Complex Langevin, thimbles and tensor networks as solutions to the sign problem

Atis YOSPRAKOB

PhD thesis

Department of Particle and Nuclear Physics,
School of High Energy Accelerator Science,
The Graduate University for Advanced Studies, SOKENDAI

Abstract

In this thesis, I discuss various ways to overcome the numerical sign problem, which has been a recurring obstacle that appears in Monte Carlo studies of many interesting physical systems, such as real-time quantum dynamics, finite-density QCD, and gauge theory with a θ term. To tackle such problems, several techniques had been introduced with different advantages and disadvantages. Firstly, we discuss the application of the complex Langevin method to the two-dimensional gauge theory with a θ term. Although the large drifts and topology freezing problems prevent us from performing the naive calculation, the simulation with a topological defect on the lattice can be done with results agreeing with analytical prediction up to $\theta = \pi$. Next, we consider the application of the generalized thimble method to real-time quantum mechanics. We discuss the new techniques, including the backpropagating HMC and the preconditioned flow equation, that makes it possible to perform the simulation. We also discuss some results for time evolution under quartic and double-well potentials. And finally, we discuss the application of the tensor renormalization group method for two-dimensional gauge theory. A criterion for restricting the number of representations in the character expansion construction is introduced. We also discuss the behavior of singular values at large-N and finite θ. A new interpretation of volume independence based on the tensor-network viewpoint is also introduced.

Contents

Chapter 1

Introduction

1.1 Nonperturbative approach to physics

From statistical mechanics to quantum theory, path integration is an important concept that stands at the center of contemporary physics. Although the detailed interpretations are different, the key object—the partition function—can always be similarly written in the same form:

$$Z = \int dx \exp(-S[x]). \tag{1.1.1}$$

Here, x is the configuration; the microscopic description of the physical state and the action $S[x]$ can be either real or complex. In general, it is not possible to perform this integration directly. A common practice is to consider a system of interest as a small perturbation of another system where one can evaluate the path integral exactly. This perturbative approach proves to be very successful as it has been used throughout the development of the standard model of particle physics. Indeed, the several orders of perturbation for quantum electrodynamics give one of the most precise predictions in physics [1, 2].

Despite its overwhelming success, there are still many questions that cannot be answered with perturbative formalism. Importantly, it has been experimentally observed that quantum chromodynamics has a mass gap, which is a result of the confinement of quarks and gluons [3, 4]. This effect is invisible to perturbative quantum field theory. In superstring theory, developments during the 'second superstring revolution' taught us that some form of nonperturbative formulation is required for the theory to make sense [5–11]. And apart from these, the study of nontrivial vacua [12] or topological quantum field theories [13, 14] cannot be understood without nonperturbative treatments.

To study quantum field theory non-perturbatively, a common approach is to define it on a lattice [15, 16]. Specifically, field configuration $\phi(x)$, which is originally defined on the continuous space-time, is 'discretized' as $\phi(n)$ where n is the lattice site on a finite-volume lattice. In turn, this makes the number of degrees of freedom finite, which makes the path integral rigorously well-defined. This also allows us to study it on a computer. With the use of Monte Carlo methods, lattice field theory simulations have been producing many important nonperturbative results in quantum field theories, condensed matter physics,

and superstring theories [16–24].

1.2 Monte Carlo simulation

1.2.1 Markov chain Monte Carlo

As previously mentioned, Monte Carlo simulation is an important tool for evaluating the multidimensional integrals of the form

$$\langle \mathcal{O}(x) \rangle = \frac{\int d^N x \mathcal{O}(x) e^{-S(x)}}{\int d^N x e^{-S(x)}}. \tag{1.2.1}$$

Naively, one can attempt to perform a discrete approximation to evaluate such an integral. An immediate drawback to this is that the cost of the computation grows exponentially with the number of degrees of freedom of the system. Let us assume for a moment that the action $S(x)$ is real. In this case, one can generate a set of configurations $\{x_i\}$ with probability

$$P(x) = \frac{e^{-S(x)}}{\int d^N x e^{-S(x)}}. \tag{1.2.2}$$

The integral (1.2.1) can then be approximated by a summation

$$\langle \mathcal{O}(x) \rangle = \lim_{N_{\text{samp}} \to \infty} \frac{1}{N_{\text{samp}}} \sum_{i=1}^{N_{\text{samp}}} \mathcal{O}(x_i). \tag{1.2.3}$$

This technique is known as Monte Carlo integration. The question now boils down to how to effectively draw an ensemble of configurations from an arbitrary distribution $\rho(x) = \exp(-S(x))$.

One of the most popular methods for this task is the Markov chain Monte Carlo (MCMC). By constructing a Markov chain with some specific equilibrium, one can obtain an ensemble of configurations directly from the states of the chain. With the length of the chain approaching infinity, the collection of states obtained will start to approach the desired distribution. Let us first understand what exactly is a Markov chain. A Markov chain describes an evolution of a state x to another state x' with some specific transition probability $P(x'|x)$. Namely, if the probability that a state is given by x at step n is $P(x;n)$, then it follows that the probability that it will arrive at x' in the next step is given by the Chapman–Kolmogorov equation

$$P(x';n+1) = \int dx P(x'|x) P(x;n). \tag{1.2.4}$$

The limiting/stationary distribution $P(x)$ can then be defined by

$$P(x) = \lim_{n \to \infty} P(x;n) \tag{1.2.5}$$

If such a limit exists and is uniquely equal to our target distribution, this Markov chain can thus be used as the generator for our configuration ensemble. The Markov chain can

be guaranteed to have a unique stationary distribution if it satisfies the following two conditions [25]:

1. The detailed balance condition: the probability of being in a state x and transitioning to x' is equal to the probability of being in a state x' and transitioning to x

$$P(x'|x)P(x) = P(x|x')P(x'). \tag{1.2.6}$$

2. The stationary distribution is unique. This will be the case if the Markov chain is irreducible; i.e. any two states in the chain are connected, and all of the states in the chain are ergodic.

In the following sections, we will discuss some of the important MCMC algorithms.

1.2.2 Real Langevin algorithm

In this algorithm (see, e.g. Ref [26]), the Markov chain is governed by a stochastic equation of motion known as the Langevin equation

$$\dot{x} = -\frac{\partial S}{\partial x} + \eta(t). \tag{1.2.7}$$

Here, $S(x) = -\log P(x)$, and t is the fictitious Langevin time which can be discretized into the step n of the chain, and $\eta(t)$ is a Gaussian random number satisfying

$$\langle \eta(t)\eta(t') \rangle = 2\delta(t - t'). \tag{1.2.8}$$

It is known that the probability distribution at time t; $P(x;t)$, evolves according to the Fokker-Planck equation

$$\dot{P}(x;t) = \frac{\partial}{\partial x}\left(P\frac{\partial S}{\partial x} + \frac{\partial P}{\partial x} \right). \tag{1.2.9}$$

For the derivation of this equation, see Appendix A. The stationary solution to this equation can be obtained by setting the left-hand side to zero, which can be quickly verified to be

$$P(x) = e^{-S(x)}. \tag{1.2.10}$$

This concludes that the resulting distribution drawn from the chain of configurations $x(t)$ indeed matches with our desired target distribution $\rho(x) = \exp(-S(x))$.

1.2.3 Metropolis-Hastings algorithm

This algorithm [27] is composed of two steps:

1. Starting with the configuration $x_n = x$, propose a new configuration x' according to some criteria that must be symmetric. Namely, if x' is proposed according to an arbitrary density $g(x'|x)$, then it is required that $g(x'|x) = g(x|x')$; e.g. the Gaussian distribution $g(x'|x) \sim \exp(-a(x - x')^2)$.

2. Define the acceptance ratio

$$\alpha = \frac{P(x')}{P(x)} \tag{1.2.11}$$

and generate a uniform random number $u \in [0,1]$. If $u \leq \alpha$, the candidate configuration x' is accepted as the next state of the chain; $x_{n+1} = x'$. If $u > \alpha$, the proposed configuration is rejected, and the previous configuration is used in the next step of the chain; $x_{n+1} = x$.

To show that this specific Markov chain satisfies the two conditions of having a unique stationary distribution, we rewrite the detailed balance condition as

$$\frac{P(x'|x)}{P(x|x')} = \frac{P(x')}{P(x)}. \tag{1.2.12}$$

In case of the Metropolis-Hastings algorithm, we can write the transition probability as a product of the proposal probability $g(x'|x)$ and the acceptance probability $\alpha(x', x)$, from which we can rearrange the detailed balance condition as

$$\frac{\alpha(x', x)}{\alpha(x, x')} = \frac{P(x')g(x|x')}{P(x)g(x'|x)} = \frac{P(x')}{P(x)}. \tag{1.2.13}$$

The second equality follows from the fact that $g(x'|x)$ is symmetric. The acceptance function according to the algorithm given above can be written as

$$\alpha(x', x) = \min\left(1, \frac{P(x')}{P(x)}\right). \tag{1.2.14}$$

One can then straightforwardly show that (1.2.13) is always satisfied with this choice of acceptance function.

1.2.4 Hybrid Monte Carlo algorithm

Instead of using an ad hoc proposal probability $g(x'|x)$, which is running risk of giving a bad acceptant rate if choosen poorly, one can instead obtain a proposal configuration from a fictitious dynamics that keeps the acceptant rate as high as possible. One such algorithm is the Hybrid/Hamiltonian Monte Carlo (HMC) algorithm which was introduced for efficient calculations in lattice quantum chromodynamics [28]. To do this, one introduces an auxiliary momentum to the integral:

$$\int d^N x e^{-S(x)} \longrightarrow \int d^N x d^N p e^{-p^2/2 - S(x)}. \tag{1.2.15}$$

The pair (x, p) can then be interpreted as a canonical pair of a Hamilton system with the Hamiltonian

$$H(x, p) = \frac{p^2}{2} + S(x). \tag{1.2.16}$$

To generate the chain $\{x_n\}$, one first randomize the momentum p with the distribution $P(p) \sim e^{-p^2/2}$. Next, one performs the fictitious time evolution of the pair (x, p) with $x = x_n$ according to Hamilton's equations. Since Hamilton's dynamics is reversible (implying that the corresponding $g(x'|x)$ is symmetric) and strictly preserves the Hamiltonian, which now acts as the action of the auxiliary system, the acceptant ratio (1.2.11) is kept at 1. The final configuration (x', p') is then used as the next configuration in the chain; $x_{n+1} = x'$ with p' discarded and re-randomized in the next step of Markov chain.

The Hamilton's equations are given by

$$\begin{aligned} \frac{dx}{ds} &= p, \\ \frac{dp}{ds} &= -\frac{\partial S}{\partial x} \equiv F(x), \end{aligned} \tag{1.2.17}$$

where s is the fictitious Hamiltonian time. The standard discretization of these equations are the leapfrog integrator:

$$\begin{aligned} p(s + \Delta s/2) &= p(s) + \frac{\Delta s}{2} F(x(s)), \\ x(s + \Delta s) &= x(s) + \Delta s p(s + \Delta s/2), \\ p(s + \Delta s) &= p(s + \Delta s/2) + \frac{\Delta s}{2} F(x(s + \Delta s)). \end{aligned} \tag{1.2.18}$$

Due to the discretization error (which is of order Δs^2 in the case of leapfrog), the Hamiltonian is not exactly conserved during the process. Therefore, it is naturally expected that the acceptant ratio is not 1 at finite Δs. But the important thing is we can always adjust Δs so that the acceptant ratio is as high as desired.

1.3 Numerical sign problem

So far, we assume that the weight is real and positive, which can then be interpreted as a probability density. However, it is often the case that the weight becomes complex. The most well-known example is when we study QCD with nonzero chemical potential [29]

$$Z = \int DAD[\psi, \bar{\psi}] \exp\left\{-S_g[A] - \int d^4x \left(\bar{\psi}(\not{D} - m)\psi - \mu\psi^\dagger\psi\right)\right\} \tag{1.3.1}$$

$$= \int DAD[\psi, \bar{\psi}] e^{-S_g[A] - \int \bar{\psi} M(\mu) \psi}, \tag{1.3.2}$$

where $S_g[A]$ is the pure gauge action and $\not{D}$ is the gauge-covariant Dirac operator. With the traditional Monte Carlo method, we cannot directly simulate the fermionic degrees of freedom. Instead, we directly evaluate the fermionic integral with Berezin integration, giving

$$Z_{\mathrm{QCD}} = \int DA \det M(\mu) e^{-S_g[A]}. \tag{1.3.3}$$

The fermion determinante det$M$ is known to be complex at finite μ. Essentially, we are considering the complex-valued effective action

$$S = S_g - \log\det M \tag{1.3.4}$$

with the partition function

$$Z_{\rm QCD} = \int DA e^{-S[A]}. \tag{1.3.5}$$

To perform these kind of simulations, one uses the reweighting method to calculate the expectation value. Namely, one can separate the real and imaginary parts, and use $\exp(-{\rm Re}S)$ as the probability distribution instead. In general, if x is the configuration of the system, the reweighting method is

$$\langle\mathcal{O}(x)\rangle_S = \frac{\int d^N x \mathcal{O}(x) e^{-i{\rm Im}S(x)} e^{-{\rm Re}S(x)}}{\int d^N x e^{-i{\rm Im}S(x)} e^{-{\rm Re}S(x)}} = \frac{\langle\mathcal{O}(x)e^{-i{\rm Im}S(x)}\rangle_{{\rm Re}S}}{\langle e^{-i{\rm Im}S(x)}\rangle_{{\rm Re}S}}. \tag{1.3.6}$$

The biggest problem with this method happens when the phase $\exp(-i{\rm Im}S)$ fluctuates violently. As a result, the expectation value in the denominator, $\langle\exp(-i{\rm Im}S)\rangle_{{\rm Re}S}$, sums up to a value smaller than its statistical error. And consequently, the statistical error for $\langle\mathcal{O}(x)\rangle_S$ itself becomes hopelessly large.

1.3.1 Phase quenching and the Silver Blaze phenomenon

Instead of dealing with the highly fluctuating phase, one might be tempted to simply remove it by hand. This is known as phase quenching. Such a process can be and is usually done as a temporary measure for dealing with the sign problem in the simulations of gauge theories and matrix models with varying success [21, 30–35]. However, it is known that the phase-quenched approximation is not justified in general. A good example is the two-flavor QCD with finite chemical potential. In contrast to the usual partition function (1.3.3), the quenched theory is represented by

$$Z_{|\rm QCD|} = \int DA |\det M(\mu)| e^{-S_g[A]}. \tag{1.3.7}$$

However, this theory can be shown to be equivalent to 2-flavour QCD with isospin chemical potential, where one assigns chemical potential of the opposite signs for the two quark flavors. Note that, unlike the original theory, nonzero μ doesn't generate a baryon number, but isospin instead. This fact already implies an inconsistent prediction between the two theories. In the original theory, since the lightest particle with a nonzero baryon number is the nucleon, we expect nothing to happen with $\mu < (m_N - B)/3$, where m_N is the nucleon mass and B is the nuclear binding energy. Such a curious feature of the theory is widely known as the Silver Blaze phenomenon [36], named after a Sherlock Holmes story where a 'curious incident' of a dog doing nothing during the night can give a clue to something important. On the other hand, since the lightest particle with nonzero isospin is not the nucleon, but the pion, we expect that pion already starts to condense at $\mu = m_\pi/2 < (m_N - B)/3$, where m_π is the mass of the pion.

1.3.2 Complex Langevin method

One of the earlier attempts to solve the sign problem was proposed in Ref. [37] by considering the complex version of the Langevin algorithm. It makes an assumption that, by holomorphically complexifying the action and allowing the configurations to move into the complex plane, there would be a stationary distribution that gives a correct expectation value

$$\langle \mathcal{O}(x) \rangle = \frac{\int dx \mathcal{O}(x) e^{-S(x)}}{\int dx e^{-S(x)}} \stackrel{?}{=} \frac{\int dx dy \mathcal{O}(x+iy) P(x,y)}{\int dx dy P(x,y)}. \tag{1.3.8}$$

Ignoring the concern regarding the equivalence above, this method is very easy to be implemented. The only task one has to do is to calculate the drift term for each of the configurations and update them using the discretized Langevin equation (see (3.1.3) below).

Earlier implementations of the CLM to various exactly solvable systems were successful in some parameter regions [38–42], although it was not understood at the time what the exact condition for correct convergence really is. It was only recently that the underlying problem was understood [26, 43–47]. Additional techniques such as gauge cooling [48] and deformation technique [49] were also proposed to mitigate this problem. With this new knowledge, the unreliable aspect of the CLM was removed and many great progresses in the field of finite-density QCD were made in this direction [50–54].

In Chapter 3, we will discuss the implementation of the CLM method on 2D gauge theory with a θ term, which suffers from a severe sign problem.

1.3.3 Lefschetz thimble approach

An alternative approach to solving the sign problem is through the use of Picard-Lefschetz theory [55, 56]. Although its usefulness for evaluating highly oscillating integrals has been known for a long time, its application to physics was pointed out in the context of Chern-Simons and Liouville theory [57,58]. The basic idea is to analytically deform the integration manifold into the complex plane in a specific way so that the oscillation of the weight becomes minimized. This specifically deformed manifold is known as the Lefschetz thimbles. Picard-Lefschetz theory states that the imaginary part of the action on each of these thimbles is constant so the fluctuation from the action becomes minimized[1]. A typical example of its usefulness is its application to the Airy function, which can be defined via the oscillating integral

$$\mathrm{Ai}(s) = \int_{-\infty}^{+\infty} \frac{dx}{2\pi} \exp\left(i\frac{x^3}{3} + isx \right). \tag{1.3.9}$$

The corresponding action for the Airy function is

$$S(x) = -i\frac{x^3}{3} - isx, \tag{1.3.10}$$

[1] In general, there is also a complex phase coming from Jacobian determinant that we should keep our eyes on. This will be further discussed in Chapter 4.

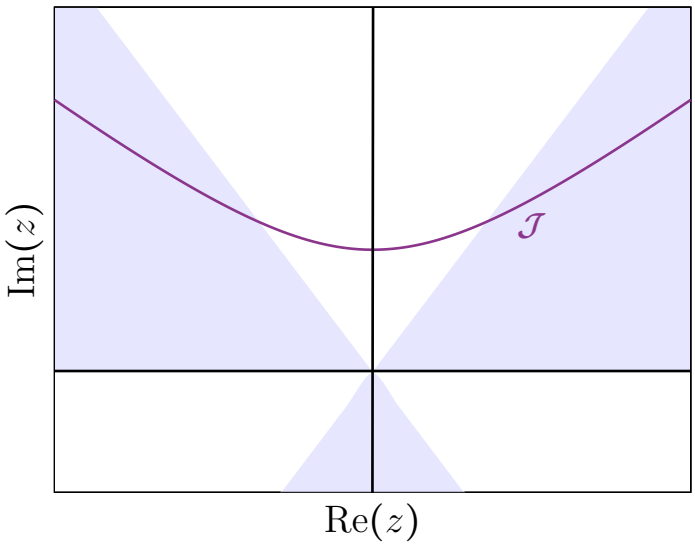

Figure 1.1: The Lefschetz thimble of the Airy integral Ai(s) with $s > 0$. The shaded area represents the direction which has a nice convergent property Re$S \to \infty$.

which is pure imaginary for real x. By deforming the original manifold, which is $\mathbb{R}$, into the Lefschetz thimble $\mathcal{J}$ (see Fig. 1.1), the integral has a nice convergent property Re$S \to \infty$. Due to the compactness of the support, the integrand no longer oscillates infinitely. Thus, the sign problem is solved.

Recently, this technique has been gaining a lot of attention in various directions such as in the context of Thirring model, the toy model for finite density QCD [59, 60], quantum tunneling [61–63], and quantum gravity [64]. Some suitable numerical technique is thus required in order to explore the case with large degrees of freedom.

Although the thimble method promises to eliminate the sign problem on the thimble, it is also known that performing a Monte Carlo simulation with thimbles is not straightforward. This is mainly because an infinitely high potential barrier between any two thimbles prevent the configuration from crossing. This essentially breaks ergodicity of the simulation. Also, the traditional implementation of the HMC algorithm for thimble calculation is also known to be significantly slower than the same calculation via the complex Langevin method due to the need to calculate the Jacobian matrix. We will discuss these issues and how to overcome them in Chapter 4.

1.3.4 Tensor renormalization group approach

Another approach known as the tensor renormalization group [65] and its variants is another promising direction to solving the sign problem. In this approach, we avoid the Monte Carlo algorithm entirely. And because of its non-stochastic nature, there is no sign problem by construction. This technique was introduced to calculate the partition function of many-body systems in lower dimensions. If one can represent the partition function as a tensor network, one can perform coarse-graining and remove irrelevant degrees of freedom through

volume scaling. In this sense, the TRG approach is also attractive because it can access large volumes very easily. Another advantage of the TRG approach is that, unlike the Monte Carlo methods, the fermionic degrees of freedom can be handled directly without having to be integrated out first [66–76]. This means that we don't have to deal with the fermion determinant, which is usually costly and time consuming.

On the other hand, the drawback of the TRG method is that the cost of performing coarse-graining is usually growing exponentially with the space-time dimensions [77–79]. And only local observables can be calculated, unlike Monte Carlo simulations where one can calculate practically any observables. It is also noticed that the TRG method might not be suitable to study the system near the critical point, where interactions are long-range and non-local. Nevertheless, the TRG method is still a powerful tool that helps us learn about physics from a new perspective as will be demonstrated in Chapter 5.

