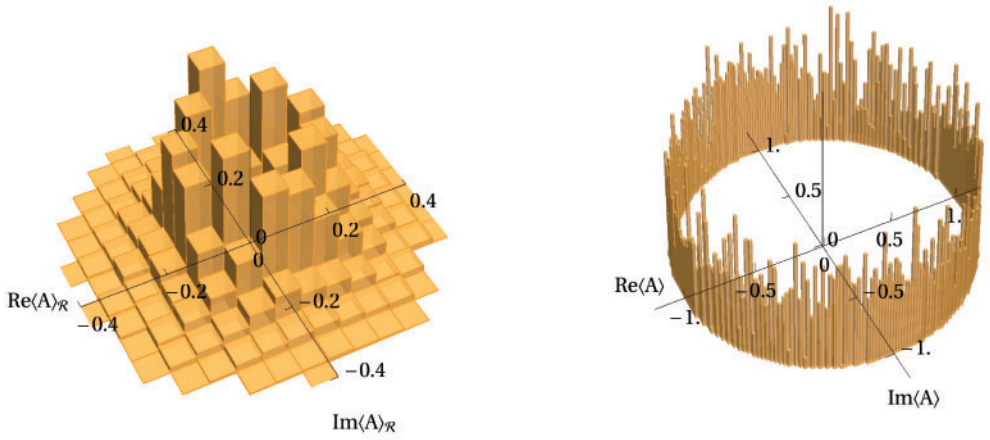

Fig. 12. (Left) Histogram of $A(z) = \alpha^{-1}(z)\, e^{i\varphi(z) - i\,\mathrm{Im}\,S(z)}$ obtained from the WV-TLTM. (Right) Histogram of the phase factor $e^{i\varphi(z) - i\,\mathrm{Im}\,S(z)}$ obtained from the reweighting. Both figures are at $\mu = 0.575$.

on a unit circle, giving a vanishingly small average phase factor. The statistical errors are then of $O(1/\sqrt{N_{\mathrm{conf}}})$, because the standard deviation of phase factors for uniformly distributed phases is of $O(1)$. One thus needs a huge sample size of $e^{O(N)}$ to make the statistical errors relatively small compared to the mean value. In the WV-TLTM, in contrast, the reweighting factor $A(z)$ is distributed in a 2D way (not on a circle), and the contributions of the radius α^{-1} enter both the mean value and the statistical errors, and also both the numerator and the denominator. Thus, the effect of small radius cancels out in a ratio of reweighted averages. Therefore, no additional problem is caused by

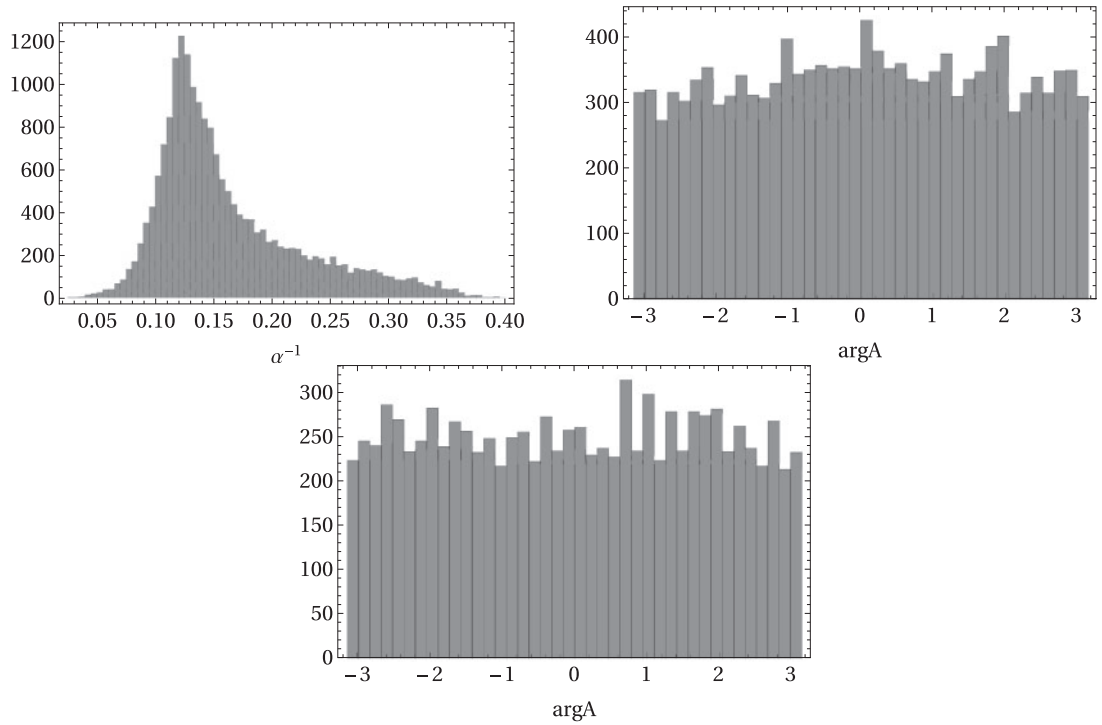

Fig. 13. (Top) Histograms of the modulus $|A(z)| = \alpha^{-1}(z)$ and the phase $\arg A(z) = \varphi(z) - \mathrm{Im}\, S(z)$ from the WV-TLTM. (Bottom) Histogram of the phase from the reweighting method. All figures are at $\mu = 0.575$.

the smallness of the reweighting factor, and the extent of the sign problem is still governed by the phase factor, which is reduced by taking sufficiently large flow times.