1.4 Outline of the thesis

In Chapter 2, I describe the physical importance of two of the topics that will be mainly discussed in this thesis. The first of which is the two-dimensional gauge theory with a θ term, which has an interesting vacuum structure in the T-θ plane. In the second topic, real-time quantum mechanics, I outline the path integral formulation. Then I describe the tunneling problem from various points of view and explain ongoing works.

In Chapter 3, I describe the complex Langevin method in more detail, its justification, and its application to the 2D U(1) gauge theory. I explain the difficulties in the naive implementation of the method and how to overcome them by modifying the lattice topology.

Chapter 4 describes the Lefschetz thimble method and various technical developments for numerical simulations. In particular, I explain the backpropagating HMC algorithm that helps perform the simulation on the original contour at a fast speed. I also describe a newly developed preconditioning formalism that helps regulate the flow equation at large flow time. And finally, I present the numerical result for the real-time evolution of a wave function under the quartic and double-well potential.

In Chapter 5, the tensor renormalization group approach is explained. Various implementations for 2D gauge theories are also described. I present a way to cut off the number of irreducible representations when the tensor is constructed via character expansion. I then give the analysis of the singular values at large N and finite θ. A new interpretation of volume independence from the viewpoint of TRG is also given.

Chapter 2

Selected systems with the sign problem

2.1 Two-dimensional gauge theories with a θ term

One of the main goals of the lattice community is to explore quantum chromodynamics in all parameter regions. However, many of these interesting regions are still inaccessible due to various technical difficulties, including the sign problem. Thus, it is insightful to study some semi-realistic models with similar sign problems and develop useful tools to handle them. Two-dimensional pure gauge theory is one of the most suitable theories for this. It is exactly solvable[1] for any gauge group, on any manifold, and with any finite volume [80–82]. And one particular theory that has been gaining a lot of attention in the past few years is the theory with a θ term.

The physics of θ vacua was first noticed in the context of the strong CP problem. In physics, it was widely believed that 'everything allowed by the laws of nature must actually exist' [83]. And one such thing is the θ term in the QCD action

$$S_\theta = -i\theta Q; \tag{2.1.1}$$

$$Q = \frac{1}{32\pi^2}\epsilon_{\mu\nu\rho\sigma}\int d^4x \mathrm{tr} F^{\mu\nu}F^{\rho\sigma}. \tag{2.1.2}$$

The quantity Q is known as the topological charge and takes integer values on a compact space. This θ term is renormalizable by power counting and is thus perfectly sensible to be included in the QCD action. It, however, breaks parity and time-reversal symmetries, and hence the CP symmetry. The existence of the θ term implies a non-vanishing electric dipole moment of a neutron, which is not observed in experiments—the upper bound on the θ angle obtained so far is $|\theta| \lesssim 10^{-10}$ [84]. Such unreasonably small value without a good reason is the stem of the strong CP problem. A popular solution to this is the Peccei-Quinn mechanism [85–88]. It introduces a pseudo-Nambu-Goldstone boson of a hypothetical global

[1] One variation of the derivation can be done by representing the partition function as a tensor network using character expansion, which is explored in the case of a two-dimensional torus in Section 5.2.3.

$U(1)_{PQ}$ symmetry, known as the axions. In this mechanism, the potential for the axions induced by QCD automatically chooses a vacuum to be CP invariant, namely with $\theta = 0$.

2.1.1 θ vacua

The θ term itself is interesting from the theoretical point of view. To illustrate this, let us discuss the 'hidden' vacua of gauge theory. Consider the topological charge in two dimensions;

$$Q = \frac{1}{4\pi}\int d^2x \epsilon_{\mu\nu} \mathrm{tr} F^{\mu\nu}, \tag{2.1.3}$$

which can be descritized on a lattice as

$$Q = \frac{1}{2\pi i}\sum_{n\in\Lambda} \log \det P_n \tag{2.1.4}$$

The logarithm here is defined on the principal branch; $-\pi \leq \mathrm{Im}\log x < \pi$, and the plaquette P_n is a parallel transport along a closed loop at site n:

$$P_n = U_{1,n} U_{2,n+\hat{1}} U^\dagger_{1,n+\hat{2}} U^\dagger_{n,2}; \tag{2.1.5}$$

$$U_{\mu,n} = \exp(i\epsilon A_\mu(\vec{n}\epsilon)). \tag{2.1.6}$$

In the continuum limit, the plaquette is therefore nothing but the curvature $F_{\mu\nu}$, which leads us back to the continuum definition (2.1.3).

One can show that (2.1.4) actually only takes integer values. First, let φ_n denotes the U(1) part of the plaquette P_n, then

$$Q = \frac{1}{2\pi i}\sum_{n\in\Lambda} \log e^{iN\varphi_n} = \frac{1}{2\pi i}\left(2\pi i q + \log\prod_{n\in\Lambda} e^{iN\varphi_n}\right) = q. \tag{2.1.7}$$

The additional term $2\pi i q$ comes from possible branch crossing after combining two principal logarithms[2]. Also, the second term in the parenthesis vanishes since it topologically connected to the trivial configuration. Next, its derivative also vanishes which can be seen by first separating the U(1) phase of the link variable from the SU(N) part, then

$$\delta Q = \frac{1}{2\pi i}\sum_{n\in\Lambda} iN(\delta\varphi_{1,n} + \delta\varphi_{2,n+\hat{1}} - \delta\varphi_{1,n+\hat{1}} - \delta\varphi_{2,n}). \tag{2.1.8}$$

Because of the summation over the lattice site, all four terms exactly cancel. As a result, the topological charge can take any integer value.

Formally, the topological charge can be considered as a winding number in the context of the homotopy group of the gauge transformation Ω. For example, in the 4D SU(2) gauge theory, the gauge transformation is a map from a spatial manifold $\mathbb{R}^3 \cup \{\infty\} \cong S^3$ to a group manifold $\mathrm{SU}(2) \cong S^3$. The associating homotopy group is $\pi_3(\mathrm{SU}(2)) = \mathbb{Z}$. As such, for each

[2]For example, $1.8i\pi = \log(\exp(0.9i\pi)) + \log(\exp(0.9i\pi)) = 2i\pi + \log(\exp(1.8i\pi)) = 2i\pi + (-0.2i\pi)$.

Ω, we can count the number of times, n, the 'spatial' S^3 is mapped onto the 'group' S^3. Gauge transformations Ω_n with $n = 0$ are known as the small gauge transformations, while those with $n \neq 0$ are known as the large gauge transformation. Since we can classify the gauge field configurations into different sectors, identified by the winding number q, the gauge theory then has many vacuum states $|q\rangle$, referred to as topological vacua.

For a state to be the vacuum state of the gauge theory, it has to be an eigenstate of both the small and large gauge transformations. The topological vacua $|q\rangle$ are, however, not the vacua of the gauge theory since it transforms under the large gauge transformation Ω_n as $\Omega_n|q\rangle = |q+n\rangle$. However, we can define the θ-vacua

$$|\theta\rangle = \sum_q e^{iq\theta}|q\rangle, \tag{2.1.9}$$

which transforms under Ω_n as $\Omega_n|\theta\rangle = e^{-in\theta}|\theta\rangle$. θ-vacua, parametrized by the angle θ, are thus the vacua of the gauge theory.

One can solve the equation of motion for the gauge theory and obtain classical solutions with finite action known as instantons. These instantons, each can be associated with winding number ν, are what responsible for a tunneling from one topological vacuum to another [89], i.e. $|n\rangle \to |n+\nu\rangle$. θ-vacua would be degenerate without these tunnelings. Because of the instantons, the ground state energy depends explicitly on θ, making various θ-vacua physically different from each other.

2.1.2 Phase diagram

Gauge theory with non-zero topological angle θ is given by the partition function

$$Z_\theta = \int dA e^{-S_g[A]+i\theta Q[A]}. \tag{2.1.10}$$

Since Q takes integer values, the partition function and all of its derivatives with respect to θ must be periodic in θ with the period of 2π. In two dimensions, the partition function can be evaluated exactly and is thoroughly understood [82, 90]. In particular, (2.1.10) in the continuum limit can be shown to take the form

$$Z_\theta = \sum_{q\in\mathbb{Z}} \exp\left\{-\frac{V\chi}{2}(\theta - 2\pi q)^2\right\}, \tag{2.1.11}$$

where V is the physical volume of the system and χ is the topological susceptibility which will be defined in later sections. It can easily be shown that this function develops kinks in the large volume limit at θ equals to odd-multiples of π, as shown in Fig. 2.1. These special points have physical significants—they are the points where parity symmetry is spontaneously broken. There are many pieces of evidence that similar kinks are also developed in 4 dimensions [91, 92] for SU(2). In fact, it is shown to be the case at large N in Ref. [93–95]. In these cases, the kinks correspond to the points where CP symmetry is spontaneously broken.

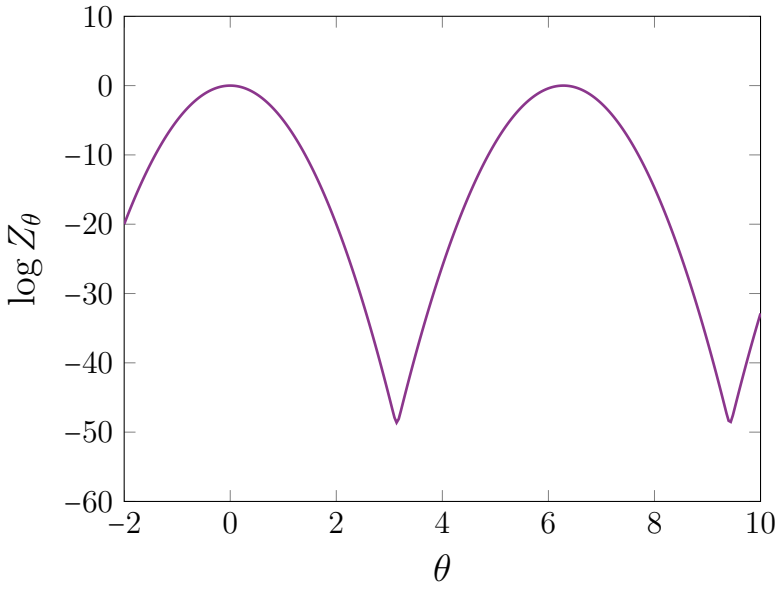

Figure 2.1: Partition function with nonzero θ and $V\chi = 10$.

Note that the story so far concerns the quantum field at zero temperature. At high temperature, which corresponds to the case where the length of the compactified imaginary time becomes small, it is a different story. When the temperature is sufficiently high, two important things happen: the gauge field becomes deconfined, and the 'instanton gas' becomes diluted [96]. It is known that in such a region, the kinks that we discuss so far do not appear at any θ.

The study from large-N analysis, holography, and 't Hooft anomaly matching condition helps us shed some light on the T-θ phase diagram [97–104]. At large N, the diagram is already understood. Namely, the region are separated horizontally by the deconfinement transition at $T_{\rm dec}(\theta)$ which can depend on the topological angle θ. And in the confined region, there is another line which indicates the first-order phase transition at $\theta = \pi$ going from $T = 0$ to $T = T_{\rm dec}(\pi)$. In other words, we have $T_{\rm dec}(\pi) = T_{\rm CP}$. The diagram is shown in Fig. 2.2 (Left). However, there are still some confusion left at finite N. In particular for SU(2), most of the diagram looks the same as the large-N case's, except at the CP restoration point. A study based on supersymmetric Yang-Mills theory suggests that $T_{\rm dec}(\pi) < T_{\rm CP}$ [97]. It is thus important to investigate gauge theory close to $\theta = \pi$ from first principles.

As a first step toward this interesting problem in four dimensions, in one of our works [105], we consider the application of the CLM to the two-dimensional U(1) lattice gauge theory with a θ term, which still suffers from the sign problem despite its simplicity. The naive reweighting method, for instance, can reach up to $\theta \sim 2.2$ on a 16×16 lattice [106]. The region near $\theta = \pi$ is thus still a seemingly difficult task for this method, especially at larger lattice sizes. And our other work [90], we investigate a more general 2D Yang-Mills theory using the tensor renormalization group method. Both of our works can reproduce

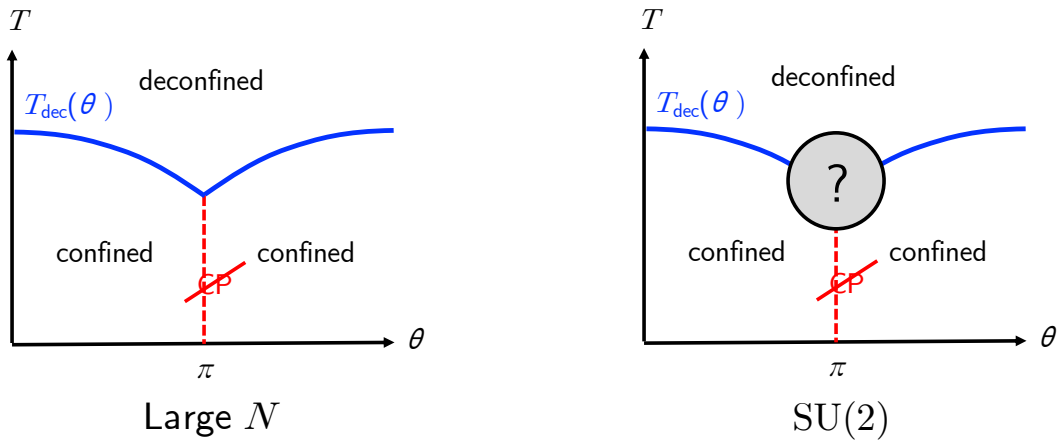

Figure 2.2: T-θ phase diagrams of 4D Yang-Mills theory. (Left) The diagram obtained at large N. (Right) There is still an unknown region near the CP restoration point at $\theta = \pi$ that requires numerical investigation.

satisfying results up to $\theta = \pi$.

2.1.3 Definition on a two-dimensional lattice

Let us go into details of the definition of the theory. In two dimensions, gauge theory with a θ term is given by a partition function

$$Z = \int DU e^{-S_g[P]+i\theta Q[P]}; \tag{2.1.12}$$

$$S_g[P] = -\frac{N}{\lambda}\sum_{n\in\Lambda}\mathrm{tr}(P_n + P_n^\dagger), \tag{2.1.13}$$

$$P(n) = U_{n,1}U_{n+\hat{1},2}U^\dagger_{n+\hat{2},1}U^\dagger_{n,2}. \tag{2.1.14}$$

where N is the rank of the gauge group $\mathrm{U}(N)$ or $\mathrm{SU}(N)$ and $\lambda = 2Ng^2\epsilon^2$ is the 't Hooft coupling constant. Out of the large-N context, it is more convenient to use the inverse temperature β instead of the 't Hooft coupling, which are related via

$$\beta = \frac{N^2}{\lambda}. \tag{2.1.15}$$

As for the topological charge, there are multiple possible definitions, for which we will consider two of them: the 'log' definition and the 'sine' definition [105].

$$Q_{\log} = \frac{1}{2\pi i}\sum_{n\in\Lambda}\log P_n \sim \frac{1}{2\pi}\sum_{n\in\Lambda}\epsilon^2 F_{12}(\vec{n}\epsilon), \tag{2.1.16}$$

$$Q_{\sin} = \frac{1}{2\pi i}\sum_{n\in\Lambda}(P_n - P_n^{-1}) \sim \frac{1}{2\pi}\sum_{n\in\Lambda}\sin\left(\epsilon^2 F_{12}(\vec{n}\epsilon)\right). \tag{2.1.17}$$

We have already discussed the first definition above. The second definition, however, does not take integer values at finite lattice spacing ϵ. The two definitions will only coincide in the continuum limit. But the sine definition is actually important in the sense that it is analogous to the clover-leaf definition of the topological charge, which is one of the most popular definitions in four dimensions based on Wilson loops [107].

2.2 Real-time quantum mechanics

Perhaps the most basic system with the sign problem is nothing but simple real-time quantum mechanics in one dimension. The first course of quantum mechanics taught us that a wavefunction ψ in a system defined by a Hamiltonian H is evolved according to the Schrödinger equation

$$\frac{\partial}{\partial t}\psi = -iH\psi. \tag{2.2.1}$$

If the Hamiltonian is not explicitly time-dependent, the solution to this equation is

$$\psi(T) = \exp(-iHT)\psi(0). \tag{2.2.2}$$

Practically, one can perform the matrix exponentiation via matrix diagonalization

$$e^{-iHT} = U \cdot \mathrm{diag}(e^{-i\lambda_1 T}, .., e^{-i\lambda_N T}) \cdot U^{-1}; \tag{2.2.3}$$

$$H = U \cdot \mathrm{diag}(\lambda_1, .., \lambda_N) \cdot U^{-1}, \tag{2.2.4}$$

where N is the size of the Hilbert space. The cost of this process is usually of order $\mathrm{O}(N^3)$ in complexity, but it is a reliable numerical methods for solving the Schrödinger equation.

2.2.1 Path integral quantization

An alternative approach to solving the Schrödinger equation is to use path integral [108], which can be advantageous over direct diagonalization if the size of the Hilbert space is very large. Note that the time evolution operator can be written as a transition amplitude

$$\langle y|e^{-iHT}|x\rangle \equiv K(x,y;T) = \int_{x(0)=x}^{x(T)=y} Dx \exp\left(iS_{\mathrm{QM}}[x;T]\right). \tag{2.2.5}$$

Here, the function $S_{\mathrm{QM}}[x;T]$ is the time-integral of the Lagrangian;

$$S_{\mathrm{QM}}[x;T] = \int_0^T dt \left\{\frac{1}{2}\dot{x}^2 - V(x)\right\} \tag{2.2.6}$$

which is known as the action, in the original sense. Here, the measure Dx indicates the sum over all possible function $x(t)$ with a fixed boundary condition $x(0) = x$ and $x(T) = y$.

This transition amplitude is nothing but a partition function, whose associated 'action' is a pure imaginary number $S[x] = -iS_{\mathrm{QM}}[x;T]$:

$$Z(T) = \int Dx e^{-S[x]}. \tag{2.2.7}$$

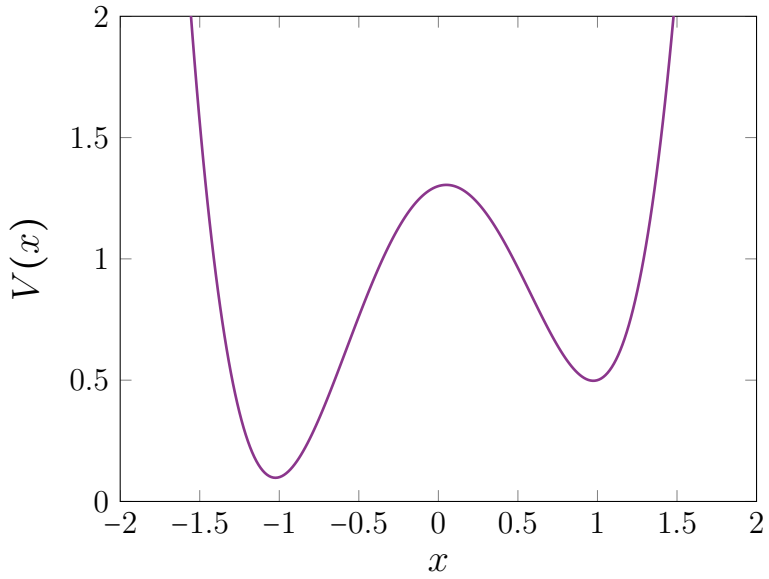

Figure 2.3: An example of a system with an unstable vacuum near $x = 1$ and a stable vacuum near $x = -1$.

This causes a maximum sign problem since $\mathrm{Re}S[x] = 0$, which makes the phase of the weight fluctuate almost uniformly. This system is therefore an attractive target for testing new algorithms aiming to solve the sign problem.

2.2.2 Tunneling problem

One interesting topic in quantum mechanics is the study of quantum tunneling in real-time formalism, which plays an important role in quantum cosmology [109–116]. Assuming that the potential of interest has two local minima (e.g., Fig. 2.3), if an initial state is localized near one with higher energy, the state will decay into the true vacuum given a sufficiently long time. The tunneling rate can be estimated by several methods, which will be given below.