5. Conclusion and outlook

In this paper, we have proposed an HMC algorithm on the worldvolume $\mathcal{R}$ of an integration surface Σ_t, where the flow time t changes in the course of molecular dynamics, and thus the multimodal problem is resolved without introducing replicas. Furthermore, computation of the Jacobian is not necessary in generating a configuration. We applied this algorithm to a chiral random matrix model (the Stephanov model) and confirmed that it reproduces the correct results, solving both the sign and multimodal problems simultaneously.

The validity of this algorithm should be further investigated by applying it to other systems that also have the sign problem, including finite density QCD, strongly correlated electron systems, and real-time quantum field theories as well as frustrated spin systems like the antiferromagnetic Heisenberg model on the triangular lattice and the Kitaev model on the honeycomb lattice.

It is also important to keep developing the algorithm in order to perform large-scale calculations for such systems as those listed above with lower computational cost. For example, it should be nice to find a more efficient algorithm to tune $W(t)$.[24] At the same time, it is worth developing an algorithm where the weight $e^{-W(t)}$ need not be introduced, as happens when one switches from simulated tempering [40] to parallel tempering [25–27]. It would also be desirable to construct an algorithm to evaluate the phase $e^{i\varphi} = \det J/|\det J|$ without computing the matrix elements of J explicitly. Furthermore, in order to make a more accurate statistical analysis, it is important to develop

[24] Machine learning may be one of the possible tools.

a systematic method to estimate numerical errors that are necessarily introduced in integrating the antiholomorphic gradient flow and in solving Newton's method iteratively (Step 2 in Sect. 3.5).

The modification of the flow equation (6) also deserves intensive investigation for various reasons. To see this, note that Eq. (6) is not the only possible equation deforming the original integration surface $\mathbb{R}^N$ so as to approach a union of Lefschetz thimbles. For example, it can be modified with a positive Hermitian matrix $G^{ij}(z,z^*)$ to the form

$$\frac{dz_t^i}{dt} = G^{ij}(z,z^*)\,[\partial_j S(z_t)]^*, \tag{113}$$

without changing the structure of the thimbles. However, this modification changes the flows of the configurations off the thimbles, and can be designed so that the flowed configurations approach zeros of $e^{-S(z)}$ only very slowly (see, e.g., Ref. [41]). We have investigated this type of modification, proposing to take G^{ij} of the following simple form [42]:

$$G^{ij}(z,z^*) = \frac{\delta^{ij}}{1+(|\partial S(z)|/\Lambda)^\alpha} \quad (\alpha \geq 2). \tag{114}$$

This actually removes zeros from $\mathcal{R}$ for finite flow times, and is sometimes helpful in iteratively solving the constraint (66). However, it seems that the obtained gain does not exceed the increased complexity of the algorithm, and also that the functional form of G^{ij} needs to be fine-tuned, depending on the parameters of each model. This is the reason why we did not pursue this possibility in this paper. However, it will be essentially important when one develops a Metropolis–Hastings algorithm, described in Appendix D, because the configuration space $\tilde{\mathcal{R}} = \{\xi = (t,x^a) \mid T_0 \leq t \leq T_1\}$ comes to have a simple structure if points to be mapped to zeros do not exist in the region.

Another possible application of modifying the flow equation is to provide a mechanism to solve the so-called global sign problem (cancellation among contributions from different thimbles). In fact, since $\alpha^{-1}(z)$ increases exponentially in the vicinity of a Lefschetz thimble [see the comment below Eq. (B.8)], a change of flows caused by the modification may significantly shift the distribution of $A(z)$ and distort the balanced contributions from different thimbles, which was the origin of the global sign problem.

A study along these lines is now in progress and will be reported elsewhere.

Acknowledgements

The authors thank Hitotsugu Fujii, Issaku Kanamori, Yoshio Kikukawa, Yusuke Namekawa, and especially Naoya Umeda for useful discussions. This work was partially supported by Japan Society for the Promotion of Science (JSPS) KAKENHI (Grant Numbers 20H01900, 18J22698) and by SPIRITS 2020 of Kyoto University (PI: M.F.).

Funding

Open Access funding: SCOAP[3].

Appendix A. Geometry of the RATTLE algorithm

In this appendix, we clarify the geometrical aspects of the RATTLE algorithm and prove a few statements necessary for discussions in the main text.

As in the main text, let $\mathcal{R}$ be an m-dimensional manifold embedded in the flat space $\mathbb{R}^M = \{z = (z^I)\}$ $(I = 1,\ldots,M)$.[25] With local coordinates $\xi = (\xi^\mu)$ $(\mu = 1,\ldots,m)$ of $\mathcal{R}$, the embedding is

[25] When $\mathcal{R}$ is the worldvolume of an integration surface in $\mathbb{C}^N = \mathbb{R}^{2N}$, we set $M = 2N$ and $m = N+1$.

expressed by functions $z^I = z^I(\xi)$. The vectors

$$E_\mu = \left(E^I_\mu = \partial z^I/\partial \xi^\mu\right) \tag{A.1}$$

form a basis of the tangent space $T_z\mathcal{R}$ at $z \in \mathcal{R}$, and give the induced metric $ds^2 = g_{\mu\nu}(\xi)\, d\xi^\mu d\xi^\nu$ as

$$g_{\mu\nu} = E_\mu \cdot E_\nu. \tag{A.2}$$

Furthermore, denoting the inverse of $(g_{\mu\nu})$ by $(g^{\mu\nu})$, and defining another basis of $T_z\mathcal{R}$ by[26]

$$E^\mu \equiv g^{\mu\nu} E_\mu, \tag{A.3}$$

we also introduce local coordinates $\eta = (\eta_\mu)$ on $T_z\mathcal{R}$ as coefficients in an expansion with respect to E^μ.[27] Namely, for $\pi \in T_z\mathcal{R}$, its coordinates $\eta = (\eta_\mu)$ are defined through the relation

$$\pi = \eta_\mu\, E^\mu, \tag{A.4}$$

whose explicit forms are given by

$$\eta_\mu = E_\mu \cdot \pi. \tag{A.5}$$

The line element on $T_z\mathcal{R}$ then takes the form

$$(d\pi^I)^2\big|_{T_z\mathcal{R}} = g^{\mu\nu}(\xi)\, d\eta_\mu\, d\eta_\nu, \tag{A.6}$$

and thus the volume elements of $T\mathcal{R}$ are given by

$$dV_{T\mathcal{R}} = Dz\, D\pi = D\xi\, D\eta \tag{A.7}$$

with

$$Dz \equiv D\xi \equiv \sqrt{g}\, d\xi, \quad D\pi \equiv D\eta \equiv d\eta/\sqrt{g}. \tag{A.8}$$

We think of the tangent bundle $T\mathbb{R}^M = \{(z^I, \pi^I)\}$ (*not* the cotangent bundle) as the phase space of motions in $\mathbb{R}^M$ with a symplectic form $\Omega \equiv d\pi^I \wedge dz^I$. The sub-bundle $T\mathcal{R} = \{(z^I, \pi^I) \,|\, z \in \mathcal{R},\ \pi \in T_z\mathcal{R}\}$ is then regarded as the phase space of constrained motions on $\mathcal{R}$. Its symplectic form is given by[28]

$$\omega \equiv \Omega|_{T\mathcal{R}} = d\eta_\mu \wedge d\xi^\mu, \tag{A.9}$$

which defines the Poisson bracket as $\{\xi^\mu, \eta_\nu\} = \delta^\mu_\nu$, $\{\xi^\mu, \xi^\nu\} = \{\eta_\mu, \eta_\nu\} = 0$. The volume element (A.7) agrees with the phase-space volume element associated with ω:

$$dV_{T\mathcal{R}} = \frac{\omega^m}{m!}, \tag{A.10}$$

[26] Note that $E^\mu \cdot E^\nu = g^{\mu\nu}$.

[27] As in the main text, we denote a function on $\mathcal{R}$ by $f(z)$ and $f(\xi)$, interchangeably, with the understanding that $z = z(\xi)$. The transition matrix is also written as $P(z'|z)$ and $P(\xi'|\xi)$ for $z = z(\xi)$, $z' = z(\xi') \in \mathcal{R}$. Similarly, a function on $T\mathcal{R}$ is written as $f(z, \pi)$ and $f(\xi, \eta)$, interchangeably.

[28] This can be proved as follows: $\omega = d\pi^I \wedge dz^I|_{T\mathcal{R}} = \left[d\pi^I|_{T_z\mathcal{R}}\right] \wedge dz^I(\xi) = \left[d\pi^I|_{T_z\mathcal{R}}\right] \wedge E^I_\mu(\xi)\, d\xi^\mu = d\left[\pi^I|_{T_z\mathcal{R}}\, E^I_\mu(\xi)\right] \wedge d\xi^\mu = d\eta_\mu \wedge d\xi^\mu$, where we have used $d\left[E^I_\mu(\xi)\right] \wedge d\xi^\mu = \left(\partial^2 z^I/\partial\xi^\nu \partial\xi^\mu\right) d\xi^\nu \wedge d\xi^\mu = 0$.

because $\omega^m/m! = d\xi\, d\eta = (\sqrt{g}\, d\xi)\, (d\eta/\sqrt{g}) = D\xi\, D\eta$.

We now consider molecular dynamics on $T\mathcal{R}$ with a Hamiltonian

$$H(\xi,\eta) = \frac{1}{2}\,(\pi^I)^2 + V(z)\Big|_{T\mathcal{R}} = \frac{1}{2}\, g^{\mu\nu}(\xi)\,\eta_\mu\eta_\nu + V(z(\xi)), \tag{A.11}$$

with which the time evolution of (ξ,η) is given by

$$\partial_s\xi^\mu = \{\xi^\mu, H\} = \partial_{\eta_\mu}H(\xi,\eta), \quad \partial_s\eta_\mu = \{\eta_\mu, H\} = \partial_{\xi^\mu}H(\xi,\eta). \tag{A.12}$$

It is easy to see that this evolution preserves the Hamiltonian H, the symplectic form ω, and thus also the phase-space volume element $dV_{T\mathcal{R}}$:

$$\partial_s H = 0, \quad \partial_s\omega = 0, \quad \partial_s(dV_{T\mathcal{R}}) = 0. \tag{A.13}$$