Approximation from energy gap

The most straightforward method is to approximate the initial wave function as a linear combination between the ground state and the first excited state:

$$\psi(0) \approx \frac{1}{\sqrt{2}} \left(\psi_0 + \psi_1\right). \tag{2.2.8}$$

With the time evolution, we have

$$\psi(t) \approx \frac{1}{\sqrt{2}} e^{-iE_0 t} \left(\psi_0 + e^{-i\Delta E t}\psi_1\right), \tag{2.2.9}$$

where $\Delta E = E_1 - E_0$ is the gap between the first excited state and the ground state. The tunneling is deemed complete when the relative phase between the two states flips the sign, namely with

$$t_T = \frac{\pi}{\Delta E}. \tag{2.2.10}$$

This is nothing but the inverse of the tunneling rate[3] Therefore, we have

$$\Gamma \sim \frac{\Delta E}{\pi}. \tag{2.2.11}$$

WKB approximation

A more precise value can be obtained via the WKB approximation. Consider a slowly-varying potential barrier $V(x)$ and a particle of energy E, an approximate solution to the time-independent Schrödinger equation

$$\frac{d^2}{dx^2}\psi(x) = -2(E - V(x))\psi(x) \tag{2.2.12}$$

is

$$\psi(x) \sim \exp\left\{\pm i \int dx \sqrt{2(E - V(x)}\right\}. \tag{2.2.13}$$

In the region where the tunneling happens; i.e., $E \le V(x)$ for $x_1 \le x \le x_2$, the wave function decays exponentially. We can then estimate the tunneling rate via

$$\Gamma = \left|\frac{\psi(x_2)}{\psi(x_1)}\right|^2 \sim \exp\left\{-2\int_{x_1}^{x_2} dx \sqrt{2(V(x) - E)}\right\}. \tag{2.2.14}$$

2.2.3 Coleman's instanton

Coleman [89] introduced an alternative way to compute the transition amplitude through the use of instantons. In this context, we will consider the tunneling that happens during an infinitely long time interval: $[-T/2, T/2]$ with $T \to \infty$. Or more precisely, we consider the imaginary time evolution

$$\langle y|e^{-HT}|x\rangle = \int Dx e^{-S_{\mathrm{QM}}[x;-iT]}. \tag{2.2.15}$$

On the left-hand side, one can see that only the ground state contributes at large T;

$$\langle y|e^{-HT}|x\rangle \approx e^{-E_0 T}\langle y|E_0\rangle\langle E_0|x\rangle. \tag{2.2.16}$$

[3]This can be seen by first writing the particle number density $N(t) = \exp(-t/t_T)$. Then the tunneling rate becomes

$$\Gamma = -\frac{1}{N}\frac{dN}{dt} = \frac{1}{t_T}.$$

As for the right-hand side, since

$$S[x] = S_{\mathrm{QM}}[x; -iT] = \int_{-T/2}^{T/2} dt \left\{ \frac{1}{2}\dot{x}^2 + V(x) \right\} \tag{2.2.17}$$

is real, the dominating part of the path integral is given by the semi-classical path that solves the Euclidean equation of motion

$$\frac{\delta S}{\delta x} = -\ddot{x} + V'(x) = 0. \tag{2.2.18}$$

Note that this equation is different from the Lorentzian equation of motion in that the sign of the potential energy is flipped $V(x) \to -V(x)$. In this sense, we transform the problem of the double well into the problem of the double hill. Indeed, the constant of motion is

$$E = \frac{1}{2}\dot{x}^2 - V(x), \tag{2.2.19}$$

with the potential flipped.

In any case, let $\bar{x}(t)$ be a semi-classical solution to (2.2.18). And let us define an operator

$$\Lambda f = -\frac{\partial^2}{\partial t^2} f + V''(\bar{x}) f \tag{2.2.20}$$

for any function $f(t)$ satisfying the boundary condition $f(\pm T/2) = 0$. We also define x_n to be eigenvectors of Λ with eigenvalues λ_n, also satisfying $x_n(\pm T/2) = 0$. Then we can write any path in the path integral as

$$x(t) = \bar{x}(n) + \sum_n c_n x_n(t). \tag{2.2.21}$$

The path integral measure is then formally given by

$$Dx = \mathcal{N} \prod_n \frac{dc_n}{\sqrt{2\pi}} \tag{2.2.22}$$

for some normalization constant $\mathcal{N}$. We now estimate the path integral in the semi-classical limit as a product of many Gaussian integrals centered at $\bar{x}$, giving

$$\int Dx e^{-S} = \mathcal{N} e^{-S[\bar{x}]} \prod_n \lambda_n^{-1/2} = \frac{\mathcal{N} e^{-S[\bar{x}]}}{\sqrt{\det \Lambda}}. \tag{2.2.23}$$

To summarise, the imaginary-time transition amplitude is given by

$$W(x, y; -iT) \approx e^{-E_0 T} \langle y | E_0 \rangle \langle E_0 | x \rangle \approx \frac{\mathcal{N} e^{-S[\bar{x}]}}{\sqrt{\det \Lambda}}. \tag{2.2.24}$$

When x and y are on the opposite side of the barrier, the transition amplitude is related to the tunneling rate via

$$\Gamma(x \to y) \propto |W(x, y; -iT)|^2. \tag{2.2.25}$$

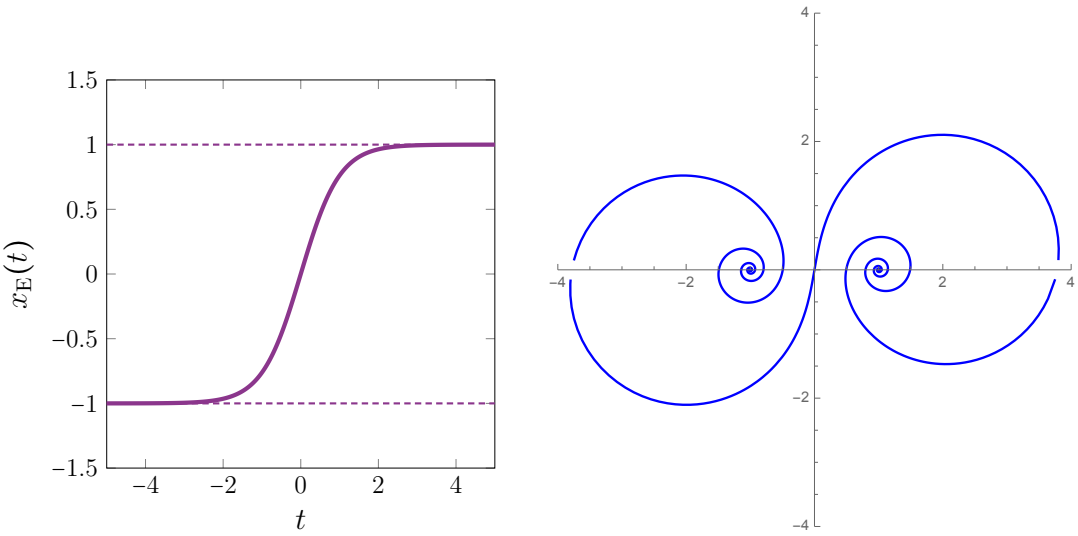

Figure 2.4: (Left) The instanton solution in the imaginary time (Euclidean) formalism associated to the potential $V(x) = 1 + \frac{1}{2}x^4 - x^2$. The dashed lines show the locations of the two vacua. (Right) The parametric plot on the complex plane of the analytic continuation of the instanton solution; $x(t) = x_{\mathrm{E}}(t \times \exp(1.4i))$. The singular behaviour can be seen for $t \to \pm\infty$ as the particle wildly spiralling into the two vacua $x = \pm 1$.

Instanton solution

Consider a symmetric double well potential $V(x)$ with the two vacua located at $x = \pm a$. We will also level the potential so that $V(\pm a) = 0$. As stated above, in the imaginary-time formalism, the potential is flipped and we are considering a double hill instead. There are four particular transitions that are notably interesting: $a \to a$, $-a \to -a$, $a \to -a$, and $-a \to a$. In all of these cases, we will consider the solutions that are asymptotically static; i.e., with $E = 0$, or

$$\frac{dx}{dt} = \sqrt{2V(x)}. \tag{2.2.26}$$

One example of such solutions with the boundary condition $-1 \to 1$ for $V(x) = 1 + \frac{1}{2}x^4 - x^2$ and $T = \infty$ is $x = \tanh(t)$ as shown in Fig. 2.4 (Left). This type of solution is called an instanton as the point of interest only happens at an instant in time (as opposed to a soliton). One can also define an anti-instanton as an instanton with the opposite boundary conditions; $1 \to -1$. And in general, one can combine several instantons and anti-instantons together and get the approximate classical trajectories, as long as the turning points are not too close together and the boundary conditions are unchanged.

What is so special about these (anti-)instanton solutions is that they sit at the local minima of the action, which can be evaluated analytically as

$$S[\bar{x}] = \int_{-T/2}^{T/2} dt \left\{ \frac{1}{2}\dot{\bar{x}}^2 + V(\bar{x}) \right\} = \int_{-T/2}^{T/2} dt \dot{\bar{x}}^2 = \int_{-a}^{a} dx \sqrt{2V(x)}. \tag{2.2.27}$$

The last equality comes from the reparametrization $t \to x(t)$, as well as using the identity

(2.2.26). Note that if we substitute this into the rightmost expression of (2.2.24), we get the exponent similar to that of the WKB approximation's (2.2.14). This is by no means accidental. The careful proof that the instanton approach is equivalent to the WKB approximation is shown in Ref. [89].

2.2.4 Unsolved conundrum: Instantons vs. Sphalerons

According to Coleman's analysis, tunneling occurs via instantons. This is, however, technically in the imaginary time formalism. The question still remains: what is the most contributing trajectory in the real-time path integral? This question cannot be confidently answered without the simulation from the first principle. However, there are several attempts to answer this question.

Koike-Tanizaki's explanation

In Ref. [61], the authors attempted to evaluate the real-time path integral using Lefschetz thimbles. This involves identifying all saddle points on the complex planes. However, to analytically compute the path integral, one must be able to calculate the intersection number $\mathfrak{n}_\sigma$ (see (4.1.4) below), which is a highly nontrivial task. The authors then explore some of the lowest-lying modes at finite time evolution for both real and imaginary time. Not surprisingly, the imaginary time solutions highly resemble the instanton solutions even with finite time intervals. On the other hand, the real-time solutions are those we call the sphaleron, which are the solutions where the particle has the energy barely enough to climb the barrier, rest on the top of the barrier, and then fall down on the other side at the end (see Fig. 2.5). The authors argue that such solutions are real and will not give reasonable suppressions in the real-time path integral. They conjectured that the trajectories that truly contribute to the real-time path integral are those with very large oscillations, which is analogous to the solutions obtained from analytically continuing instantons to real-time. Fig 2.4 (Right) shows an example of the instanton after a partial Wick rotation.

Instanton complexification

The possibility that this type of solutions is important in real-time path integral was already hinted in an earlier work [117]. The authors considered the real-time path integral as the limit of the Wick rotation of imaginary-time path integral. They show that despite the fact that the instanton wildly fluctuates on its way between the two vacua, the action remains constant throughout the complexification. However, it should be noted that although the saddle points can be straightforwardly complexified via analytic continuation, this is not the case for the associated thimbles. On this note, Ref. [62] developed a technique to approximate an analytically continued path integral up to the first quantum corrections. Remarkably, they were able to reproduce the one-loop tunneling amplitude given by Callan and Coleman [111].

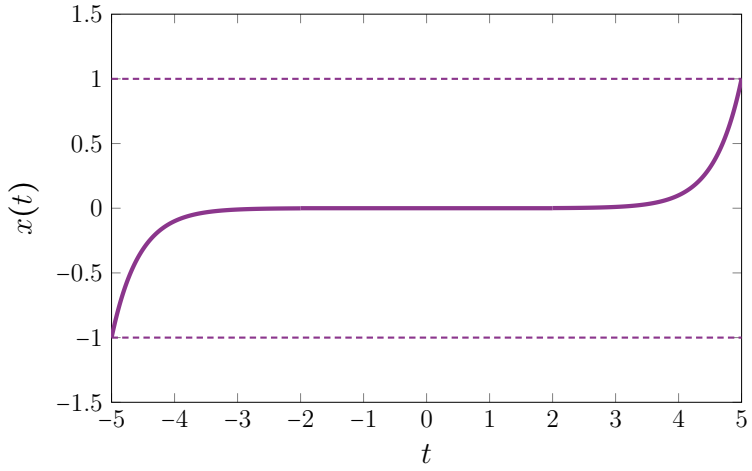

Figure 2.5: Sphaleron solution with the dashed lines showing the location of the vacua.

Mou-Saffin-Tranberg's approach

In Ref. [63], the authors tackle the path integral from a different angle. Namely, they consider the Keldysh formalism

$$\langle \mathcal{O}(t) \rangle = \mathrm{tr}[\rho_0 \mathcal{O}(t)] = \frac{\int Dx \rho_0 e^{iS_{\mathrm{QM}}} \mathcal{O}}{\int Dx \rho_0 e^{iS_{\mathrm{QM}}}} \tag{2.2.28}$$

with ρ_0 representing the initial density matrix. The path of time in the action starts from $t = 0$, runs to some late time, and returns to $t = 0$. The path integral is then split into two parts: the initial density matrix part which involves drawing a trajectory from an ensemble with specific initial conditions, and the dynamic part which is associated with quantum corrections. They found that since the initial points are real, the corresponding trajectory which is unique must also be real, contradicting the previous complexified instanton argument. Furthermore, the calculation of the expected position with this method matches the one solved from direct diagonalization (2.2.3) very well.

Closing remarks

It is thus an open question whether the real (sphalerons) or complex (instantons) trajectories really contribute to the real-time path integral. We discuss this issue in section 4. Using the technique we developed, we were able to compute the path integral using Lefschetz thimbles for the double-well potential when the barrier is not too high. The dominant trajectories so far are regular and are consistent with Mou-Saffin-Tranberg's conclusion. We plan to further investigate the problem with a higher barrier in the future to make a more concrete conclusion.

Chapter 3

Complex Langevin method

3.1 Overview

The idea of the CLM is to generalize the well-studied real Langevin method to the system of complex variables [37]. Similar to the usual Monte Carlo integral, the assumption is that it is possible to generate a set of complex configurations z_i such that the approximation

$$\frac{\int d^N x \mathcal{O}(x) e^{-S(x)}}{\int d^N x e^{-S(x)}} \approx \frac{1}{N_{\text{samp}}} \sum_{i=1}^{N_{\text{samp}}} \mathcal{O}(z_i) \tag{3.1.1}$$

holds for sufficiently large N_{samp}. Note that since the Boltzmann weight $\exp(-S(x))$ is no longer treated as a probability weight, we do not encounter the sign problem even if $S(x)$ is complex.

The configurations z_i are generated as a solution to the complex Langevin equation,

$$\dot{z}(t) = -\frac{\partial S(z(t))}{\partial z(t)} + \eta(t), \tag{3.1.2}$$

where $S(z)$ is the analytic continuation of $S(x)$. The derivative $\partial S/\partial z$ will be refered to as the drift term from now on. In practical calculation, the complex Langevin process is discretized with time interval ϵ, and the configurations are generated via

$$z_{i+1} = z_i - \epsilon \frac{\partial S}{\partial z_i} + \sqrt{\epsilon}\tilde{\eta}_i \tag{3.1.3}$$

with $\langle \tilde{\eta}_i \tilde{\eta}_j \rangle = 2\delta_{ij}$.

3.2 Justification of the method

Although the complex Langevin method is very convenient and fast, it is known not to converge or to converge at the wrong values in some parameter regions. The reason for this has been very well studied [26, 43–47], with a few proposed criteria of correctness. We will discuss two of them here.

3.2.1 Boundary effects

In this analysis (for more careful and complete analysis, see Ref. [43–46]), we define $P(x,y;t)$ to be the *real* probability distribution of the Markov chain after the complexification $x \to x+iy$, and $\rho(x;t)$ to be the original *complex* distribution before the complexification. The expectation values of the two cases are given by

$$\langle \mathcal{O} \rangle_{P(t)} = \int dxdy P(x,y;t)\mathcal{O}(x+iy), \tag{3.2.1}$$

$$\langle \mathcal{O} \rangle_{\rho(t)} = \int dx \rho(x;t)\mathcal{O}(x). \tag{3.2.2}$$

In the first equation, the function $\mathcal{O}(z)$ is a holomorphic generalization of $\mathcal{O}(x)$ in the second equation. We also assume that both $P(x,y;t)$ and $\rho(x;t)$ are normalized. The two distributions can be shown to satisfy the Fokker-Planck equations[1]

$$P(x,y;t) = e^{tL^\top} P(x,y;0), \tag{3.2.3}$$

$$\rho(x+iy;t) = e^{tL_c^\top} \rho(x+iy;0), \tag{3.2.4}$$

where

$$L^\top = \partial_x(\partial_x + \mathrm{Re}\partial S) + \partial_y \mathrm{Im}\partial S, \tag{3.2.5}$$

$$L_c^\top = \partial_x(\partial_x + \partial_x S). \tag{3.2.6}$$

Notice that we have enabled the random walk in the imaginary direction (y) for $L^\top$, which doesn't affect the simulation but is useful in the following analysis. The initial conditions for (3.2.3) and (3.2.4) are assumed to match

$$P(x,y;0) = \rho(x;0)\delta(y-y_0). \tag{3.2.7}$$

Next, we define

$$F_{\mathcal{O}}(t,\tau) = \int dxdy P(x,y;t-\tau) e^{\tau L}\mathcal{O}(x+iy). \tag{3.2.8}$$

At $\tau = 0$, we straightforwardly obtain $F_{\mathcal{O}}(t;0) = \langle \mathcal{O} \rangle_{P(t)}$. And at $\tau = t$ it is equal to $\langle \mathcal{O} \rangle_{\rho(t)}$;

$$\begin{aligned} F_{\mathcal{O}}(t;t) &= \int dxdy P(x,y;0) e^{tL}\mathcal{O}(x+iy) \\ &= \int dxdy \rho(x;0)\delta(y-y_0) e^{tL}\mathcal{O}(x+iy) \\ &= \int dx \left(e^{tL_c^\top}\rho(x;0)\right)\mathcal{O}(x+iy_0) \\ &= \int dx \rho(x;t)\mathcal{O}(x+iy_0) \\ &= \langle \mathcal{O} \rangle_{\rho(t)}. \end{aligned} \tag{3.2.9}$$

[1] For the derivation of the Fokker-Planck equation in case of the real action, see Appendix A.

In the derivation above, we perform a partial integration once in the third line, which is possible because of the assumption that $\rho(x;0)$ must decay exponentially or faster at infinities. As a result, $F_{\mathcal{O}}(t,\tau)$ can be seen as an interpolation between the 'correct' expectation value $\langle\mathcal{O}\rangle_{\rho(t)}$ and the CLM expectation value $\langle\mathcal{O}\rangle_{P(t)}$. If its derivative with respect to τ always vanishes, we have a proof that the two expectation values are the same, meaning that the CLM simulation is correct. Note that

$$\frac{\partial}{\partial t}P(x,y;t) = L^{\top}P(x,y;t), \tag{3.2.10}$$

then we have

$$\frac{\partial}{\partial \tau}F_{\mathcal{O}}(t,\tau) = \int dxdy\left(-L^{\top}P(x,y;y-\tau) + P(x,y;y-\tau)L\right)e^{\tau L}\mathcal{O}(x+iy) \tag{3.2.11}$$

Performing a partial integration shows that the integral vanish. However, we implicitly assume that the boundary terms coming from the y-integral vanish, which might not always be the case. This could be the main contribution to the breakdown of the CLM. As a result, one of the criteria of correct convergence is to check if the boundary terms vanish or not.

3.2.2 Distribution of drift terms

In Ref. [26, 47], an alternative but related argument for correct convergence was proposed. Recall that the expectation value at Langevin time t is given by

$$\langle\mathcal{O}\rangle_{P(t)} = \int dxdyP(x,y;t)\mathcal{O}(x+iy). \tag{3.2.12}$$

With a time lag τ, we can formally perform an infinite series expansion in τ as

$$\langle\mathcal{O}\rangle_{P(t+\tau)} = \sum_{n=0}^{\infty}\frac{\tau^n}{n!}\int dxdyP(x,y;t)\tilde{L}^n\mathcal{O}(x+iy) \tag{3.2.13}$$

with

$$\tilde{L} = \left(\frac{\partial}{\partial z} - \frac{\partial S}{\partial z}\right)\frac{\partial}{\partial z}. \tag{3.2.14}$$

This formal expansion will be valid at finite τ under two conditions. Naturally, it is required that each of the integrals in the sum must converge. Secondly, the convergence radius for the τ-expansion must be finite. The second concern can be shown to be related to the distribution of the magnitude of the drift terms, which may be defined as

$$u(z) = \max_{1\le i\le N}\left|\frac{\partial S}{\partial z_i}\right|. \tag{3.2.15}$$

In fact, the most dominant part each integral in (3.2.13) involves

$$\int dxdyu^n(x)P(x,y;t) = \int_0^{\infty}duu^n p(u;t) \tag{3.2.16}$$

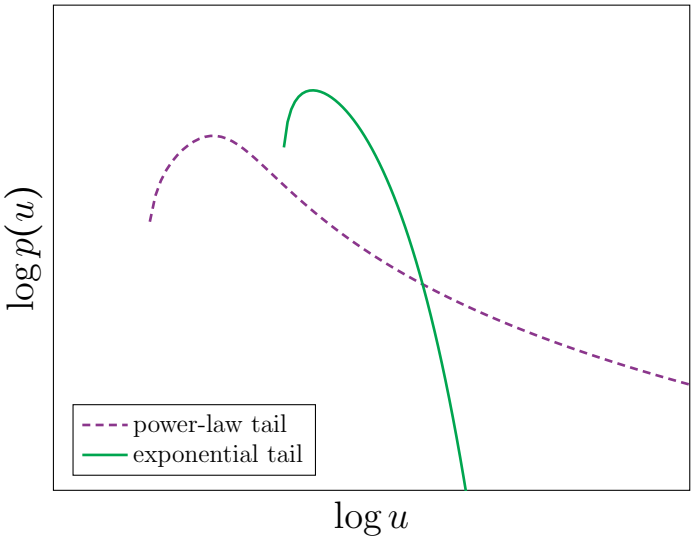

Figure 3.1: Qualitative comparison between histograms with power-law tail (dashed, purple) and exponential tail (solid, green).

where the distribution of the magnitude $p(u;t)$ is given by

$$p(u;t) \equiv \int dxdy\delta(u(z)-u)P(x,y;t) \tag{3.2.17}$$

If $p(u;t)$ falls off like a power law at large u, then (3.2.16) will diverge at large n, invalidating the τ-expansion (3.2.13). In other words, the simulation is reliable if $p(u;t)$ is exponentially suppressed or faster with u.