We write the motion from (ξ,η) to (ξ',η') with time interval s as a map Φ_s:

$$(\xi,\eta) \to (\xi',\eta') = \Phi_s(\xi,\eta) \equiv (\xi_s(\xi,\eta), \eta_s(\xi,\eta)). \tag{A.14}$$

Since the Hamiltonian satisfies the relation $H(\xi,-\eta) = H(\xi,\eta)$, the map Φ_s preserves the reversibility. Namely, if $(\xi',\eta') = \Phi_s(\xi,\eta)$ is a motion, we have $(\xi,-\eta) = \Phi_s(\xi',-\eta')$. Furthermore, due to the volume preservation, the kernel

$$\Phi_s(\xi',\eta'|\xi,\eta) \equiv \delta\big(\xi' - \xi_s(\xi,\eta)\big)\,\delta\big(\eta' - \eta_s(\xi,\eta)\big) \tag{A.15}$$

satisfies the relation

$$\Phi_s(\xi',\eta'|\xi,\eta) = \Phi_s(\xi,-\eta|\xi',-\eta'). \tag{A.16}$$

For stochastic processes on $\mathcal{R}$, we define the probability density $p(z(\xi)) = p(\xi)$ on $\mathcal{R}$ with respect to the volume element $Dz = D\xi = \sqrt{g}\, d\xi$, which is thus normalized as

$$\int_{\mathcal{R}} D\xi\, p(\xi) = 1. \tag{A.17}$$

One can then easily show that the transition matrix

$$P_s(\xi'|\xi) \equiv \int_{T_{z'}\mathcal{R}} D\eta' \int_{T_z\mathcal{R}} D\eta\, \Phi_s(\xi',\eta'|\xi,\eta)\, \frac{1}{(2\pi)^{m/2}}\, e^{-g^{\mu\nu}(\xi)\,\eta_\mu\eta_\nu/2} \tag{A.18}$$

satisfies the detailed balance condition[29]

$$P_s(\xi'|\xi)\, e^{-V(z(\xi))} = P_s(\xi|\xi')\, e^{-V(z(\xi'))} \tag{A.19}$$

and the normalization condition

$$\int_{\mathcal{R}} D\xi'\, P_s(\xi'|\xi) \;= 1. \tag{A.20}$$

The Gaussian distribution $e^{-g^{\mu\nu}(\xi)\,\eta_\mu\eta_\nu/2}/(2\pi)^{m/2}$ in Eq. (A.18) can be obtained by first generating $\tilde\pi = (\tilde\pi^I) \in T_z\mathbb{R}^M$ from the Gaussian distribution $e^{-\tilde\pi^2/2}/(2\pi)^{M/2}$ and then projecting it onto $T_z\mathcal{R}$. In fact, for the orthogonal decomposition $\tilde\pi = \pi + \pi_\perp$, $\tilde\pi^2$ is written as $\tilde\pi^2 = \pi^2 + \pi_\perp^2$, and the

[29] This will be proved for a more complicated case in Eq. (A.39).

integration measure of $T_z\mathbb{R}^M$ is factorized as $d^M\tilde{\pi} \equiv d\tilde{\pi}^1 \cdots d\tilde{\pi}^M = D\pi\, D\pi_\perp$. By integrating out only the normal component, the distribution of $\pi \in T_z\mathcal{R}$ is left in the desired form:

$$\int_{N_z\mathcal{R}} d^M\tilde{\pi}\, \frac{1}{(2\pi)^{M/2}} e^{-\tilde{\pi}^2/2} = D\pi \int_{N_z\mathcal{R}} D\pi_\perp\, \frac{1}{(2\pi)^{M/2}}\, e^{-(\pi^2+\pi_\perp^2)/2} = \frac{1}{(2\pi)^{m/2}}\, e^{-\pi^2/2}\, D\pi\,. \tag{A.21}$$

Note that the projector $\Pi_{\mathcal{R}}$ from $\tilde{\pi} \in T_z\mathbb{R}^M$ to $\pi = \Pi_{\mathcal{R}}\,\tilde{\pi} \in T_z\mathcal{R}$ is given by

$$\Pi_{\mathcal{R}} = E_\mu (E^\mu)^T = g^{\mu\nu} E_\mu E_\nu^T. \tag{A.22}$$

We now assume that $\mathcal{R}$ is characterized by $M - m$ independent constraint equations, $\phi^r(z) = 0$ $(r = 1, \ldots, M-m)$. Then, Hamilton's equations (A.12) can be expressed as constrained motions in $\mathbb{R}^M$ by using Lagrange multipliers λ_r:

$$\partial_s z = \pi, \tag{A.23}$$

$$\partial_s \pi = -\partial V(z) - \lambda_r\, \partial\phi^r(z), \tag{A.24}$$

$$\phi^r(z) = 0, \tag{A.25}$$

$$\pi \cdot \partial\phi^r(z) = 0. \tag{A.26}$$

Here, $\partial \equiv (\partial_{z^I})$ is the gradient in $\mathbb{R}^M$.