In practice, for each Langevin step, we record $u(z)$, from which we make a histogram. The simulation is reliable only if the tail decays exponentially or faster. If the tail is power-law-like, the result has to be discarded. An example of such a histogram is shown in Fig. 3.1.

3.3 Implementation to the 2D U(1) gauge theory

In case of the lattice gauge theory, the complex Langevin simulation amounts to updating the link variables with

$$\mathcal{U}_{n,\mu} \to \mathcal{U}_{n,\mu} \exp\left[i\left\{-\epsilon D_{n,\mu}S + \sqrt{\epsilon}\eta_{n,\mu}\right\}\right]. \tag{3.3.1}$$

Notice that we use the symbol $\mathcal{U}$, the complexified link variables, which belong to $\mathbb{C}\setminus\{0\}$, to distinguish it from the original link variables U, which belong to U(1). The drift term is defined by

$$D_{n,\mu}S = \lim_{\phi\to 0}\frac{S(e^{i\phi}U_{n,\mu}) - S(U_{n,\mu})}{\phi} \tag{3.3.2}$$

The gauge part of the drift term can be straightforwardly derived:

$$D_{n,1}S_g = -i\frac{\beta}{2}(P_n - P_n^{-1} - P_{n-\hat{2}} + P_{n-\hat{2}}^{-1}), \tag{3.3.3}$$

$$D_{n,2}S_g = +i\frac{\beta}{2}(P_n - P_n^{-1} - P_{n-\hat{1}} + P_{n-\hat{1}}^{-1}). \tag{3.3.4}$$

As for the θ term, one can already encounter the problem with the log definition. Since the topological charge is always an integer, its derivative will be zero all the time except at the instant of topology change, making the drift term singular at that point. Such a delta-like function is also difficult to be defined as a holomorphic function. It is, therefore, more informative to use the sine definition instead, whose drift term is well-behaved and easily obtained:

$$D_{n,1}S_\theta = -i\frac{\theta}{4\pi}(P_n + P_n^{-1} - P_{n-\hat{2}} - P_{n-\hat{2}}^{-1}), \tag{3.3.5}$$

$$D_{n,2}S_\theta = +i\frac{\theta}{4\pi}(P_n + P_n^{-1} - P_{n-\hat{1}} - P_{n-\hat{1}}^{-1}). \tag{3.3.6}$$

3.3.1 Gauge cooling

In order to stabilize the simulation, at every Langevin step, we perform a gauge transformation to minimize non-unitarity of the configuration [48]. This process can be done without affecting the argument for justifying the CLM. To describe the process, let us first define the unitarity norm

$$\mathcal{N} = \frac{1}{2L^2}\sum_{n,\mu}\left\{|\mathcal{U}_{n,\mu}|^2 + |\mathcal{U}_{n,\mu}|^{-2} - 2\right\} \tag{3.3.7}$$

which identically vanishes when all the link variables are unitary.

Next, the infinitesimal gauge transformation on a link variable is given by

$$\delta\mathcal{U}_{n,\mu} = (\epsilon_n - \epsilon_{n+\hat{\mu}})\mathcal{U}_{n,\mu} \tag{3.3.8}$$

for some real field ϵ_n. The change of the unitarity norm according to such transformation is

$$\begin{aligned}\delta\mathcal{N} &= \frac{1}{2L^2}\sum_{n,\mu}\left\{2(\epsilon_n - \epsilon_{n+\hat{\mu}})|\mathcal{U}_{n,\mu}|^2 - 2(\epsilon_n - \epsilon_{n+\hat{\mu}})|\mathcal{U}_{n,\mu}|^{-2}\right\} \\ &= \frac{1}{2L^2}\sum_n 2\epsilon_n G_n,\end{aligned} \tag{3.3.9}$$

where G_n is defined as

$$G_n = \sum_\mu\left\{|\mathcal{U}_{n,\mu}|^2 - |\mathcal{U}_{n,\mu}|^{-2} - |\mathcal{U}_{n-\hat{\mu},\mu}|^2 + |\mathcal{U}_{n-\hat{\mu},\mu}|^{-2}\right\}. \tag{3.3.10}$$

As a result, it is sensible to choose the field ϵ_n to be positively proportional to $-G_n$ to maximize the change of the unitarity norm. Specifically, we propose the gauge cooling procedure

$$\mathcal{U}_{n,\mu} \to g_n\mathcal{U}_{n,\mu}g_{n+\hat{\mu}}^{-1}; \qquad g_n = \exp(-\alpha G_n) \tag{3.3.11}$$

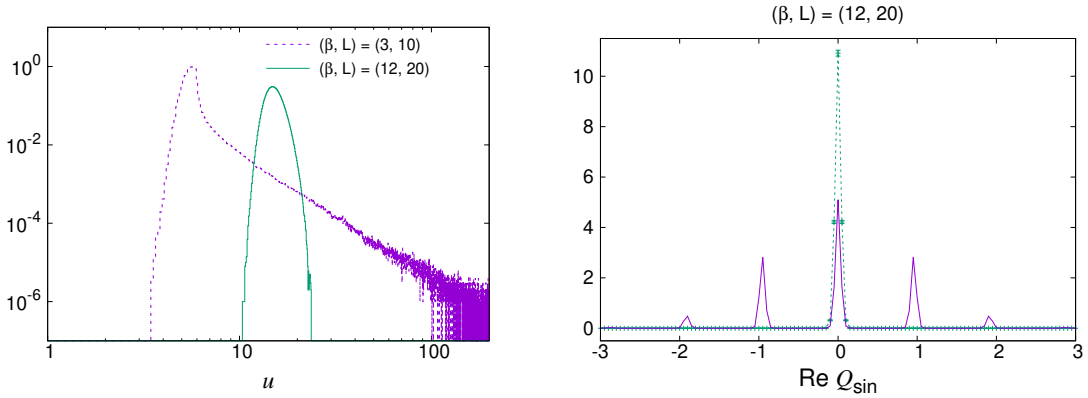

Figure 3.2: (Left) histogram of the drift terms with $(\beta, L) = (3, 10)$ and $(12, 20)$ with $\theta = \pi$ and (Right) histogram of ReQ with $(\beta, L) = (12, 20)$ and $\theta = \pi$. The exact result obtained with $\theta = 0$ is also shown by the solid line for comparison.

with some positive parameter α.

The optimal value of α can be estimated by first writing the unitarity norm as a function of α:

$$\mathcal{N}(\alpha) = \frac{1}{2L^2} \sum_{n,\mu} \left\{ |\mathcal{U}_{n,\mu}|^2 e^{-2\alpha(G_n - G_{n+\hat{\mu}})} + |\mathcal{U}_{n,\mu}|^{-2} e^{+2\alpha(G_n - G_{n+\hat{\mu}})} - 2 \right\}. \tag{3.3.12}$$

Expanding this function up to the second order in α gives us the estimated minimum of

$$\alpha = \frac{1}{2} \left(\sum_n G_n^2 \right) \left(\sum_{n,\mu} (G_n - G_{n+\hat{\mu}})^2 (|\mathcal{U}_{n,\mu}|^2 + |\mathcal{U}_{n,\mu}|^{-2}) \right)^{-1}. \tag{3.3.13}$$

We repeat this process until the unitarity norm converges with $\delta\mathcal{N}/\mathcal{N} < 10^{-5}$.

3.4 Topology freezing and the large-drift problem

Unfortunately, we found that naive CLM simulations will never work due to two different problems that are unexpectedly connected in some way. To see this, let us look at some results. We perform simulations at various θ at $(\beta, L) = (3, 10)$ 'the small β', and $(\beta, L) = (12, 20)$ 'the large β'. These two sets of parameters are actually corresponding to the same physical volume $V_{\rm phys} = L^2/\beta = 10^2/3$. In the result below, we will discuss only $\theta = \pi$ since it suffers from the sign problem the most. The situation at smaller θ is qualitatively the same as the case at $\theta = \pi$.

At small β, we find that the histogram of the drift term falls off with a power-law tail. This signifies that the simulation is unreliable at this parameter region. This is shown as the purple-dashed line of Fig. 3.2 (left). And as we approach the continuum limit with large β, the criterion for correct convergent is satisfied as expected.

However, another problem arises in place of the large-drift problem at large β. If we

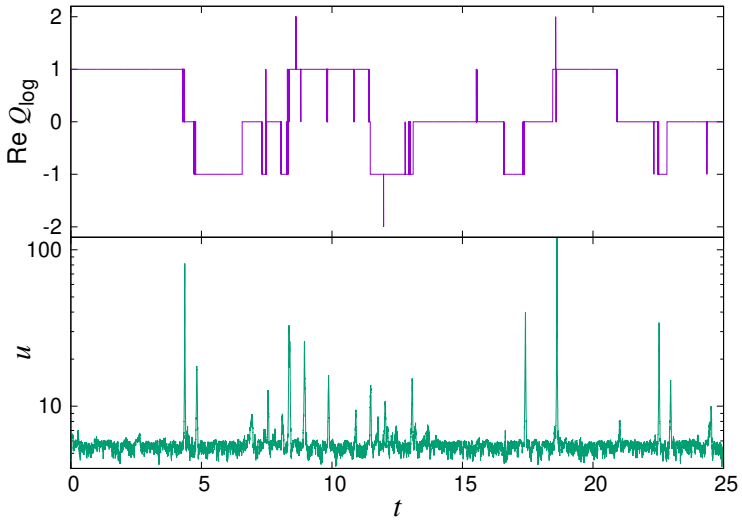

Figure 3.3: The upper plot show the history of the log histogram of the topological charge while the lower plot show the history of the magnitude of the drift term in log scale. The simulation is done with $(\beta, L) = (3, 10)$ and $\theta = \pi$.

look at the histogram of the real part of the topological charge[2], we can see that its value is critically frozen at zero, contrasting with the small-β case where the charge fluctuates in the region of $|\mathrm{Re}Q| \lesssim 2$. This is shown in Fig. 3.2 (right).

As already preluded, the large-drift problem and the topology freezing problem are in fact a trade-off of each other. Namely, one cannot find a parameter region where both problems simultaneously disappear. To see this, note that a strong correlation between the topology change (log definition) and the large drift terms can be observed as shown in Fig. 3.3. When the topological charge jumps, a peak in the drift term appears.

To understand this better, we consider a drift term from one particular link variable, e.g. $\mathcal{U}_{n,1}$. The adjacent plaquette to this link variables are $P(n)$ and $P(n-\hat{2})$. For simplicity of the analysis, we set $P_{n-\hat{2}} = 1$ and consider the drift term v as a function of $P_n = \exp(i\phi)$:

$$v(\phi) = \beta \sin\phi - i\frac{\theta}{2\pi}(\cos\phi - 1). \tag{3.4.1}$$

Let $\beta = \theta = 1$. The flow diagram of ϕ according to the drift term at different complex ϕ is shown in Fig. 3.4 (left). When $\mathrm{Re}\phi$ is close to $\pm\pi$, namely when the topology change happens, the phase ϕ tends to flow further away from the real line. But as shown in Fig. 3.4 (right), the magnitude of the drift term increases exponentially with $|\mathrm{Im}\phi|$. In other words, the drift term tends to diverge if the topological transition happens.

[2]The real part of the topological charge is not a physically sensible observable, but it is a good indicator for the topology freezing problem.

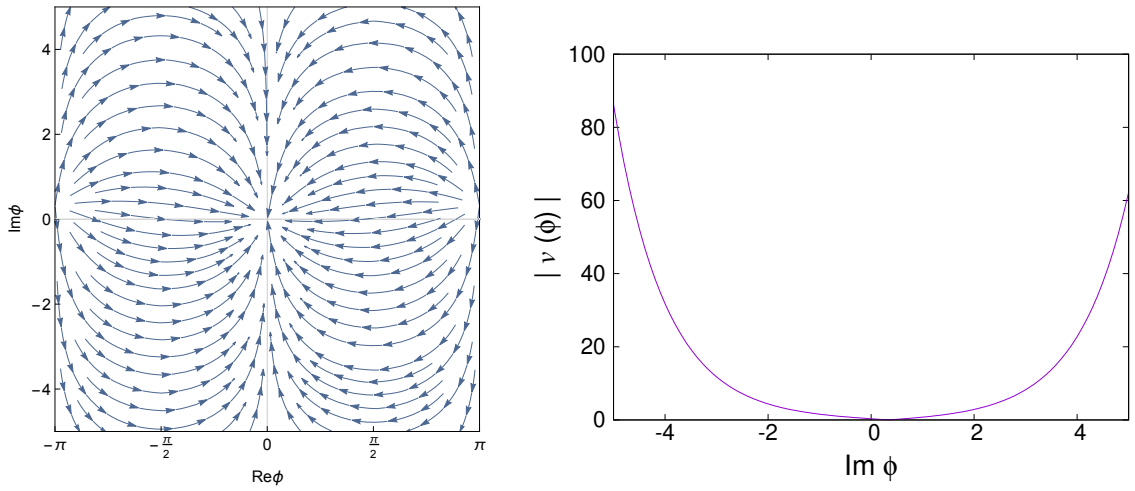

Figure 3.4: (Left) flow diagram of the drift term at different complex values of ϕ and (Right) the magnitude of the drift term as a function of Imϕ with Re$\phi = \pi$.

Introducing a topological defect

Since there are links between lattice topology and drift term's behavior, we consider introducing a topological defect to the lattice. Specifically, we introduce an open boundary around a specific plaquette, which we will call a 'puncture'. Since the configuration is no longer restricted by the close-manifold topology, the topology freezing is expected to be solved. To implement this, we remove one of the plaquettes, denoted as P_K from the action. Consequently, the drift term for the log definition can also be made sense in the punctured model due to the non-compactness of the manifold.

Specifically, to implement the puncture, we only need to modify the drift terms around the puncture to compensate for the removal of the plaquette P_K. In the followings, we will only discuss the log definition as the sine definition has similar qualitative behavior. The drift term in the presence of the puncture becomes

$$D_{n,1}S = \begin{cases} -i\frac{\beta}{2}(P_n - P_n^{-1}) - i\frac{\theta}{2\pi} & ;\ n = K + \hat{2}, \\ +i\frac{\beta}{2}(P_{n-\hat{2}} - P_{n-\hat{2}}^{-1}) + i\frac{\theta}{2\pi} & ;\ n = K, \\ -i\frac{\beta}{2}(P_n - P_n^{-1} - P_{n-\hat{2}} + P_{n-\hat{2}}^{-1}) & ;\ \text{else}, \end{cases} \tag{3.4.2}$$

$$D_{n,2}S = \begin{cases} -i\frac{\beta}{2}(P_n - P_n^{-1}) + i\frac{\theta}{2\pi} & ;\ n = K + \hat{1}, \\ +i\frac{\beta}{2}(P_{n-\hat{1}} - P_{n-\hat{1}}^{-1}) - i\frac{\theta}{2\pi} & ;\ n = K, \\ -i\frac{\beta}{2}(P_n - P_n^{-1} - P_{n-\hat{1}} + P_{n-\hat{1}}^{-1}) & ;\ \text{else}. \end{cases} \tag{3.4.3}$$

Using the same parameters as in the previous simulations but with the introduction of the puncture, the situation is greatly improved. In the small β region, the CLM is still prone to break down because the plaquettes throughout the lattice tend to oscillate more than those in the large β region. Such oscillations make it more likely that a plaquette will cross the branch cut at $\phi = \pm\pi$ (see the final part of section 3.4), making the drift terms large. But in the large β region, the fluctuation is greatly suppressed, and the tail of the

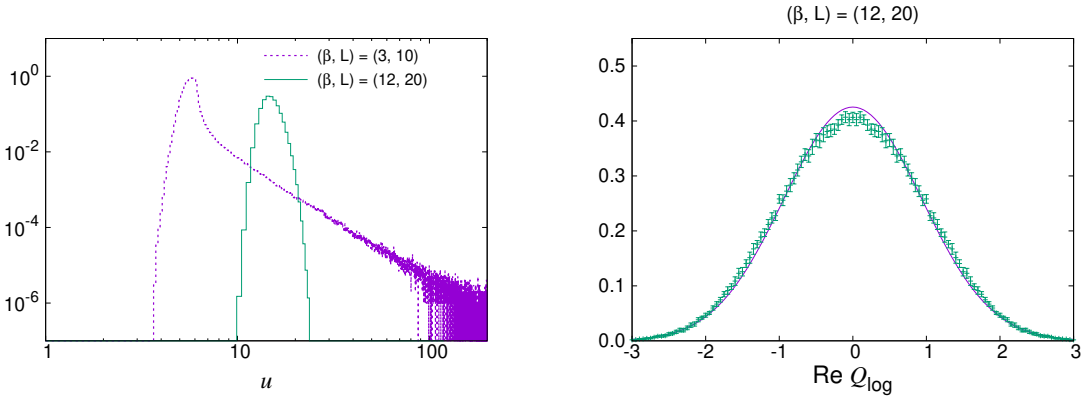

Figure 3.5: (Left) histogram of the drift terms with $(\beta, L) = (3, 10)$ and $(12, 20)$ with $\theta = \pi$ after introducing a puncture and (Right) histogram of ReQ with $(\beta, L) = (12, 20)$ and $\theta = \pi$, also with a puncture. The exact result obtained with $\theta = 0$ is also shown by the solid line for comparison.

histogram falls exponentially. This is shown in Fig. 3.5 (left).

As for the topological charge, because of the puncture, they are no longer integers. Therefore, the charge can fluctuate freely as a real number, even at large β. This is show in Fig. 3.5 (right), where the charge fluctuates between -3 and $+3$.

Observables at finite θ

We will consider three observables: average plaquette, topological charge density, and topological susceptibility, defined by

$$w = \frac{1}{V}\frac{\partial}{\partial\beta}\log Z = \frac{1}{2V}\sum_{n\neq K}\langle P_n + P_n^{-1}\rangle_{\mathrm{CLM}}, \tag{3.4.4}$$

$$\frac{\langle Q\rangle}{V} = \frac{1}{iV}\frac{\partial}{\partial\theta}\log Z = \frac{1}{2\pi iV}\sum_{n\neq K}\langle \log P_n\rangle_{\mathrm{CLM}}, \tag{3.4.5}$$

$$\chi = -\frac{1}{V}\frac{\partial^2}{\partial\theta^2}\log Z = \frac{1}{V}\left(\langle Q^2\rangle_{\mathrm{CLM}} - \langle Q\rangle^2_{\mathrm{CLM}}\right), \tag{3.4.6}$$

where $\langle\,\cdot\,\rangle_{\mathrm{CLM}}$ stands for the expectation value obtained from the simulation. The results are shown in Fig. 3.6. As already discussed, the result slightly deviated from the exact values at small β because of the branch-crossing problem which is directly related to the large-drift problem. Interestingly, despite the fact that the average plaquette shows a clear deviation from the exact result, the topological charge density and the topological susceptibility agree very well. On the other hand. the large-drift problem totally disappear for large β, where the numerical results agree with the exact value very well. All of this can be understood from our analysis on drift terms previously.

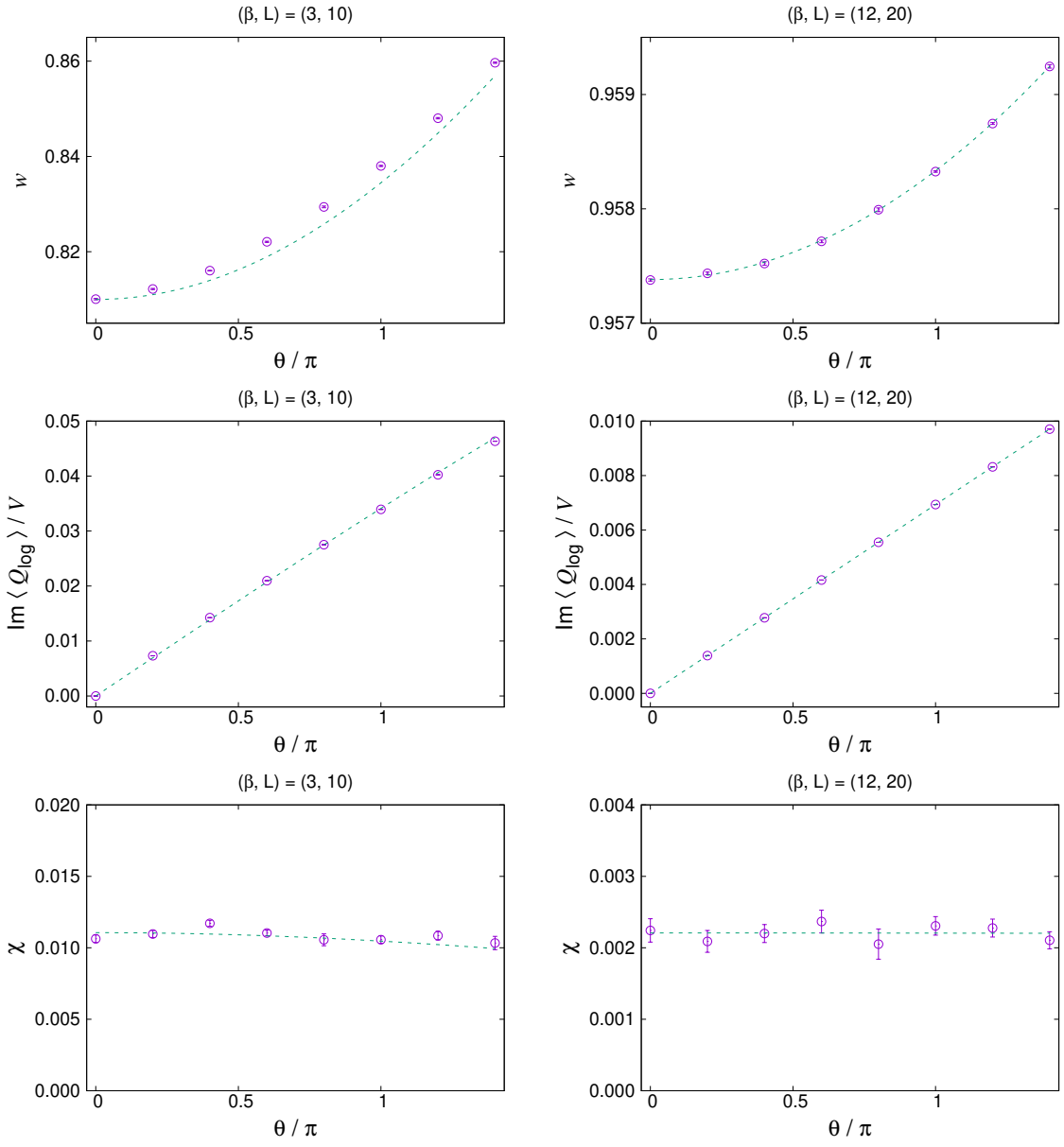

Figure 3.6: The three observables calculated in the puncture model with $(\beta, L) = (3, 10)$ (left column) and $(\beta, L) = (12, 20)$ (right column). The exact results for each plot are also shown in dashed line for comparison.