The RATTLE algorithm [35,36] is an algorithm that discretizes Eqs. (A.23)–(A.26) preserving the symplecticity and the reversibility (below Δs is the step size):

$$\pi_{1/2} = \pi - \frac{\Delta s}{2}\, \partial V(z) - \lambda_r\, \partial\phi^r(z), \tag{A.27}$$

$$z' = z + \Delta s\, \pi_{1/2}, \tag{A.28}$$

$$\pi' = \pi - \frac{\Delta s}{2}\, \partial V(z') - \lambda'_r\, \partial\phi^r(z'). \tag{A.29}$$

Here, λ_r and λ'_r are determined, respectively, so that the following constraints are satisfied:

$$z' \in \mathcal{R} \quad (\text{i.e., } \phi^r(z') = 0), \tag{A.30}$$

$$\pi' \in T_{z'}\mathcal{R}. \tag{A.31}$$

One can easily show that the map $\Phi_{\Delta s} : (z, \pi) \to (z', \pi')$ actually satisfies the symplecticity and the reversibility (with λ_r and λ'_r interchanged):

- $\omega(z', \pi') = \omega(z, \pi)$, (A.32)
- $(z', \pi') = \Phi_{\Delta s}(z, \pi) \Rightarrow (z, -\pi) = \Phi_{\Delta s}(z', -\pi')$, (A.33)

which means that

$$\Phi_{\Delta s}(z', \pi' | z, \pi) = \Phi_{\Delta s}(z, -\pi | z', -\pi'). \tag{A.34}$$

The Hamiltonian is conserved to the order of Δs^2, i.e., $H(z', \pi') - H(z, \pi) = O(\Delta s^3)$.

The HMC algorithm then consists of the following three steps for a given initial configuration $z \in \mathcal{R}$:

Step 1.Generate a vector $\tilde{\pi} = (\tilde{\pi}^I) \in T_z\mathbb{R}^M$ from the Gaussian distribution

$$\frac{1}{(2\pi)^{M/2}} e^{-\tilde{\pi}^2/2} \tag{A.35}$$

and project it onto $T_z\mathcal{R}$ to obtain an initial momentum $\pi = (\pi^I) \in T_z\mathcal{R}$.

Step 2.Calculate $\Phi_{\Delta s}(z,\pi)$ from Eqs. (A.27)–(A.31). We repeat this step n times to obtain $(z',\pi') = \Phi^n_{\Delta s}(z,\pi)$

Step 3.Update the configuration z to z' with a probability

$$\min\left(1, e^{-H(z',\pi')+H(z,\pi)}\right). \tag{A.36}$$

The above process defines a stochastic process on $\mathcal{R}$ with the following transition matrix for $z' \neq z$:

$$P(z'|z) \equiv \int_{T_{z'}\mathcal{R}} D\pi' \int_{T_z\mathcal{R}} D\pi \, \min\left(1, e^{-H(z',\pi')+H(z,\pi)}\right) \Phi^n_{\Delta s}(z',\pi'|z,\pi) \, \frac{e^{-\pi^2/2}}{(2\pi)^{m/2}}. \tag{A.37}$$

The diagonal ($z' = z$) components are determined from the probability conservation. $P(z'|z)$ can be shown to satisfy the detailed balance condition

$$P(z'|z)\, e^{-V(z)} = P(z|z')\, e^{-V(z')} \quad (z,\, z' \in \mathcal{R}) \tag{A.38}$$

as follows:

$$\begin{aligned} P(z'|z)\, e^{-V(z)} &= \int_{T_{z'}\mathcal{R}} D\pi' \int_{T_z\mathcal{R}} D\pi \, \min\left(1, e^{-H(z',\pi')+H(z,\pi)}\right) \Phi^n_{\Delta s}(z',\pi'|z,\pi) \, \frac{e^{-H(z,\pi)}}{(2\pi)^{m/2}} \\ &= \frac{1}{(2\pi)^{m/2}} \int_{T_{z'}\mathcal{R}} D\pi' \int_{T_z\mathcal{R}} D\pi \, \min\left(e^{-H(z,\pi)}, e^{-H(z',\pi')}\right) \Phi^n_{\Delta s}(z,-\pi|z',-\pi') \\ &= P(z|z')\, e^{-V(z')}, \end{aligned} \tag{A.39}$$

where we have used Eq. (A.34) n times to get the second line, and have made the change of integration variables, $\pi \to -\pi$ and $\pi' \to -\pi'$, with the relation $H(z,-\pi) = H(z,\pi)$ to obtain the third line.

Appendix B. Analytical expressions for the Gaussian case

We present analytical expressions for some geometrical quantities defined in Sect. 3.1 for the action

$$S(x) = \frac{\beta}{2} \sum_{k=1}^{N} (x^k - i)^2. \tag{B.1}$$

This has a single critical point at $z_\sigma = (z^k_\sigma = i)$, and the corresponding Lefschetz thimble is given by $\mathcal{J}_\sigma = \{z = (z^k) \in \mathbb{C}^k \mid \mathrm{Im}\, z^k = 1 \ (\forall k)\}$. Complex vectors will be used throughout this appendix.

The solution of the antiholomorphic flow equation (6) takes the form

$$z^k(t,x) = x^k e^{\beta t} + i(1 - e^{-\beta t}). \tag{B.2}$$

The Jacobian $J(t,x)$ is thus given by

$$J^k{}_a(t,x) = \frac{\partial z^k(t,x)}{\partial x^a} = e^{\beta t}\, \delta^k_a. \tag{B.3}$$

The tangent vectors E_0 and E_a are

$$E_0^k = \beta\,(z^k - i)^* = \beta\,(x^k e^{\beta t} + i e^{-\beta t}), \quad E_a^k = e^{\beta t}\,\delta_a^k. \tag{B.4}$$

Using Eq. (85), we have

$$(E_0^\perp)^k = -i\beta\,(\mathrm{Im}\, z^k - 1) = i\beta\, e^{-\beta t}. \tag{B.5}$$

The components of $g_{\mu\nu}$ are then given in the ADM parametrization (43) by

$$\gamma_{ab} = (J^\dagger J)_{ab} = e^{2\beta t}\,\delta_{ab}, \tag{B.6}$$

$$\beta^a = \gamma^{ab}\,\mathrm{Re}\,(E_0^\dagger E_b) = \beta\, x^a, \tag{B.7}$$

$$\alpha = |E_0^\perp| = \beta\sqrt{N}\, e^{-\beta t}. \tag{B.8}$$

We see that the inverse lapse is given by $\alpha^{-1} = e^{\beta t}/(\beta\sqrt{N})$ and increases exponentially in flow time t as $z(t,x)$ approaches the Lefschetz thimble.

The ideal weight function $e^{-W(t)}$ giving a uniform distribution of t is given by [see Eq. (87)]

$$e^{+W(t)} = \int_{\mathbb{R}^N} dx\,\alpha(t,x)\,|\det J(t,x)|\, e^{-\mathrm{Re}\,S(z(t,x))} = \beta\sqrt{N}\left(\frac{2\pi}{\beta}\right)^{N/2} e^{-\beta t + (\beta N/2) e^{-2\beta t}}, \tag{B.9}$$

and thus we have

$$W(t) = -\beta t + \frac{\beta N}{2}\, e^{-2\beta t}. \tag{B.10}$$

We have ignored t-independent constants. We see that the weight factor also increases exponentially, $e^{-W(t)} \simeq e^{\beta t}$, at large flow times.

Appendix C. Proof of Eq. (69)

In this appendix, we prove the equality (69). First, we construct $2N$ coordinates $(\zeta^A) \equiv (\xi^\mu, \phi^r)$ in the vicinity of $\mathcal{R}$ in $\mathbb{R}^{2N}$, by regarding (ϕ^r) as coordinates in the extra dimensions, and introduce at each point $z \in \mathcal{R}$ a basis $\{E_A\}$ of the tangent space $T_z\mathbb{R}^{2N}$ as

$$E_A = \left(E_A^I \equiv \partial z^I/\partial\zeta^A\right). \tag{C.1}$$

We further introduce the dual basis $\{\hat{E}^A\}$ to $\{E_A\}$ by

$$\hat{E}^A = \left(\hat{E}^{AI} \equiv \partial\zeta^A/\partial z^I\right), \tag{C.2}$$

which satisfies

$$\hat{E}^A \cdot E_B = \delta_B^A. \tag{C.3}$$

Note that $\hat{E}^0$ and $\hat{E}^r$ equal the gradients $\partial t(z)$ and $\partial\phi^r(z)$, respectively. Then, since the vectors $E_0^\perp$, $\hat{E}^a$, and $\hat{E}^r$ also form a basis of $T_z\mathbb{R}^{2N}$, $\hat{E}^0$ can be expanded in the form

$$\hat{E}^0 = c_\perp E_0^\perp + c_a \hat{E}^a + c_r \hat{E}^r. \tag{C.4}$$

The coefficients can be calculated by using the relations (C.3) and $E_0 \cdot E_0^\perp = (E_0^\perp)^2$ to be

$$c_\perp = \frac{1}{(E_0^\perp)^2}, \quad c_a = 0, \quad c_r = -(E_r \cdot E_0^\perp) \times c_\perp = -\frac{E_r \cdot E_0^\perp}{(E_0^\perp)^2}, \tag{C.5}$$

and thus we find that $\hat{E}^0 = \partial t(z)$ takes the form

$$\partial t(z) = \frac{1}{(E_0^\perp)^2} E_0^\perp - \frac{E_r \cdot E_0^\perp}{(E_0^\perp)^2} \partial \phi^r(z). \tag{C.6}$$

Appendix D. Another version of WV-TLTM with a Metropolis–Hastings algorithm

In this appendix, we give another version of WV-TLTM that also does not require the computation of the Jacobian in generating a configuration. This is a Metropolis–Hastings algorithm on a subspace in the parameter space (not in the target space), $\tilde{\mathcal{R}} \equiv \{\xi = (t, x^a) \mid T_0 \le t \le T_1\}$.