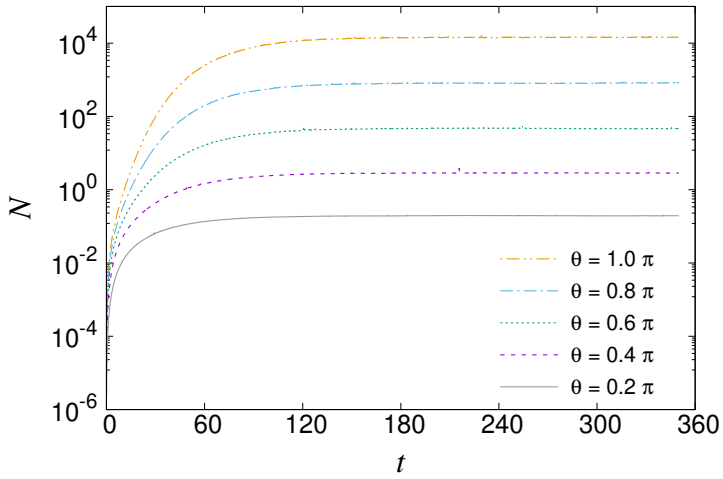

Figure 3.7: The history of unitarity norm in case of the punctured model with various θ and $(\beta, L) = (5, 16)$.

3.5 Remarks on unitarity norm

Note that the drift terms associated with the θ term (see (3.4.2) and (3.4.3)) directly modify the unitarity of the link variables surrounding the puncture. Specifically, at each Langevin step, two of the links are multiplied by $\exp(\theta\epsilon/2\pi)$, and the other two are multiplied by $\exp(-\theta\epsilon/2\pi)$. Furthermore, such a non-unitary factor can also propagate outward into the interior of the lattice, making the unitarity norm as a whole increase significantly. It is thus a valid concern that this unitarity issue will grow exponentially with the Langevin time.

On the other hand, our simulations suggest that the growth of the unitarity norm, even though initially grows exponentially, will eventually saturate at a finite value at some point, as shown in Fig. 3.7. On that note, it is important to wait for unitarity norm saturation before we can use the configurations to calculate observables.

In fact, we find that the non-unitarity is distributed not uniformly across the lattice, but actually is localized near the puncture. To see this, we define the 'local unitarity norm' to be the measure of non-unitarity of the link variables surrounding a plaquette at a certain lattice site n:

$$\mathcal{N}(n) = \frac{1}{4} \sum_{(k,\mu)\in P_n} \left\{ |\mathcal{U}_{k,\mu}|^2 + |\mathcal{U}_{k,\mu}|^{-2} - 2 \right\}. \tag{3.5.1}$$

The global unitarity norm (3.3.7) is actually an average of the local unitarity norm across the lattice, including the removed one; namely $\mathcal{N} = \frac{1}{L^2} \sum_n \mathcal{N}(n)$. Fig. 3.8 shows $\mathcal{N}(n)$ across the lattice.

Despite this large unitarity norm, the observables agree with the analytical prediction very well. In fact, what actually matters for the validity of the CLM is not so much concerning the unitarity of the link variables, but actually more about the unitarity of the plaquettes, which are what contribute most of the important observables. If we plot the

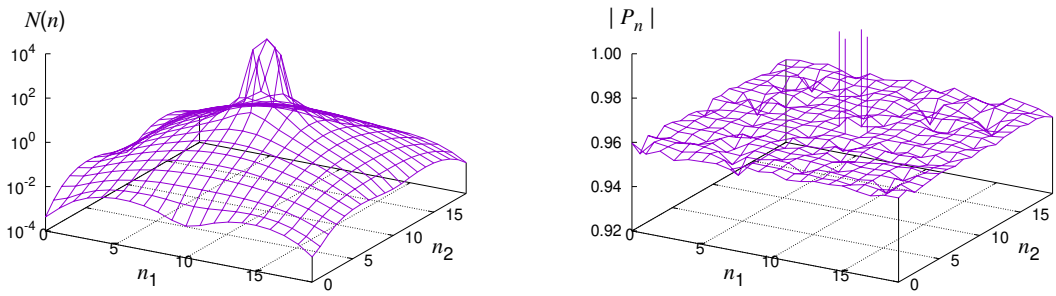

Figure 3.8: (Left) the local unitarity norm (3.5.1) plotted against the lattice site $n = (n_1, n_2)$ with the puncture located at the center $K = (10, 10)$. (Right) Absolute value of the plaquette of the same lattice. The parameters for both plots are $(\beta, L) = (12, 20)$ and $\theta = \pi$.

absolute value of the plaquette in a similar fashion to the local unitarity norm, we find that all of the plaquettes share the non-unitarity, which is surprisingly small compared to $\mathcal{N}(n)$, uniformly. This is shown in Fig. 3.8 (right). Actually, the plaquette at the puncture P_K is the one that is exponentially deviated from unitarity, but it is inconsequential since P_K is included in neither the action nor the observables.

3.6 Brief summary

In this chapter, I discuss the application of the complex Langevin method to the two-dimensional U(1) gauge theory with a θ term. In section 3.1, I give an overview of how the CLM is formulated. In section 3.2, I explain the justification condition of the method and explain the situation where the CLM breaks down. This includes the consideration of the so-called boundary term, whose existence can be shown to spoil the equivalence between the complex Langevin's equilibrium and the original theory's equilibrium. On the other hand, one can also determine if the CLM converges correctly or not by looking at the distribution of the drift term. I then proceed to explain how to implement the CLM for the 2D U(1) gauge theory with a θ term in section 3.3. In section 3.4, I present the naive implementation of the CLM and show that the simulation suffers from the topology freezing problem for large β and also the large-drift problem for small β. We show that the two problems are actually a trade-off. We then introduce a topological defect on the lattice, which turns out to solve both of the problems in the continuum limit. And despite the fact that I altered the topological structure of the lattice, the expectation value of all variables agrees with the exact result very well in the continuum limit up to $\theta = \pi$. In section 3.5, I explain a possible concern about non-unitarity of link variables. This however is not a problem as I show that the divergent-looking behavior of link variable non-unitarity all cancels out in gauge-invariant observables.

Chapter 4

Generalized Lefschetz thimble method

4.1 Picard-Lefschetz theory

The goal of the Picard-Lefschetz theory is to study the topology of a complex manifold $\mathcal{M}$ by looking at a set of saddle points of a holomorphic function on $\mathcal{M}$ [55, 56]. One of its applications is that it helps rewrite multidimensional integrals with complex weight in a way that the Monte Carlo simulation suffers least of the sign problem [57, 58].

Specifically, consider the deformation of the integration contour

$$\int_{\mathbb{R}^N} d^N x e^{-S(x)} \to \int_{\mathcal{M}} d^N z e^{-S(z)}. \tag{4.1.1}$$

The deformation is defined via the so-called anti-holomorphic gradient flow equation

$$\frac{dz}{d\tau} = \overline{\frac{\partial S}{\partial z}} \tag{4.1.2}$$

with the initial condition $z(0) = x$. Because of the analyticity of the flow, the two integrals are equivalent by virtue of Cauchy's theorem. The Picard-Lefschetz theory states that the deformed manifold $\mathcal{M}$ will be decomposed into many submanifolds $\mathcal{J}^{(\sigma)}$, each associated with different saddle points $z_\star^{(\sigma)}$. These submanifolds are known as the Lefschetz thimbles. And due to the fact that the flow equation preserves the imaginary part of $S(z)$ along the flow:

$$\frac{dS}{d\tau} = \frac{\partial S}{\partial z}\frac{dz}{d\tau} = \left|\frac{\partial S}{\partial z}\right|^2 > 0, \tag{4.1.3}$$

one can shows that in the infinite flow time limit, configurations on each thimble all produces the same imaginary part of the action, say $\mathrm{Im}S(z) = \phi_\sigma$ for $z \in \mathcal{J}^{(\sigma)}$. Consequently, one can write

$$\int_{\mathcal{M}} d^N z e^{-S(z)} = \sum_\sigma \mathfrak{n}_\sigma e^{-i\phi_\sigma} \int_{\mathcal{J}^{(\sigma)}} d^N z e^{-\mathrm{Re}S(z)}. \tag{4.1.4}$$

Theoretically, the intersection number $\mathfrak{n}_\sigma \in \mathbb{Z}$ can be computed via Morse theory, but the computation can be nontrivial in practical situations.

4.2 Technical developments

In this section, we will describe the method in a more practical way, as well as the developments of numerical techniques used in the thimble simulations. In the followings, the term generalized Lefschetz thimble (GLT) method will be referred to a class of algorithms where one makes a deformation of the original contour with finite flow time; i.e. $\mathbb{R}^N \to \mathcal{M}_\tau \subset \mathbb{C}^N$,

$$\int_{\mathbb{R}^N} d^N x e^{-S(x)} \to \int_{\mathcal{M}_\tau} d^N z e^{-S(z)}. \tag{4.2.1}$$

In the original generalized Lefschetz thimble method, the deformed manifold is parametrized by the 'flow time' τ, denoted $\mathcal{M}_\tau$, on which the configuration $z(\tau)$ is solved from the flow equation

$$\frac{dz_a}{d\sigma} = \overline{\frac{\partial S(z)}{\partial z_a}} \tag{4.2.2}$$

from $\sigma = 0$ to $\sigma = \tau$ with the initial condition $z(0) = x \in \mathbb{R}^N$. The corresponding Jacobian $J_{ab} = \partial z_a / \partial x_b$ can be evaluated by taking derivative of the flow equation with respect to x_b:

$$\frac{\partial J_{ab}}{\partial \sigma} = \overline{H_{ac} J_{cb}} \tag{4.2.3}$$

where $H_{ab} \coloneqq \partial_a \partial_b S$ is the Hessian, and the initial condition for J is $J(0) = \mathbb{1}$.

In practice, the flow equations for both z and J need to be discretized as

$$z(\sigma + \epsilon) = z(\sigma) + \epsilon \overline{\partial S}(\sigma), \tag{4.2.4}$$

$$J(\sigma + \epsilon) = J(\sigma) + \epsilon \overline{H}(\sigma) \overline{J}(\sigma), \tag{4.2.5}$$

with the same initial condition as in the continuum case. Note that the finiteness of ϵ has no effect on the physics. However, if ϵ is too large, the sign problem will not be resolved at large τ.

As already stated, due to (4.1.3), the real part of the action always increases with the flow, while the imaginary part remains constant. Because of this, the Boltzmann weight $\exp(-S(z))$ far away from the fixed points z^* (the points where the derivative of the action vanishes) will be exponentially suppressed. As a result, the configurations obtained from the Monte Carlo simulation will be dominantly around the fixed points, where the phase of the reweighting factor is not violently fluctuating, essentially solving the sign problem.

4.2.1 HMC algorithm on the original manifold

There are two ways to implement the HMC algorithm for thimble simulation. The traditional method is to implement the Hamilton dynamics on the original integration domain. Specifically, one first define the HMC expectation value

$$\langle \mathcal{O}(z) \rangle_{\mathrm{HMC}} = \frac{\int d^N x d^N p \mathcal{O}(z(x)) e^{-\mathrm{Re} S(z(x)) - p^2/2}}{\int d^N x d^N p e^{-\mathrm{Re} S(z(x)) - p^2/2}}. \tag{4.2.6}$$

Using this, the expectation value of interest is given by

$$\langle \mathcal{O}(z) \rangle = \frac{\int d^N z \mathcal{O}(z) e^{-S(z)}}{\int d^N z e^{-S(z)}} = \frac{\langle \mathcal{O}(z) \det J e^{-i \mathrm{Im} S(z)} \rangle_{\mathrm{HMC}}}{\langle \det J e^{-i \mathrm{Im} S(z)} \rangle_{\mathrm{HMC}}} \tag{4.2.7}$$

where $\langle \cdot \rangle_{\mathrm{HMC}}$ is the expectation value obtained from the simulation and $J_{ab} = \partial z_a / \partial x_b$ is the Jacobian of the map $z : \mathbb{R}^N \to \mathcal{M}_\tau$. One may notice that this is essentially the reweighting formula (1.3.6). However, due to the fact that the only dominant configurations are those around the fixed points, the expectation value $\langle \det J e^{-i \mathrm{Im} S(z)} \rangle_{\mathrm{HMC}}$ no longer cancels out to be a small number, thus avoiding the sign problem.

In this formulation, the canonical pair is (x, p) with the Hamiltonian

$$H(x, p) = \frac{p^2}{2} + \mathrm{Re} S(z(x)). \tag{4.2.8}$$

The force exerted on the variable x_a is then given by

$$F_a = -\frac{\partial}{\partial x_a} \mathrm{Re} S(z(x)) = -\mathrm{Re}\left(\frac{\partial S}{\partial z_b} \frac{\partial z_b}{\partial x_a} \right) = -\mathrm{Re}(\partial_b S \cdot J_{ab}). \tag{4.2.9}$$

One apparent obstacle to this formulation is that, in order to compute the force, it is necessary to calculate the Jacobian, which is known to have the computational cost of order $O(N^3)$, or $O(N^2)$ if the Hessian is sparse. Compared to the measly cost of $O(N)$ in the case of the complex Langevin, the traditional Lefschetz simulation was seen as less favorable when studying large systems.

Apart from the problem of high computational cost, the thimble technique is also known to suffer from the ergodicity problem, which occurs when the system has many fixed points. At large flow time, the potential barrier separating the region around two nearby fixed points becomes exponentially high, making the probability of the configuration crossing the regions exponentially suppressed. This means that the sampling will be mostly trapped near one of the fixed points. The earlier solution was to find the intermediate flow time where both the ergodicity problem and the sign problem are not prominent, but we cannot guarantee that such a region always exists, especially as we increase the system size.

An alternative solution to the ergodicity problem is to introduce several replicas of generalized thimbles $\mathcal{M}_\tau$, each with a different flow time [118–120]. Each replica will undergo its own HMC dynamics. And every so often, a swap operation is done between two replicas if it is accepted by the Metropolis-Hastings algorithm. This operation is known as parallel tempering, which is a useful method for solving the ergodicity problem in general. The acceptance rate is however depending on the number of replicas and how close they are together. As the number of replicas increases, not only does it takes up more computational time, it also takes up more memory. A better approach is thus highly desired.

4.2.2 HMC algorithm on the deformed manifold

A more sophisticated version of the tempering approach was recently proposed [121] to solve both the high computation cost and the ergodicity problems simultaneously. The idea is

to implement the Hamiltonian dynamics on the deformed manifold instead of the original domain. One can also naturally extend the dynamical space from $\mathcal{M}_\tau$ to $\mathcal{R} = \mathcal{M}_\tau \times \{\tau\}$, allowing the configuration to travel to different 'timeslices'. The technique was aptly named the worldvolume approach to the tempered Lefschetz thimble method (WV-TLTM).

In this formulation, the canonical pair is (z, π) where $z_I = (\mathrm{Re}z_a, \mathrm{Im}z_a)$ and π_I is the conjugated momentum to ξ_I. The Hamiltonian is

$$H(z,\pi) = \frac{\pi^2}{2} + \mathrm{Re}S(z). \tag{4.2.10}$$

The force exerted on z_I can then be straightforwardly computed without the involvement of the Jacobian:

$$F_I = \left(-\frac{\partial \mathrm{Re}S}{\partial \mathrm{Re}z_a}, -\frac{\partial \mathrm{Re}S}{\partial \mathrm{Im}z_a}\right). \tag{4.2.11}$$

There is one complication that needs to be addressed: the usual leapfrog integrator needs to be modified in order for the configuration to be confined onto the manifold $\mathcal{R}$ at all time. And that is to add an appropriate amount of normal forces to the Hamilton's equations:

$$\begin{aligned} \pi(s+\Delta s/2) &= \pi(s) + \frac{\Delta s}{2} F(z(s)) - \lambda \mathcal{N}(z(s)), \\ z(s+\Delta s) &= z(s) + \Delta s \pi(s+\Delta s/2), \\ \pi(s+\Delta s) &= \pi(s+\Delta s/2) + \frac{\Delta s}{2} F(z(s+\Delta s)) - \tilde{\lambda} \mathcal{N}(z(s+\Delta s)). \end{aligned} \tag{4.2.12}$$

The computation of the normal forces $\mathcal{N}$ and the Lagrange multipliers λ, $\tilde{\lambda}$ are considerably complicated and will not be fully expressed here, but the point is they can be computed with the cost of $O(N^2)$ or $O(N)$ for sparse Hessian. Another advantage of this approach is the fact that the so-called overlap problem, which happens when the modulus of the reweighting factor and the observables may not have significant overlap, can be avoided entirely since the modulus of the Jacobian is already included in the HMC dynamics.

Because the thimble approach doesn't have the wrong convergence problem and can be done with the same complexity as the CLM, this development made the thimble simulation a very attractive solution to the sign problem. Although it should be noted that the residual sign problem from the phase of the Jacobian can still cause a problem at large system sizes, which is not a problem for the CLM.

4.2.3 Backpropagating HMC

In our recent work [122], we noticed that it is possible to overcome the high-cost problem even if we implement the HMC algorithm on the original manifold. Such algorithm can also be implemented with the worldvolume technique (see next section). An advantage of this approach over the implementation on the deformed manifold is its simplicity since there is no need to perform a complicated procedure of manifold projection during the leapfrog process.

To understand the improvement of the computational cost, let us rewrite the force in the HMC process as

$$\begin{aligned} F_a &= -\frac{\partial z_b(0)}{\partial x_a}\frac{\partial}{\partial z_b(0)}\mathrm{Re}S(z(\tau)) - \frac{\partial \bar{z}_b(0)}{\partial x_a}\frac{\partial}{\partial \bar{z}_b(0)}\mathrm{Re}S(z(\tau)) \\ &= -\frac{\partial}{\partial z_b(0)}\mathrm{Re}S(z(\tau))J_{ba}(0) + \text{c.c.} \\ &= -\frac{\partial}{\partial z_a(0)}\mathrm{Re}S(z(\tau)) + \text{c.c.}, \end{aligned} \tag{4.2.13}$$

where we have used the fact that $J(0) = \mathbb{1}$. It is convenient to define a quantity

$$f_a(\sigma) = -\frac{\partial}{\partial z_a(\sigma)}\mathrm{Re}S(z(\tau)), \tag{4.2.14}$$

so we have $F = f(0) + \bar{f}(0)$. Using chain rule, we can write a recurrence relation

$$\begin{aligned} f_a(\sigma-\epsilon) &= -\frac{\partial z_b(\sigma)}{\partial z_a(\sigma-\epsilon)}\frac{\partial}{\partial z_b(\sigma)}\mathrm{Re}S(z(\tau)) - \frac{\partial \bar{z}_b(\sigma)}{\partial z_a(\sigma-\epsilon)}\frac{\partial}{\partial \bar{z}_b(\sigma)}\mathrm{Re}S(z(\tau)) \\ &= -\delta_{ab}\frac{\partial}{\partial z_b(\sigma)}\mathrm{Re}S(z(\tau)) - \epsilon H_{ab}(\sigma-\epsilon)\frac{\partial}{\partial \bar{z}_b(\sigma)}\mathrm{Re}S(z(\tau)) \\ &= f_a(\sigma) + \epsilon H_{ab}(\sigma-\epsilon)\bar{f}_b(\sigma). \end{aligned} \tag{4.2.15}$$

In the derivation above, we have used the following identities:

$$\frac{\partial z_a(\sigma+\epsilon)}{\partial z_b(\sigma)} = \delta_{ab}, \tag{4.2.16}$$

$$\frac{\partial \bar{z}_a(\sigma+\epsilon)}{\partial z_b(\sigma)} = \epsilon H_{ab}(\sigma), \tag{4.2.17}$$

which can be easily derived from the discretized flow equation (4.2.4).

Using (4.2.15), one can compute $f(0)$ iteratively 'backward' from $f(\tau)$, and thus the name backpropagation[1]. The cost of computation in each step of (4.2.15) is of order $O(N^2)$, or $O(N)$ for sparse Hessian. Therefore, this effectively reduces the total cost of the HMC algorithm from the usual calculation by the order of $O(N)$.

Alternatively, one can understand backpropagation by looking at the flow equation of the Jacobian (4.2.5), which can be rewritten as

$$\begin{pmatrix} J(\sigma+\epsilon) \\ \bar{J}(\sigma+\epsilon) \end{pmatrix} = \underbrace{\begin{pmatrix} \mathbb{1} & \epsilon\bar{H}(\sigma) \\ \epsilon H(\sigma) & \mathbb{1} \end{pmatrix}}_{\equiv\,\mathcal{U}(\sigma)} \begin{pmatrix} J(\sigma) \\ \bar{J}(\sigma) \end{pmatrix}. \tag{4.2.18}$$

[1]This process is similar to the technique of the same name in the field of machine learning, which is a heavy inspiration for our technique.

Using this recurrence relation, one can rewrite the force as

$$F = -\mathrm{Re}(\partial S(\tau)\cdot J(\tau)) = -\frac{1}{2}\begin{pmatrix}\partial S(\tau) & \overline{\partial S}(\tau)\end{pmatrix}\begin{pmatrix}J(\tau)\\ \bar{J}(\tau)\end{pmatrix}$$
$$= -\frac{1}{2}\begin{pmatrix}\partial S(\tau) & \overline{\partial S}(\tau)\end{pmatrix}\mathcal{U}(\tau-\epsilon)\mathcal{U}(\tau-2\epsilon)\cdots\mathcal{U}(2\epsilon)\mathcal{U}(\epsilon)\begin{pmatrix}\mathbb{1}\\ \mathbb{1}\end{pmatrix}. \tag{4.2.19}$$

If one were to evaluate this quantity by multiplying the matrices from the right (forward propagation), which is what is traditionally done, we cannot avoid the repeated operation of matrix-matrix multiplications, which is of order $O(N^3)$ in complexity (or $O(N^2)$ for sparse Hessian). By virtue that $\begin{pmatrix}\partial S(\tau) & \overline{\partial S}(\tau)\end{pmatrix}$ is a vector, multiplying the matrices from the left (backward propagation) reduces the cost of operation by an order of $O(N)$.

4.2.4 Integrating over the flow time

To avoid the ergodicity problem, we promote the parameter τ in $\mathcal{M}_\tau$ to be dynamical. By doing this, the configuration is then able to cross between 'sectors' through the small-τ region. And since such a region suffers from the sign problem, we only include the large-τ region in the calculation of observables. The general idea was known as the worldvolume approach [121] to thimble simulation and was first introduced with the HMC simulation on the deformed manifold. However, it is also possible to apply it to the HMC simulation on the original manifold, as we will explain in this section.

In terms of the integral, the new formulation amounts to the modification

$$\int_{\mathcal{M}_\tau} d^N z e^{-S(z)} \to \int_{\tau_{\min}}^{\tau_{\max}} d\tau \int_{\mathcal{M}_\tau} d^N z e^{-S(z)-W(\tau)} \tag{4.2.20}$$

where $W(\tau)$ is an additional 'potential' for τ that is introduced for technical benefit and is physically inconsequential. The role of the potential $W(\tau)$ is to make the configuration uniformly distributed in our region of interest. We will discuss how to choose this function below.

Because τ is now dynamical, we need to introduce the corresponding force, which is given by

$$F_\tau = -\frac{\partial}{\partial\tau}\mathrm{Re}S - \frac{d}{d\tau}W(\tau) = -\mathrm{Re}\left(\frac{\partial S(\tau)}{\partial z_a}\frac{\partial z_a}{\partial\tau}\right) - \frac{d}{d\tau}W(\tau). \tag{4.2.21}$$

In the continuum flow-time limit, we can simply invoke the flow equation $\partial z/\partial\tau = \overline{\partial S/\partial z}$ and get $F_\tau = |\partial S|^2 + \dot{W}$. However, this is incorrect for finite ϵ, where one needs to write the flow equation for $\dot{z} = \partial z/\partial\tau$ separately.

More specifically, we define

$$\dot{z}(\sigma) = \frac{\epsilon}{\sigma}\frac{\partial z(\sigma)}{\partial\epsilon}. \tag{4.2.22}$$

Using this, the flow equation for $\dot{z}$ is

$$\dot{z}(\sigma+\epsilon) = \frac{\epsilon}{\sigma+\epsilon}\overline{\partial S}(\sigma) + \frac{\sigma}{\sigma+\epsilon}(\mathbb{1}+\epsilon\bar{H}(\sigma))\dot{z}(\sigma). \tag{4.2.23}$$

Choosing the W-potential

To identify the auxiliary function $W(\tau)$, we first perform a preliminary simulation with some initial guess $W_0(\tau)$ and then make a histogram plot of τ. Suppose that $\rho(\tau)$ is the probability distribution of τ obtained from the histogram, we then update the W-potential with

$$W(\tau) = W_0(\tau) + \log \rho(\tau). \tag{4.2.24}$$

Keep repeating this process until the histogram is sufficiently flat.

HMC algorithm with variable mass

Unlike the implementation on the deformed manifold, naively implementing the HMC algorithm on the original manifold $\mathbb{R}^N \times \{\tau\}$ has some technical difficulty when the interval of τ is large. Due to the nature of the flow equation, perturbing the initial condition by δx gives rise to the perturbation of the final point δz, which grows exponentially with τ. This in itself is not a problem. However, this will critically slow down the simulation because the Hamiltonian conservation will become exponentially worse, i.e. the acceptance rate in the Metropolis-Hastings step will be almost 0%. One solution is to decrease the leapfrog parameter Δs until the acceptance rate at $\tau_{\max}$ is acceptable. But then the acceptance rate at $\tau_{\min}$ will become mostly 100%, which is in itself inefficient.

Our solution is to introduce a τ-dependent mass to the HMC Hamiltonian:

$$H = \frac{p_i^2}{2m(\tau)} + \frac{p_\tau^2}{2} + \mathrm{Re}S(z(x;\tau)) + W(\tau). \tag{4.2.25}$$

It is important that the mass for p_τ is kept to be 1. By making $m(\tau)$ large at large τ, we can finely tune the acceptance rate to be constant at all τ.

Specifically, the mass function can be estimated using

$$m(\tau) \sim |\det J|^{2/N} \tag{4.2.26}$$

with the reasoning that the perturbation δz is proportional to $|\det J|^{1/N}$, while the effective step size of the leapfrog process is proportional to $1/\sqrt{m(\tau)}$. Setting the product of the two quantities to be of order unity gives $|\det J|^{1/N}/\sqrt{m(\tau)} \sim 1$, and thus (4.2.26). And as will be seen below, using non-constant mass function will have a direct effect on the HMC dynamics, meaning that it will affect the distribution of τ. Therefore, it is important to choose the mass function before tuning the W-potential.

The Hamilton's equations for this system are

$$\frac{dx}{ds} = \frac{p}{m(\tau)} \equiv A(\tau, p), \tag{4.2.27}$$

$$\frac{dp}{ds} = F(x;\tau), \tag{4.2.28}$$

$$\frac{d\tau}{ds} = p_\tau, \tag{4.2.29}$$

$$\frac{dp_\tau}{ds} = F_\tau(x;\tau) + \frac{m'(\tau)}{2m(\tau)}p^2 \equiv B(x,\tau,p). \tag{4.2.30}$$

Note that this calls for a modification to the leapfrog process because p is now coupled with τ. The modified leapfrog process that possesses reversibility and the preservation of the phase space volume is

$$p_\tau(s+\Delta s/2) = p_\tau(s) + \frac{\Delta s}{2}B(x(s),\tau(s),p(s)), \tag{4.2.31}$$

$$x(s+\Delta s/2) = x(s) + \frac{\Delta s}{2}A(\tau(s),p(s)), \tag{4.2.32}$$

$$\tau(s+\Delta s/2) = \tau(s) + \frac{\Delta s}{2}p_\tau(s+\Delta s/2), \tag{4.2.33}$$

$$p(s+\Delta s) = p(s) + \Delta s F(x(s+\Delta s/2);\tau(s+\Delta s/2)), \tag{4.2.34}$$

$$\tau(s+\Delta s) = \tau(s+\Delta s/2) + \frac{\Delta s}{2}p_\tau(s+\Delta s/2), \tag{4.2.35}$$

$$x(s+\Delta s) = x(s+\Delta s/2) + \frac{\Delta s}{2}A(\tau(s+\Delta s),p(s+\Delta s)), \tag{4.2.36}$$

$$p_\tau(s+\Delta s) = p_\tau(s+\Delta s/2) + \frac{\Delta s}{2}B(x(s+\Delta s),\tau(s+\Delta s),p(s+\Delta s)). \tag{4.2.37}$$

If the leapfrog process moves the configuration out of the interval $[\tau_{\min}, \tau_{\max}]$, then we need to revert the configuration back to before (4.2.31), flip the sign of p_τ, and then redo the leapfrog process again.

4.3 Preconditioned flow equation

Although the antiholomorphic gradient flow equation

$$\frac{dz_a}{d\tau} = \overline{\frac{\partial S}{\partial z_a}} \tag{4.3.1}$$

is known to have worked well in many systems, it has a serious problem that can hinder practical calculations. To see this, consider the perturbation of the flow equation:

$$z_a \to z_a + \delta z_a;$$
$$\frac{d\delta z_a}{d\tau} = \overline{\frac{\partial^2 S}{\partial z_a \partial z_b}} \cdot \delta \bar{z}_b.$$

From this equation, one can see that the growth of the flow is governed solely by the eigenvalues of the Hessian $H_{ab}(z) = \partial_a \partial_b S$. The problem arises when the largest and the smallest eigenvalues are too different that one of the modes already diverges before the smallest mode flows sufficiently close to the saddle point. In turn, this makes the Jacobian diverges even before the sign problem can be solved. We will refer to this problem as the anisotropy problem of the Hessian.

One solution is to try generalizing the flow equation in a way that all the nice properties are kept intact while the problem of Hessian anisotropy is also solved. The most general way to do this is to introduce a prefactor matrix A in front of the gradient

$$\frac{dz_a}{d\tau} = A_{ab}(z, \bar{z}) \overline{\frac{\partial S}{\partial z_b}} \tag{4.3.2}$$

At the moment, we will not impose any holomorphicity on A. However, we require A to be a positive-definite Hermitian matrix because we want the imaginary part of the action to remain constant during the flow. In other words, the τ-derivative of the action must be positively real:

$$\frac{dS}{d\tau} = \frac{\partial S}{\partial z_a}\frac{dz_a}{d\tau} = \frac{\partial S}{\partial z_a} A_{ab} \overline{\frac{\partial S}{\partial z_b}} > 0. \tag{4.3.3}$$

Next, one can also show that the perturbation of the modified flow equation is

$$\frac{d\delta z_a}{d\tau} = \delta A_{ab} \overline{\frac{\partial S}{\partial z_b}} + A_{ab} \bar{H}_{bc} \delta \bar{z}_c. \tag{4.3.4}$$

At large τ, where the anisotropy problem appears, the configuration will be forced to be near the saddle point. As a result, only the second term dominates at large τ in practical simulations, which effectively gives

$$\frac{d\delta z_a}{d\tau} \sim A_{ab} \bar{H}_{bc} \delta \bar{z}_c. \tag{4.3.5}$$

Therefore, the ideal modification is to find a positive matrix A such that the product $A\bar{H}$ has eigenvalues of the same modulus; e.g. making $A\bar{H}$ unitary.

On a side note, the ratio between the largest and the smallest singular values of a matrix M is known in the field of numerical analysis as the condition number. Such a condition number is a good indication of the stability of a linear system when solving with iterative methods. The process of minimizing the condition number by means of multiplying M by another matrix A beforehand is called preconditioning, and A is called a preconditioned matrix. The task is essentially the same in our case. As such, we will refer to the modified flow equation as the preconditioned flow equation.

4.3.1 Constant preconditioning

If the Hessian is not fluctuating too much during the simulation, we can approximate it to be a constant matrix defined at some configuration z_0; $H_0 = H(z_0) \sim H(z)$. This

is indeed the case if the system is weakly coupled. The implementation is subsequently straightforward—we replace all gradient ∂S (in the flow equation, forces, etc.) with the preconditioned gradient $v = A\partial S$. In a simple quantum mechanical system where H is tridiagonal, we find that the following choice works the best:

$$A^{-1} = \mathrm{trim}\left(U \cdot \mathrm{diag}(|\lambda_1|, |\lambda_2|, \cdots, |\lambda_N|) \cdot U^\dagger\right) \tag{4.3.6}$$

where λ's are complex eigenvalues of H_0 and U is the diagonalization matrix; $H_0 = U \cdot \mathrm{diag}(\lambda_i) \cdot U^{-1}$. The operation $\mathrm{trim}(\ \cdot\ )$ removes all the non-tridiagonal entries from the matrix. Since the matrix A is constant, it can be evaluated beforehand. And finally, the product $v = A\partial S$ can be calculated using some tridiagonal linear solver for the linear equation

$$A^{-1}v = \partial S. \tag{4.3.7}$$

4.3.2 Dynamical preconditioning

In case the Hessian is not approximately constant during the simulation, the constant preconditioning may be insufficient. This is especially the case with strong couplings. In this case, we choose the preconditioner matrix to be depending on the configuration:

$$A = (\bar{H}H^\intercal)^{-1/2}, \tag{4.3.8}$$

where H is the Hessian of the current configuration. Since the matrix A keeps changing during the simulation, the calculation is considerably slower from the previous case, but it works for arbitrarily large couplings. To see that this choice of preconditioning precisely minimizes the condition number of the flow, we note that the Hessian can be decomposed under the singular value decomposition:

$$H = U\Sigma V \tag{4.3.9}$$

where U and V are some unitary matrices. Then the preconditioner matrix becomes

$$A = (\bar{H}H^\intercal)^{-1/2} = (\bar{U}\Sigma\bar{V}V^\intercal\Sigma U^\intercal)^{-1/2} = (\bar{U}\Sigma^2 U^\intercal)^{-1/2} = \bar{U}\Sigma^{-1}U^\intercal. \tag{4.3.10}$$

It then follows directly that

$$A\bar{H} = \bar{U}\Sigma^{-1}U^\intercal\bar{U}\Sigma\bar{V} = \overline{UV} \tag{4.3.11}$$

is a unitary matrix, and thus has eigenvalues with unit modulus.

Finally, for proof that the preconditioned flow equation doesn't modify the value of the integral, see Appendix B.

Rational approximation

In practice, it is more useful to consider A to be a rational function of $\bar{H}H^\intercal$ instead. We will also write $H^\intercal = H$ from now on because of its symmetricity. To do this, we consider

q	c_q	m_q
0	0.006451003957458213	
1	0.4642523382124909	0.11305025198040564
2	0.6884319202695277	1.5424312774960316
3	0.9119429380769698	6.225923007008075
4	1.196185819926601	17.12763492872428
5	1.6241687265578753	40.03062038540964
6	2.370812308934214	88.4236572105766
7	3.933837875175048	201.19537040130655
8	8.32637138686609	534.9611414154086
9	31.905549463140776	2358.169432739363

Table 4.1: Parameters used in the rational approximation (4.3.12).

the rational approximation, which is a useful technique in lattice calculations involving fermions [123–126],

$$X^{-1/2} = c_0\mathbb{1} + \sum_{q=1}^{Q} c_q (X + m_q\mathbb{1})^{-1}, \tag{4.3.12}$$

with $X = \bar{H}H$, which very well approximates A within the range $a \le \lambda_i \le b$ where λ_i is any of the eigenvalues of X. One particular set of parameters with $Q = 9$, $a = 0.2$, and $b = 2\times10^3$ is given in Table 4.1. The usefulness of this approximation is twofold. Firstly, since the matrix A is not sparse, the multiplication Av takes $O(N^2)$ operations to compute. And secondly, the derivative of an inverse square-root matrix is quite complicate compared to the derivative of just an inverse matrix. In this approximation, the operation Av can be done with the cost of order $O(N)$ since the vector

$$x_q = (\bar{H}H + m_q\mathbb{1})^{-1}v \tag{4.3.13}$$

can be easily obtained from solving the equation

$$(\bar{H}H + m_q\mathbb{1})x_q = v \tag{4.3.14}$$

using the multi-shift solver algorithm [127].

The real derivative $d/d\alpha$ of A can also be evaluated:

$$\frac{dA_{ij}}{d\alpha} = G_{ijk}\frac{dz_k}{d\alpha} + \overline{G_{jik}\frac{dz_k}{d\alpha}}; \tag{4.3.15}$$

$$G_{ijk} = -\sum_{q=1}^{Q} c_q \left[(\bar{H}H + m_q\mathbb{1})^{-1}(\bar{H}\partial_k H)(\bar{H}H + m_q\mathbb{1})^{-1}\right]_{ij}. \tag{4.3.16}$$

Forces

The forces can be computed by first taking the derivative of the flow equation with respect to the real configuration $\alpha \in \{\epsilon, x_i\}$:

$$\begin{aligned}\frac{dz_i'}{d\alpha} &= \frac{dz_i}{d\alpha} + \frac{d\epsilon}{d\alpha} A_{ij}\overline{\partial_j S} + \epsilon \frac{dA_{ij}}{d\alpha}\overline{\partial_j S} + \epsilon A_{ij}\overline{H_{jk}}\frac{d\bar{z}_k}{d\alpha} \\ &= \frac{d\epsilon}{d\alpha}(A\overline{\partial S})_i + (\delta_{ik} + \epsilon G_{ijk}\overline{\partial_j S})\frac{dz_k}{d\alpha} + \epsilon(\overline{G_{jik}\partial_j S} + A_{ij}\overline{H_{jk}})\frac{d\bar{z}_k}{d\alpha}. \end{aligned} \tag{4.3.17}$$

For the temporal force, we replace $\alpha = \epsilon$ above and solve for $\partial z/\partial\epsilon$ at $\sigma = \tau$. Then the temporal force is given by

$$F_\tau = \frac{\epsilon}{\tau}\mathrm{Re}\left(\partial S \frac{\partial z}{\partial \epsilon}\right). \tag{4.3.18}$$

For other forces, we first derive the flow equation for the Jacobian, which can be done by replacing $\alpha = x_i$:

$$J_{ij}' = (\delta_{ik} + \epsilon G_{ilk}\overline{\partial_l S})J_{kj} + \epsilon(\overline{G_{lik}\partial_l S} + A_{il}\overline{H_{lk}})\overline{J_{kj}} \tag{4.3.19}$$

or in the $2N \times 2N$ matrix notation,

$$\begin{pmatrix} J_{ij}' \\ \overline{J_{ij}} \end{pmatrix} = \begin{pmatrix} (\delta_{ik} + \epsilon G_{ilk}\overline{\partial_l S}) & \epsilon(\overline{G_{lik}\partial_l S} + A_{il}\overline{H_{lk}}) \\ \epsilon(G_{lik}\partial_l S + \overline{A_{il}}H_{lk}) & (\delta_{ik} + \epsilon\overline{G_{ilk}}\partial_l S) \end{pmatrix} \begin{pmatrix} J_{kj} \\ \overline{J_{kj}} \end{pmatrix}. \tag{4.3.20}$$

The quantity that we wish to calculate is

$$F_j = \mathrm{Re}(\partial_i S J_{ij}) = \frac{1}{2}\begin{pmatrix} \partial_i S & \overline{\partial_i S} \end{pmatrix}\begin{pmatrix} J_{ij} \\ \overline{J_{ij}} \end{pmatrix} \tag{4.3.21}$$

so we want to calculate the backpropagation of the vector f_i with

$$f_k = f_k' + \epsilon G_{ijk} f_i' \overline{\partial_j S} + \epsilon G_{ijk}\partial_i S, \bar{f}_j' + \epsilon(\bar{f}'\overline{A}H)_k, \tag{4.3.22}$$

where $f(\tau) = \partial S$. Again, the primed quantities are the one associating to the flow time $\sigma + \epsilon$ while the ones without prime are those with flow time σ. This gives the final force

$$F = \mathrm{Re}(f(0)). \tag{4.3.23}$$

4.4 Application to real-time quantum mechanics

4.4.1 Quartic potential

The result in this section is reported in Ref. [122]. The system that we chose to test this algorithm is the one-dimensional quantum mechanics, using the path integral formalism. The important object here is the transition amplitude

$$\mathcal{A}(x_\mathrm{i}, x_\mathrm{f}; T) = \int_{x(0)=x_\mathrm{i}}^{x(T)=x_\mathrm{f}} Dx e^{iS[x(t)]}, \tag{4.4.1}$$

which can be used to generate the time evolution of the wave function via

$$\Psi(x_{\rm f};T) = \int dx \mathcal{A}(x,x_{\rm f};T)\Psi(x;0). \tag{4.4.2}$$

The action that generates this time evolution is

$$S[x(t)] = \int_0^T dt \left\{ \frac{m}{2}\left(\frac{dx}{dt}\right)^2 - V(x) \right\}, \tag{4.4.3}$$

where we use $m = 1$ and the potential we use is the quartic potential $V(x) = \frac{1}{4!}x^4$. We will set the initial wave function to be of Gaussian form. Specifically,

$$\Psi(x;0) = \exp\left\{ -\frac{1}{4\sigma^2}(x-1)^2 \right\}, \tag{4.4.4}$$

with $\sigma = 1$. Since the potential is beyond quadratic, the path integral cannot be exactly solved. However, it is possible to solve the Schrödinger equations for comparison.