We first rewrite Eq. (13) to the form

$$\langle \mathcal{O}(x) \rangle = \frac{\int d\xi \ \det J(\xi)\, e^{-S(z(\xi))-W(t)}\, \mathcal{O}(z(\xi))}{\int d\xi \ \det J(\xi)\, e^{-S(z(\xi))-W(t)}}, \tag{D.1}$$

where $d\xi \equiv dt\, dx \equiv dt\, dx^1 \cdots dx^N$. Then, by introducing a new positive weight and a new reweighting factor as

$$e^{-\tilde{V}(\xi)} \equiv e^{-\mathrm{Re}\, S(z(\xi))-W(t)}, \tag{D.2}$$

$$\tilde{A}(\xi) \equiv \det J(\xi)\, e^{-i\, \mathrm{Im}\, S(z(\xi))}, \tag{D.3}$$

we can rewrite Eq. (D.1) as a ratio of new reweighted averages,

$$\langle \mathcal{O}(x) \rangle = \frac{\langle \tilde{A}(\xi)\, \mathcal{O}(z(\xi)) \rangle_{\tilde{\mathcal{R}}}}{\langle \tilde{A}(\xi) \rangle_{\tilde{\mathcal{R}}}}, \tag{D.4}$$

where

$$\langle f(z) \rangle_{\tilde{\mathcal{R}}} \equiv \frac{\int_{\tilde{\mathcal{R}}} d\xi\, e^{-\tilde{V}(\xi)} f(\xi)}{\int_{\tilde{\mathcal{R}}} d\xi\, e^{-\tilde{V}(\xi)}}. \tag{D.5}$$

The weight $e^{-W(t)}$ is determined so that the function

$$\tilde{Z}(t; W) \equiv e^{-W(t)} \int_{\mathbb{R}^N} dx\, e^{-\mathrm{Re}\, S(z(t,x))} \tag{D.6}$$

is almost independent of t, as in Sect. 3.4.

The distribution $e^{-\tilde{V}(\xi)}/Z_{\tilde{\mathcal{R}}}$ ($Z_{\tilde{\mathcal{R}}} = \int_{\tilde{\mathcal{R}}} d\xi\, e^{-\tilde{V}(\xi)}$) can be obtained from a Markov chain without evaluating J explicitly, if one uses the Metropolis–Hastings algorithm to update a configuration.[30] Namely, from a configuration ξ we first propose a new configuration ξ' with a probability

$$\mathrm{Prop}(\xi'|\xi) \equiv \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-(t'-t)^2/2\sigma_t^2} \frac{1}{(2\pi\sigma_x^2)^{N/2}} e^{-(x'-x)^2/2\sigma_x^2}, \tag{D.7}$$

where we have treated t and x anisotropically. We then accept ξ' with a probability

$$\min\Big(1, \frac{\mathrm{Prop}(\xi|\xi')\, e^{-\tilde{V}(\xi')}}{\mathrm{Prop}(\xi'|\xi)\, e^{-\tilde{V}(\xi)}}\Big). \tag{D.8}$$

[30] One can also use the HMC algorithm in principle, but this requires the computation of the Jacobian $J = (J^i{}_a(\xi))$ because Hamilton's equation in molecular dynamics involves the gradient $\partial_{x^a} S(z(\xi)) = \partial_{z^i} S(z)\, J^i{}_a(\xi)$.

This algorithm gives a stochastic process with the transition matrix[31]

$$\tilde{P}(\xi'|\xi) \equiv \min\left(1, \frac{\mathrm{Prop}(\xi|\xi')\, e^{-\tilde{V}(\xi')}}{\mathrm{Prop}(\xi'|\xi)\, e^{-\tilde{V}(\xi)}}\right) \mathrm{Prop}(\xi'|\xi) \quad (\xi' \neq \xi), \tag{D.9}$$

and one can easily show that this satisfies the detailed balance condition:

$$\tilde{P}(\xi'|\xi)\, e^{-\tilde{V}(\xi)} = \tilde{P}(\xi|\xi')\, e^{-\tilde{V}(\xi')}. \tag{D.10}$$

In a generic case, $\mathrm{Re}\, S(z(t,x))$ changes rapidly at large flow times t, and thus we should better change the proposal depending on $\xi = (t, x^a)$. One way is to change (σ_t^2, σ_x^2) by randomly taking them from a set $\tilde{\mathcal{C}} = \{(\sigma_{t,c}^2, \sigma_{x,c}^2)\}$ $(c = 1, \ldots, c_{\max})$ as in Sect. 4.2. Another way is to use an asymmetric proposal $\mathrm{Prop}(\xi'|\xi)$ by making σ_t^2 and σ_x^2 t-dependent functions:

$$\mathrm{Prop}(\xi'|\xi) = \frac{1}{\sqrt{2\pi\sigma_t^2(t)}}\, e^{-(t'-t)^2/2\sigma_t^2(t)} \frac{1}{(2\pi\sigma_x^2(t))^{N/2}}\, e^{-(x'-x)^2/2\sigma_x^2(t)}. \tag{D.11}$$

In the latter case, the functional forms of $\sigma_t^2(t)$ and $\sigma_x^2(t)$ are fixed manually or adaptively from test runs.

After a sample is obtained for the region $[T_0, T_1]$, we consider subsamples for various estimation ranges $[\hat{T}_0, \hat{T}_1]$, and estimate an observable by looking at a plateau in the 2D parameter space $\{(\hat{T}_0, \hat{T}_1)\}$, as in Sect. 3.6.