To perform the simulation, we discretize the time t as $t_n = (n-1)\epsilon$ and $x_n = x(t_n)$ where $n = 1,..,N+1$ and $\epsilon = T/N$. We also define $x_{N+1} = x_{\rm f}$. Thus, the time-evolved wave function can be associated with the partition function

$$Z(x_{\rm f}) = \int d^N x e^{-S(x;x_{\rm f})} \tag{4.4.5}$$

with x_i $(i = 1,..,N)$ as the dynamical variables. The action in this integral is given by

$$S(x;x_{\rm f}) = \sum_{n=1}^{N} f(x_n, x_{n+1}) + \frac{1}{4}(x_1 - 1)^2; \tag{4.4.6}$$

$$f(x,y) = -i\epsilon \left\{ \frac{m}{2}\left(\frac{x-y}{\epsilon}\right)^2 - \frac{V(x)+V(y)}{2} \right\} \tag{4.4.7}$$

In the following simulation, we take $T = 2$ with $-1 \le x_{\rm f} \le +1$. We also discretize the time with $N = 9$. The flow equation is discretized with $N_\tau = 10$ steps. In the HMC time evolution, the fictitious time is taken to be $s = 0.25$ and is discretized with $N_s = 15$ steps.

Before doing the simulation, it is important to determine the mass function first, which can be obtained from plotting the scatter plot of $\frac{2}{N}\log|\det J|$ against the flow time τ during the testing simulation. We can then fit the data with some function to obtain the logarithm of the mass function. In Fig,4.1, we fit the data with a linear function and get $m(\tau) = \exp(11.795\tau)$. This is consistent with the fact that the typical scale of fluctuation of the configuration grows exponentially with the flow time.

It is also worth noting that the expectation value in the most generalized form is given by

$$\langle \mathcal{O}(x) \rangle = \frac{\langle \mathcal{O}(z) \det J e^{-i{\rm Im}S} e^{-\tilde{W}(\tau)} \rangle_{\rm HMC}}{\langle \det J e^{-i{\rm Im}S} e^{-\tilde{W}(\tau)} \rangle_{\rm HMC}}, \tag{4.4.8}$$

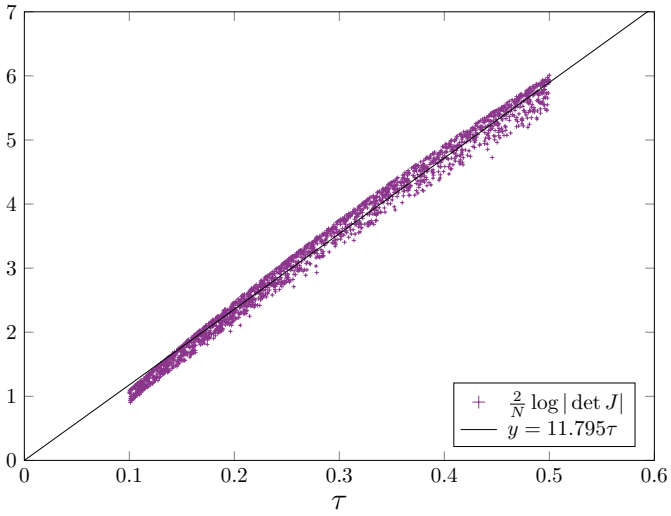

Figure 4.1: Linear fitting for the mass function (4.2.26) that is used in the modified leapfrog integrator.

where $\tilde{W}(\tau)$ (not to be confused with the W-potential) is some function that can be freely chosen in a way that the reweighting factor at different flow times are of similar orders of magnitude. In fact, a sensible choice would be

$$\tilde{W}(\tau) = \frac{N}{2}\log m(\tau), \tag{4.4.9}$$

since $m(\tau)$ is essentially the fitting of the modulus of the original reweighting factor.

The observable that we choose is the log-derivative of the time-evolved wave function

$$\frac{\partial}{\partial x_{\mathrm{f}}}\log\Psi(x_{\mathrm{f}};T) = -\left\langle\frac{\partial}{\partial x_{\mathrm{f}}}f(x_N,x_{\mathrm{f}})\right\rangle \tag{4.4.10}$$

The result at $N = 9$ is shown in Fig. 4.2, which agrees very well with the one obtained from directly diagonalizing the Hamiltonian, corresponding to the continuous time evolution ($N = \infty$).

4.4.2 Double-well potential and quantum tunneling

In the next example, we consider the real-time quantum dynamics under the double-well potential:

$$\mathcal{A}(x_{\mathrm{i}},x_{\mathrm{f}};T) = \int_{x(0)=x_{\mathrm{i}}}^{x(T)=x_{\mathrm{f}}} Dx e^{-S[x(t)]}, \tag{4.4.11}$$

with the action

$$S[x(t)] = i\int_0^T dt\left\{-\frac{m}{2}\left(\frac{dx}{dt}\right)^2 + V(x)\right\}. \tag{4.4.12}$$

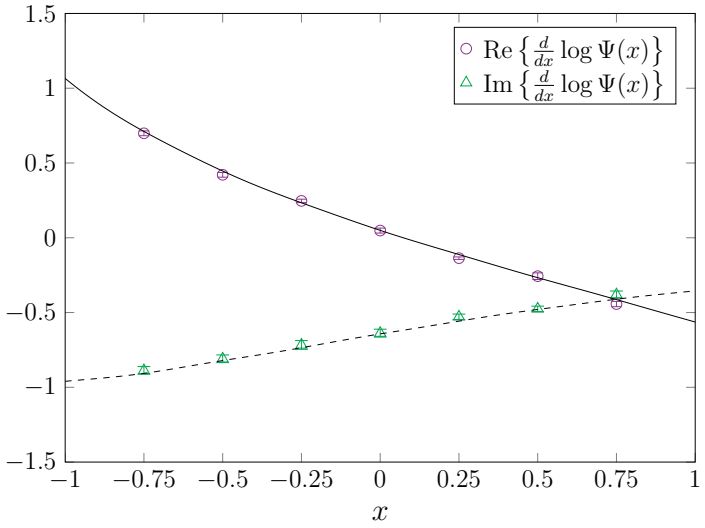

Figure 4.2: The simulated values of the real and the imaginary part of the log-derivative of the wave function, evolving under the anharmonic potential $V(x) = \frac{1}{4!}x^4$ with time $T = 2$. The initial wave function is $\Psi(x;0) = \exp(-\frac{1}{4}(x-1)^2)$. The result from directly solving the Schrödinger equation (solid line for the real part and dashed line for the imaginary part) is also shown for comparison.

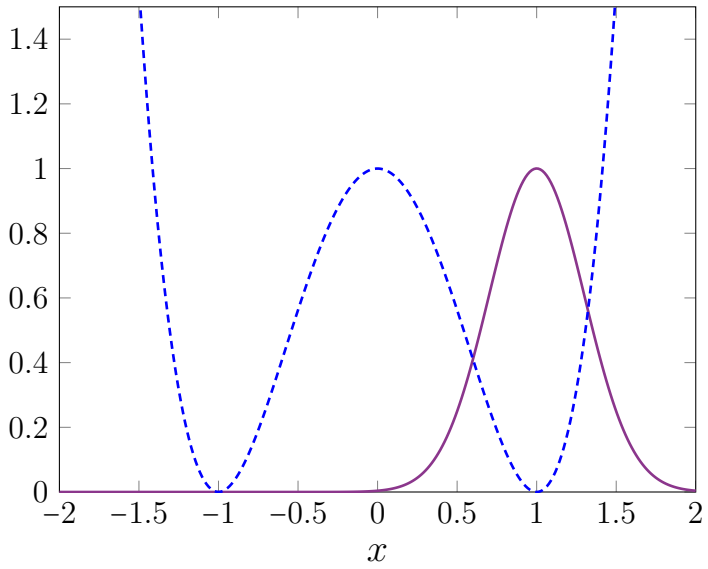

Figure 4.3: Initial distribution function (Purple, Solid) and the double-well potential (rescaled for presentation) (Blue, Dashed).

We use the initial wave function with initial momentum k;

$$\Psi(x;0) = \exp\left\{-\frac{1}{4\sigma^2}(x-1)^2 + ikx\right\}, \tag{4.4.13}$$

with $\sigma = 0.3$, $k = 0$ and 2, and the potential is $V(x) = \frac{5}{2}(1 + x^4 - 2x^2)$. The plot of the initial distribution function $|\Psi(x;0)|^2$ and the potential is shown in Fig. 4.3.

Using direct diagonalization, we can compute the time evolution of the wave function and any expectation values exactly, up to the spatial discretization of the Schrödinger equation. One relevant observable in the context of quantum tunneling is the weakly-measured trajectory, defined by

$$\langle x(t)\rangle_{\mathrm{WM}} = \frac{\langle x_{\mathrm{f}}|e^{-i\hat{H}(T-t)}\hat{x}e^{-i\hat{H}t}|\Psi(x;0)\rangle}{\langle x_{\mathrm{f}}|e^{-i\hat{H}T}|\Psi(x;0)\rangle}. \tag{4.4.14}$$

In terms of the path integral, this amounts to calculating the expectation value of each component of the configuration x_n:

$$\langle x(t_n)\rangle_{\mathrm{WM}} = \langle x_n\rangle. \tag{4.4.15}$$

It should be noted that we can estimate the probability that the particle has the energy smaller than the barrier by looking at the projection between the initial state and energy eigenstates:

$$P(E \le V_0) \coloneqq \sum_{E_n \le V_0} |\langle E_n|\Psi(x;0)\rangle|^2. \tag{4.4.16}$$

The larger the probability, the larger number of complex trajectories we expect to see. From direct diagonalization with $V_0 = 2.5$, we have

$$P(E \le V_0) = \begin{cases} 0.89 & ; k = 0, \\ 0.58 & ; k = 2. \end{cases} \tag{4.4.17}$$

The fact that the case with $k = 2$ has lower probability is expected since adding the initial momentum increases the average kinetic energy of the particle.

In the following simulations, we adopt dynamical preconditioning to the flow equation. We discretize the time with $N = 20$. The flow equation is discretized with $N_\tau = 10$ steps with $0.2 \le \tau \le 4.0$. In the HMC time evolution, the fictitious time is taken to be $s = 1.0$ and is discretized with $N_s = 10$ steps. To check the result, we calculate the weakly-measured trajectory with $x_{\mathrm{f}} = -1$ and compare it with the one solved from direct diagonalization.

As shown in Fig. 4.4, the weakly-measured trajectory for $k = 0$ is indeed complex since $P(E \le V_0)$ is as large as 0.89. But as we increase the momentum to $k = 2$, the trajectory becomes closer to the real axis, shown in Fig. 4.5.

To further explore this, we look at some of the trajectories from the simulation with the largest Boltzmann weight $w = \exp(-\mathrm{Re}S)$. Indeed, for $k = 0$, the trajectories go into the complex plane, compared to those obtained from $k = 2$. This is shown in Fig. 4.6 and 4.7, respectively. The classical solutions are obtained and classified according to Ref. [61].

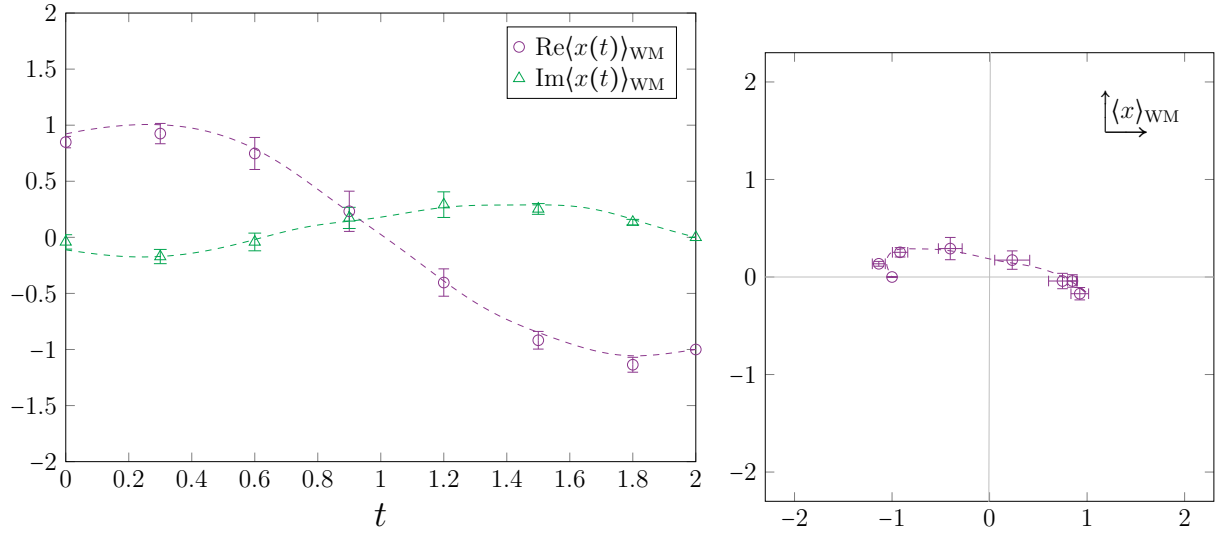

Figure 4.4: (Left) Weakly-measured trajectory for the case $k = 0$. The dashed lines are obtained from the direct diagonalization. (Right) The same trajectory shown on the complex plane.

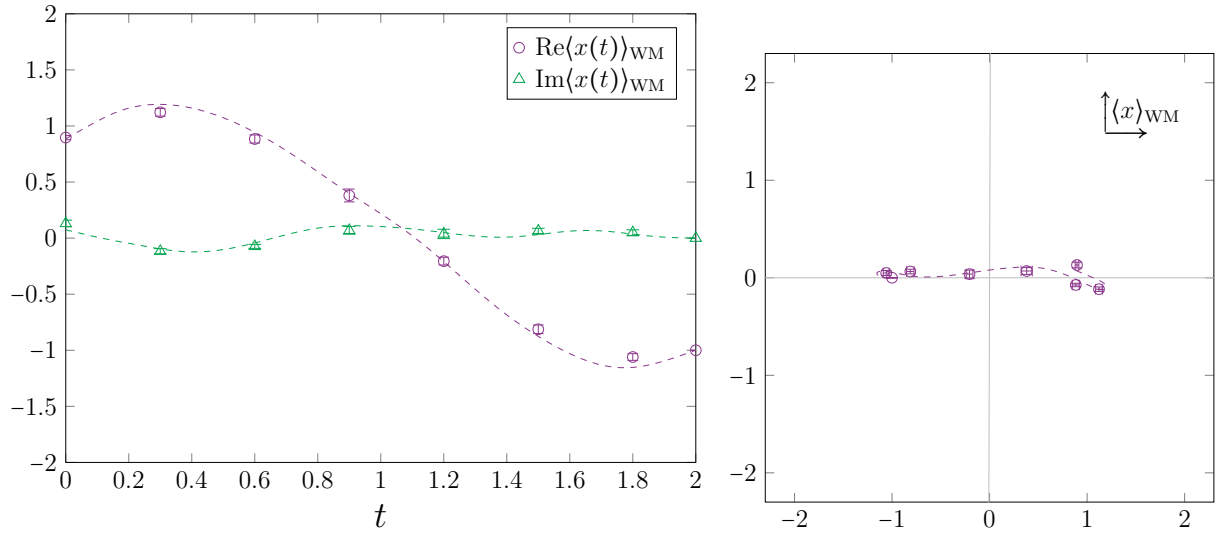

Figure 4.5: (Left) Weakly-measured trajectory for the case $k = 2$. The dashed lines are obtained from the direct diagonalization. (Right) The same trajectory shown on the complex plane.

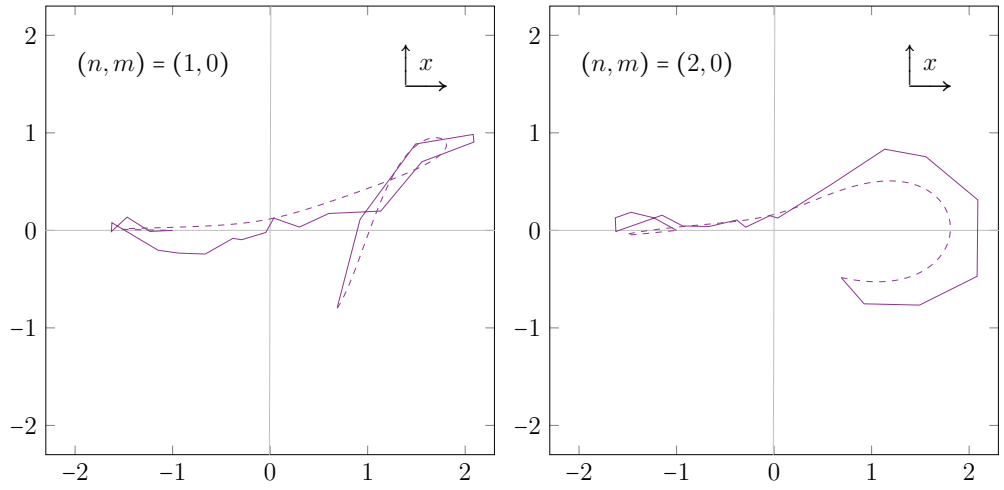

Figure 4.6: Typical trajectories from the simulation with $k = 0$, shown with the associated classical trajectories (dashed lines).

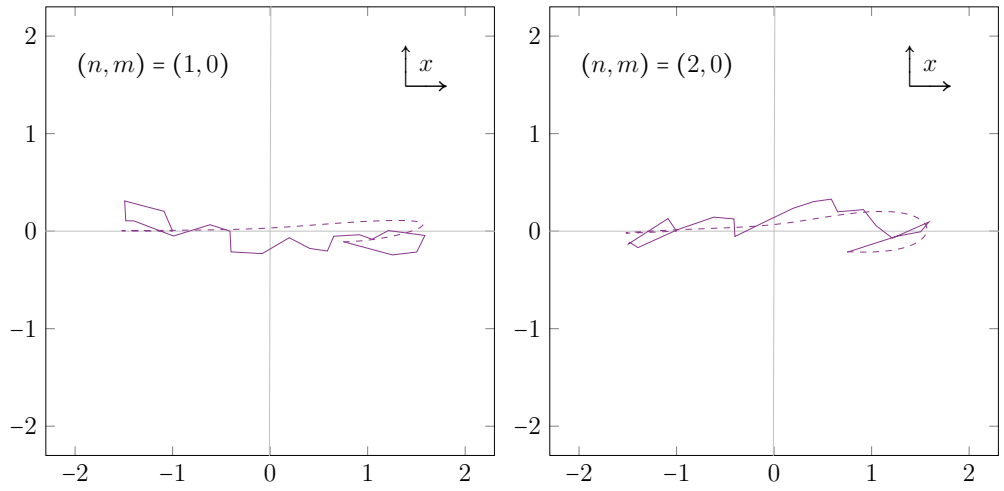

Figure 4.7: Typical trajectories from the simulation with $k = 2$, shown with the associated classical trajectories (dashed lines).

This result signifies the fact that complex saddle points (i.e., classical solutions) play an important role in the thimble approach to quantum tunneling.

It should be stressed that the use of the initial wave function is very crucial for us to obtain complex trajectories. First of all, by having the initial Gaussian wave function with finite width, we impose the cut-off to the initial momentum to those with low values. This lowers the contribution of sphaleron-type saddle points, which are real, to the path integral, and therefore increases the contribution of complex saddle points. And secondly, because the initial point of the trajectory, $x_{\rm i}$, is dynamical, whose distribution is given by the initial wave function, it is allowed to be complex in the thimble simulation. This complex initial point, in turn, makes the resulting classical trajectory complex.

Propagator limit

Quantum tunneling problem in previous literature [61, 62, 117] was discussed in the context of propagator path integral with infinite time; i.e., the two end-points are fixed to some real values with $T \to \infty$. The question is whether complex saddle points such as the infinite spiral trajectory (see Fig. 2.4 (Right)) play an important role in the Lefschetz thimble path integral or not. To consider this setup from our simulation, we have to reduce the width of the initial wave function σ to zero, which we refer to as the 'propagator limit.' Calculations in this limit suffer from various difficulties both in solving the Schrödinger equation and performing the thimble simulation. When we consider the initial wave function with zero width; $\Psi(x;0) = \delta(x - x_{\rm i})$, the distribution of the momentum becomes uniform and spans a wide range of values. High-momentum states, in particular, make solving the Schrödinger equation difficult since spatial discretization error becomes prominent. On the other hand, the thimble simulation in the propagator limit suffers from the ergodicity problem which occurs when important saddle points are too separated that the transition probability becomes exponentially suppressed as we decrease σ. This kind of ergodicity problem stems from a physical origin (similar to the topology freezing problem in lattice gauge theory) and cannot be resolved by adopting the worldvolume approach.