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]].
[3] G. Parisi, Phys. Lett. B **131**, 393 (1983).
[4] J. R. Klauder, Acta Phys. Austriaca Suppl. **25**, 251 (1983).
[5] J. R. Klauder, Phys. Rev. A **29**, 2036 (1984).
[6] G. Aarts, E. Seiler, and I.-O. Stamatescu, Phys. Rev. D **81**, 054508 (2010) [arXiv:0912.3360 [hep-lat]] [Search INSPIRE].
[7] G. Aarts, F. A. James, E. Seiler, and I.-O. Stamatescu, Eur. Phys. J. C **71**, 1756 (2011) [arXiv:1101.3270 [hep-lat]] [Search INSPIRE].
[8] G. Aarts, L. Bongiovanni, E. Seiler, D. Sexty, and I.-O. Stamatescu, Eur. Phys. J. A **49**, 89 (2013) [arXiv:1303.6425 [hep-lat]] [Search INSPIRE].
[9] K. Nagata, J. Nishimura, and S. Shimasaki, Phys. Rev. D **94**, 114515 (2016) [arXiv:1606.07627 [hep-lat]] [Search INSPIRE].
[10] C. E. Berger, L. Rammelmüller, A. C. Loheac, F. Ehmann, J. Braun, and J. E. Drut, Phys. Rept. **892**, 1 (2021) [arXiv:1907.10183 [cond-mat.quant-gas]] [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, [AuroraScience Collaboration] 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] A. Alexandru, G. Başar, and P. Bedaque, Phys. Rev. D **93**, 014504 (2016) [arXiv:1510.03258 [hep-lat]] [Search INSPIRE].
[16] 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].

[31] The diagonal components are determined by the probability conservation, $\int d\xi'\, \tilde{P}(\xi'|\xi) = 1$.

[17] M. Fukuma and N. Umeda, Prog. Theor. Exp. Phys. **2017**, 073B01 (2017) [arXiv:1703.00861 [hep-lat]] [Search INSPIRE].
[18] A. Alexandru, G. Başar, P. F. Bedaque, and N. C. Warrington, Phys. Rev. D **96**, 034513 (2017) [arXiv:1703.02414 [hep-lat]] [Search INSPIRE].
[19] A. Alexandru, G. Başar, P. F. Bedaque, and G. W. Ridgway, Phys. Rev. D **95**, 114501 (2017) [arXiv:1704.06404 [hep-lat]] [Search INSPIRE].
[20] M. Fukuma, N. Matsumoto, and N. Umeda, Phys. Rev. D **100**, 114510 (2019) [arXiv:1906.04243 [cond-mat.str-el]] [Search INSPIRE].
[21] M. Fukuma, N. Matsumoto, and N. Umeda, arXiv:1912.13303 [hep-lat] [Search INSPIRE].
[22] M. Ulybyshev, C. Winterowd, and S. Zafeiropoulos, arXiv:1906.02726 [cond-mat.str-el] [Search INSPIRE].
[23] M. Ulybyshev, C. Winterowd, and S. Zafeiropoulos, Phys. Rev. D **101**, 014508 (2020) [arXiv:1906.07678 [cond-mat.str-el]] [Search INSPIRE].
[24] A. Alexandru, "Improved algorithms for generalized thimble method," talk at the 37th international conference on lattice field theory, Wuhan, 2019.
[25] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **57**, 2607 (1986).
[26] C. J. Geyer, Markov chain Monte Carlo maximum likelihood, in *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface* (American Statistical Association, New York, 1991), p. 156.
[27] K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996).
[28] M. Fukuma, N. Matsumoto, and N. Umeda, J. High Energy Phys. **1811**, 060 (2018) [arXiv:1806.10915 [hep-th]] [Search INSPIRE].
[29] M. Fukuma, N. Matsumoto, and N. Umeda, PoS **LATTICE2019**, 168 (2019) [arXiv:2001.02028 [hep-lat]] [Search INSPIRE].
[30] M. Fukuma, N. Matsumoto, and N. Umeda, J. High Energy Phys. **1712**, 001 (2017) [arXiv:1705.06097 [hep-lat]] [Search INSPIRE].
[31] M. A. Stephanov, Phys. Rev. Lett. **76**, 4472 (1996) [arXiv:hep-lat/9604003] [Search INSPIRE].
[32] 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].
[33] J. Bloch, J. Glesaaen, J. J. M. Verbaarschot, and S. Zafeiropoulos, J. High Energy Phys. **1803**, 015 (2018) [arXiv:1712.07514 [hep-lat]] [Search INSPIRE].
[34] R. Arnowitt, S. Deser, and C. W. Misner, Gen. Relat. Gravit. **40**, 1997 (2008) [arXiv:gr-qc/0405109] [Search INSPIRE].
[35] H. C. Andersen, J. Comput. Phys. **52**, 24 (1983).
[36] B. J. Leimkuhler and R. D. Skeel, J. Comput. Phys. **112**, 117 (1994).
[37] Y. Saad and M. H. Schultz, SIAM J. Sci. Stat. Comput. **7**, 856 (1986).
[38] H. A. van der Vorst, SIAM J. Sci. Stat. Comput. **13**, 631 (1992).
[39] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001) [arXiv:cond-mat/0011174 [cond-mat.stat-mech]] [Search INSPIRE].
[40] E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992) [arXiv:hep-lat/9205018] [Search INSPIRE].
[41] Y. Tanizaki, H. Nishimura, and J. J. M. Verbaarschot, J. High Energy Phys. **1710**, 100 (2017) [arXiv:1706.03822 [hep-lat]] [Search INSPIRE].
[42] M. Fukuma, N. Matsumoto, and N. Umeda, "Applying the tempered Lefschetz thimble method to the sign problem of chiral matrix models and the estimation of the computational scaling," talk at the JPS meeting (online), September 15, 2020.