In any case, we can investigate what will happen by looking at some results from solving the Schrödinger equation in the parameter regions where the discretization error is sufficiently small. Considering the fact that the weakly-measured trajectory is a weighted average value of dominant trajectories, we can look at such observable as we approach the propagator limit. In Fig. 4.8, we compare weakly-measured trajectories in the case with $\sigma = 0.2$ and 0.12. One can see that, as we reduce σ, the trajectory becomes less smooth, indicating that higher frequency (larger momentum) modes start to contribute to the path integral, which is as we discussed above. And in Fig. 4.9, we fix the width to $\sigma = 0.15$ and increase the time interval to $T = 4$. One can see that trajectories at large T tend to have more winding than those with smaller T. This is understandable because it is natural that the particle tends to oscillate with a larger number of periods as we increase T. From this result, it is clear that, in the propagator limit with $T \to \infty$, dominant trajectories tend to have a higher frequency and naturally exhibit the spiral behavior. This could be a sign that dominant trajectories indeed become singular, similar to the complexified instanton,

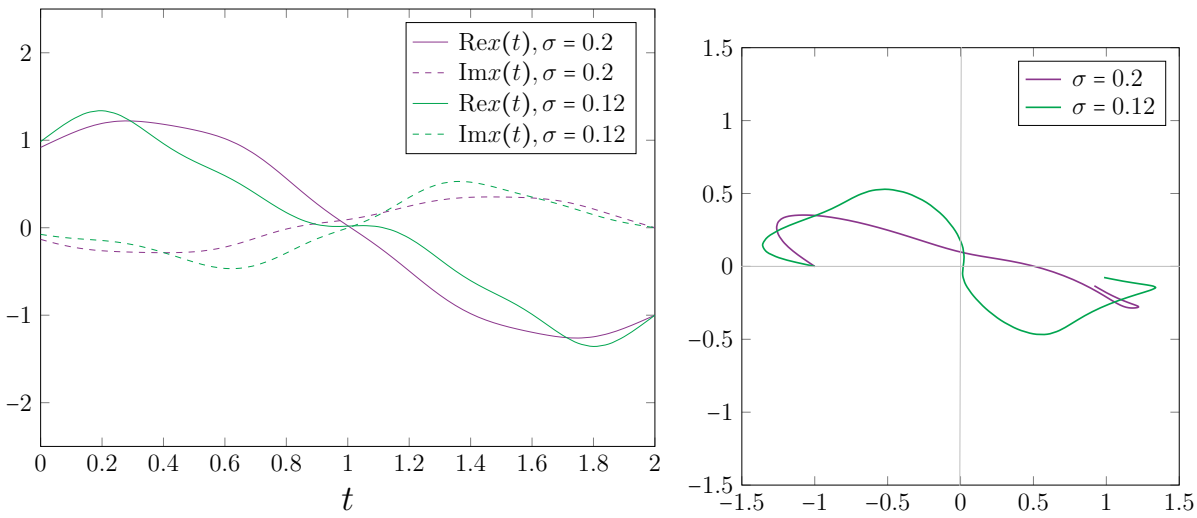

Figure 4.8: (Left) Weakly-measured trajectory with the initial width $\sigma = 0.2$ (Purple) and 0.12 (Green). Higher frequency contributions become more noticeable in smaller σ. (Right) The same trajectories on the complex plane.

in this limit. It should be noted that this singular behavior completely disappears in a more realistic setup where the initial wave function is of finite width, and with finite time evolution.

4.5 Brief summary

In this chapter, I explain the generalized Lefschetz thimble methods and their application to real-time quantum mechanics. I first introduce the Picard-Lefschetz theory in section 4.1. In section 4.2, I describe several implementations of the HMC algorithm to the thimble simulations, including the traditional simulation, the parallel tempering method, the worldvolume approach, and the newly-developed backpropagating HMC algorithm. In section 4.3, we explain a new flow equation that has a better behavior at large flow time. Specifically, we apply the preconditioner matrix to the gradient to minimize the condition number of the Hessian. And in section 4.4, we show the results of applying thimble simulations to real time evolution in quantum mechanics. The results for both quartic and double-well potentials agree with the result from direct diagonalization of the Hamiltonian very well. We also show some of the classical trajectories drawn from the simulations. It is clear from this result that the initial wave function is necessary for the existence of complex trajectories. It should be stressed that such kind of study is very difficult without the use of the generalized Lefschetz thimble technique.

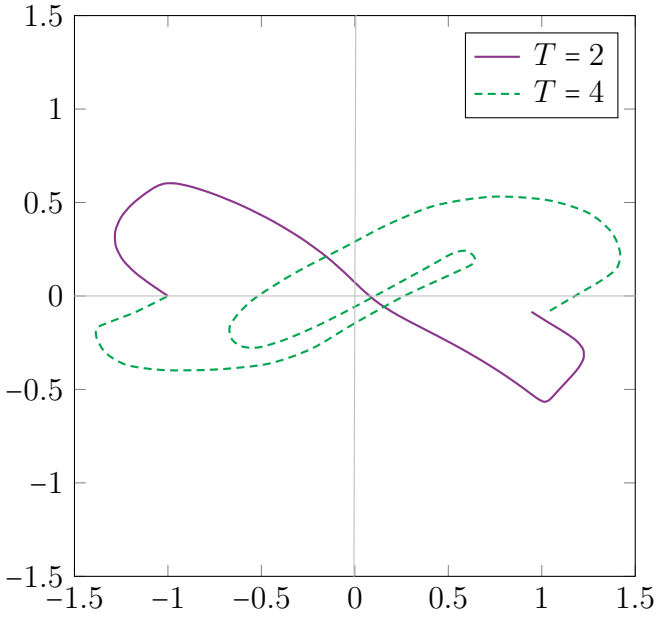

Figure 4.9: A weakly-measured trajectory in the complex plane for $\sigma = 1.5$ and $T = 2, 4$.

Chapter 5

Tensor renormalization group

5.1 Overview

In the field of quantum many-body systems, several techniques involving tensor networks had been developed with varying success. However, the first technique that is considered a true real-space renormalization was proposed in Ref. [65], the tensor renormalization group (TRG), as a useful tool to study two-dimensional systems. The technique is rather simple and allows for many generalizations to various systems and dimensions. One of the key advantages of the TRG approach is its easy accessibility to large volume lattices. This is due to the cross-graining procedure which is the key essence of the method. In this section, we will explain how it works in two dimensions.

Assuming that we have already rewritten the partition function as a tensor network, whose detailed process can be different depending on the system, the remaining task is performing coarse-graining. This consists of two main steps. The first is the decomposition of the fundamental tensor T using the singular-value decomposition (SVD). For higher-rank tensors, the decomposition is not unique, but we will consider two types:

$$T_{pqrs} = \underbrace{\sum_c S^{(1)}_{ps;c} G^{(1)}_c S^{(2)}_{rq;c}}_{\text{type I}} = \underbrace{\sum_c S^{(3)}_{pq;c} G^{(2)}_c S^{(4)}_{rs;c}}_{\text{type II}}, \tag{5.1.1}$$

where $G^{(1)}$ and $G^{(2)}$ are the two singular-value vectors. In practice, it is necessary to truncate the summation to include only a certain number of largest singular values, which we will denote it by D_{cut}. The two decompositions are depicted diagramatically in Fig. 5.1. Since singular values are positive semi-definite, we can absorb them into the $S^{(i)}$ tensors:

$$\tilde{S}^{(1)}_{pq;c} = S^{(1)}_{pq;c} \sqrt{G^{(1)}_c}, \tag{5.1.2}$$

$$\tilde{S}^{(2)}_{pq;c} = S^{(2)}_{pq;c} \sqrt{G^{(1)}_c}, \tag{5.1.3}$$

$$\tilde{S}^{(3)}_{pq;c} = S^{(3)}_{pq;c} \sqrt{G^{(2)}_c}, \tag{5.1.4}$$

$$\tilde{S}^{(4)}_{pq;c} = S^{(4)}_{pq;c} \sqrt{G^{(2)}_c}. \tag{5.1.5}$$

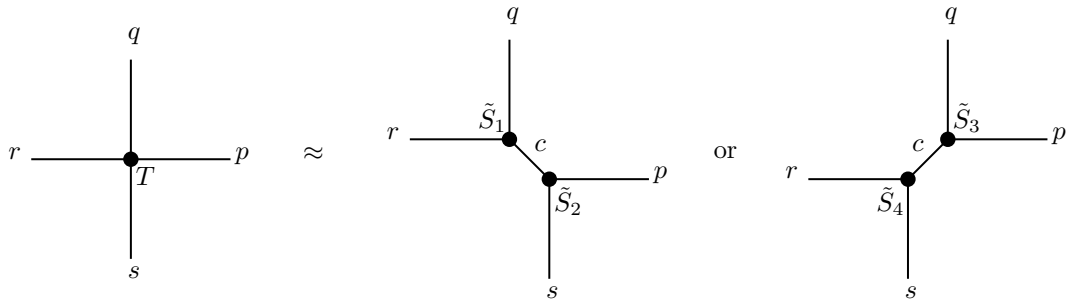

Figure 5.1: Two types of singular values decompositions in two dimensions. The left and right decompositions are type I and II, respectively.

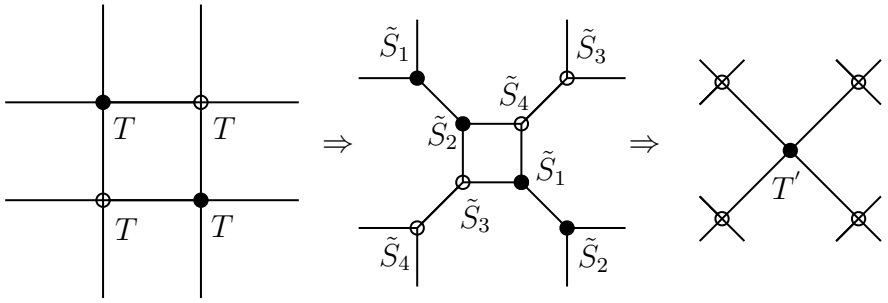

Figure 5.2: The summary of coarse-graining procedure.

In our tensor network, we apply the two decompositions on the even and odd sites alternately. This is depicted as the first step of Fig. 5.2.

In the second step, we recombine the decomposed tensor into a new rank-4 tensor

$$T'_{abcd} = \sum_{pqrs} \tilde{S}^{(1)}_{qr;a} \tilde{S}^{(3)}_{rs;d} \tilde{S}^{(2)}_{sp;c} \tilde{S}^{(4)}_{pq;b} \tag{5.1.6}$$

which represents the coarse-grained version of the original tensor T_{pqrs}. This is diagramatically depicted as the second step of Fig. 5.2.

The resulting tensor network is identical to the original tensor network, but is tilted by 45°. However, the new tensor T contains the information of two of the original T tensor, effectively reduce the size of the network from $L \times L$ into $(L/2) \times (L/2)$. In other words, if we start with a $2^n \times 2^n$ lattice, we can perform this process $2n$ times to arrive at a 'one-site' model, whose partition function is

$$Z = \sum_{p,q} T_{pqpq} \tag{5.1.7}$$

where T is the tensor obtained from the coarse-graining.

The TRG is actually applicable in more general lattices, such as honycomb or triangular types [65]. Its simplicity also allows many efficient generalizations which is applicable to higher dimensions [77–79]. One of the generalization known as the Grassmann TRG even allows us to coarse-grain fermionic fields directly without the needs to perform any high-cost 'bosonization' [66–76].

5.2 Tensor network constructions of 2D gauge theory

In this section, we will discuss different ways to construct the tensor network of the 2D gauge theory at $\theta = 0$. The construction with nonzero θ follows straightforwardly and will be discussed in greater detail in the case of character expansion construction in Section 5.5. Let us first look at the partition function of the discretized theory:

$$Z = \int DU e^{-S_g[P]}; \tag{5.2.1}$$

$$S_g[P] = -\frac{N}{\lambda} \sum_{n\in\Lambda} \mathrm{tr}(P_n + P_n^\dagger), \tag{5.2.2}$$

$$P(n) = U_{n,1} U_{n+\hat{1},2} U_{n+\hat{2},1}^\dagger U_{n,2}^\dagger. \tag{5.2.3}$$

There are two choices about how we can rewrite this into a network; working on the real space and working on the reciprocal space. In the real space, this can be done with a quadrature or a statistical sampling, while in the reciprocal space, this can be done via the character expansion.

5.2.1 Gauss-Legendre quadrature

This method is particularly useful when the gauge group of interest is of low rank. In particular, for U(1) gauge theory [128], the group integral simply become an integral of a periodic function. Thus we can replace the Haar measure by a Lebesgue measure, whose integral in turn can be approximated by a Gauss-Legendre quadrature

$$\int_{\mathrm{U}(1)} dU g(U) = \int_{-\pi}^{+\pi} \frac{d\varphi}{2\pi} g(\exp(i\varphi)) \approx \sum_{i=1}^{K} w_i g(\exp(i\varphi_i)). \tag{5.2.4}$$

Or in general, we have

$$\int_{\mathrm{U}(1)} dU g(U, U_{n,\mu}, U_{n',\mu'}, \cdots) \approx \sum_{i=1}^{K} w_i g(\exp(i\varphi_i), U_{n,\mu}, U_{n',\mu'}, \cdots). \tag{5.2.5}$$

This approximation can be performed for all the link variables, leaving the partition function as

$$Z = \mathrm{tr} \bigotimes_{n\in\Lambda} T(n), \tag{5.2.6}$$

$$T_{ijkl} = \frac{\sqrt{w_i w_j w_k w_l}}{4\pi^2} \exp\left\{ \frac{1}{\lambda} \cos(\varphi_i + \varphi_j - \varphi_k - \varphi_l) \right\}. \tag{5.2.7}$$

In the equations above, the symbol $\bigotimes_n$ indicates that the tensor T which has 4 legs is tiled periodically as depicted in Fig. 5.4, and the trace simply means that all indices are contracted with periodic boundary condition. The square root in the prefactor comes from the fact that each link variable is shared by two plaquettes. And since the tensor T is

associated with a plaquette, each of the factor w_i must be symmetrically separated into two, thus the square root. In three dimensions, for example, where a link variable is shared by 4 plaquettes, the factor should instead be of power 1/4.

It should be noted that this method is less practical in groups with higher ranks because the group manifold now becomes multi-dimensional and with nontrivial topology. Therefore, in the author's opinion, this method is most suitable only for U(1) gauge theory and other topologically similar groups.

5.2.2 Statistical sampling

In this method, the important point is the realization that we can approximate a group integral by a Monte Carlo integral [129, 130]

$$\int dU g(U, U_{n,\mu}, U_{n',\mu'}, \cdots) e^{-S_v(U)} \approx \frac{Z_v}{K} \sum_{i=1}^{K} g(U_i, U_{n,\mu}, U_{n',\mu'}, \cdots); \tag{5.2.8}$$

$$Z_v = \int dU e^{-S_v(U)}, \tag{5.2.9}$$

where U_i are drawn from the distribution $\exp(-S_v(U))$ with $S_v(U)$ being some trial action. In general, the final result should be independent of $S_v(U)$ as long as K is sufficiently large. In other words, if $S_v(U)$ is chosen poorly, the final result will still be sensitive to K. With this in mind, we rewrite the partition function as

$$Z = \int DU \exp\left(-S_g[P] + \sum_{\mu,n} S_v[U_{n,\mu}]\right) \exp\left(-\sum_{\mu,n} S_v[U_{n,\mu}]\right) \tag{5.2.10}$$

and approximate it by a Monte Carlo integral, giving

$$Z = Z_v^{2V} \mathrm{tr} \bigotimes_{n\in\Lambda} T(n), \tag{5.2.11}$$

$$T_{ijkl} = \exp\left\{\frac{2N}{\lambda} \mathrm{Retr}(U_i U_j U_k^\dagger U_l^\dagger) + \frac{1}{2}\left(S_v(U_i) + S_v(U_j) + S_v(U_k) + S_v(U_l)\right)\right\}. \tag{5.2.12}$$

This construction is similar to the case of Gauss-Legendre quadrature. The difference is essentially the way we approximate the group integral by a finite sum. The factor 1/2 in the second term comes from the fact that each link variable are shared by two plaquettes. And since the tensor T is associated to a plaquette, the action $S_v[U]$ must be symmetrically devided into two. In three dimensions, for example, where a link variable is shared by 4 plaquettes, the factor should instead be 1/4.

This technique is very versatile and doesn't have many drawbacks in going to higher dimensions or to more complicated gauge groups. However, it is still a non-trivial task to find the best trial action $S_v[U]$ which is not sensitive to the size of the ensemble K.

5.2.3 Character expansion

A different way to construct a tensor is to rewrite the theory in the reciprocal/dual space. Namely, instead of writing the partition function as an integral of group elements, we rewrite

it as the sum of irreducible representations of the plaquette. This process is known as the character expansion [131–133]. The simplest analogy of this manipulation is the fact that the integral of a periodic function can be replaced by the summation of its Fourier modes. This connection is hopefully made clear as we move along the derivation below.

As the first step, we separate the Boltzmann weight into a product of a function of a plaquette

$$e^{-S_g[P]} = \prod_n f(P_n), \tag{5.2.13}$$

$$f(P) = \exp\left\{\frac{N}{\lambda}\mathrm{tr}(P + P^\dagger)\right\}. \tag{5.2.14}$$

The function $f(P)$ has a special property that it is invariant under the conjugation map on the gauge group $G = \mathrm{U}(N)$ or $\mathrm{SU}(N)$

$$f(P) = f(g^{-1}Pg), \qquad \forall g \in G. \tag{5.2.15}$$

This type of functions is known as the class function, which can always be expanded in terms of characters of the argument

$$f(P) = \sum_r \tilde{\beta}_r \mathrm{tr}_r(P) \tag{5.2.16}$$

where the sum is taken over all irreducible representations of G. For $G = \mathrm{U}(1)$, the class function is also known as the periodic function, and the corresponding character expansion is known as the Fourier series, thus the analogy given above. From now on, we will refer to irreducible representations as 'representations' for simplicity. The expansion coefficient $\tilde{\beta}_r$ can be evaluated via

$$\tilde{\beta}_r = \int dU f(U)\mathrm{tr}_r(U^\dagger). \tag{5.2.17}$$

In our particular case, the coefficients can be evaluated as

$$\tilde{\beta}_r = \begin{cases} \det \mathcal{M}_{r,0} & ; G = \mathrm{U}(N), \\ \sum_{q\in\mathbb{Z}} \det \mathcal{M}_{r,q} & ; G = \mathrm{SU}(N). \end{cases} \tag{5.2.18}$$

The matrix $\mathcal{M}_{r,q}$ is given by

$$(\mathcal{M}_{r,q})_{ij} = \int_{-\pi}^{+\pi} \frac{d\phi}{2\pi} \cos\{(l_i + i - j + q)\phi\} \exp\left(\frac{2N}{\lambda}\cos\phi\right) = I_{l_j+i-j+q}\left(\frac{2N}{\lambda}\right) \tag{5.2.19}$$

for $i, j = 1, 2, .., N$ where $I_n(x)$ is the modified Bessel function of the first kind. The representation r is labeled by N integers $\{l_1, .., l_N\}$ in the descending order

$$l_1 \geq l_2 \geq \cdots \geq l_N \tag{5.2.20}$$

with an extra constraint $l_N = 0$ in the $\mathrm{SU}(N)$ case, where the label $\{l_i\}$ also represents the Young tableau whose i^{th} row has l_i boxes. For both $\mathrm{U}(N)$ and $\mathrm{SU}(N)$, the dimensionality of a representation $r = \{l_i\}$ is given by

$$d_r = \prod_{1\leq i<j\leq N} \left(1 + \frac{l_i - l_j}{j - i}\right). \tag{5.2.21}$$

Note that a $\mathrm{U}(N)$ representation $r^{(\mathrm{U})} = \{l'_i\}$ can be obtained uniquely from an $\mathrm{SU}(N)$ representation $r^{(\mathrm{SU})} = \{l_i\}$ via the relation

$$l'_i = l_i + q \tag{5.2.22}$$

with some integer q.

The representation matrix $D^{r^{(\mathrm{U})}}(g)$ of some $g \in \mathrm{U}(N)$ is related to $D^{r^{(\mathrm{SU})}}(\tilde{g})$ of some $\tilde{g} \in \mathrm{SU}(N)$ via

$$D^{r^{(\mathrm{U})}}(g) = e^{ip\theta} D^{r^{(\mathrm{SU})}}(\tilde{g}) \tag{5.2.23}$$

$$p \equiv \sum_{i=1}^{N} l'_i = \sum_{i=1}^{N-1} l_i + Nq. \tag{5.2.24}$$

where $g = e^{i\theta}\tilde{g}$. This transformation rule implies that the right-hand side of (5.2.23) is invariant under the transformation $\tilde{g} \to e^{2\pi ik/N}\tilde{g} \in \mathrm{SU}(N)$ and $\theta \to \theta - 2\pi k/N$ for some integer $k \in \mathbb{Z}$.

In the followings, we will use the following notation:

$$\begin{array}{lll} \mathrm{trv}^{(\mathrm{U})} & = \{0,0,..,0,0\}, & \mathrm{trv}^{(\mathrm{SU})} = \{0,0,..,0,0\}, \\ \mathrm{fnd}^{(\mathrm{U})} & = \{1,0,..,0,0\}, & \mathrm{fnd}^{(\mathrm{SU})} = \{1,0,..,0,0\}, \\ \overline{\mathrm{fnd}}^{(\mathrm{U})} & = \{0,0,..,0,-1\}, & \overline{\mathrm{fnd}}^{(\mathrm{SU})} = \{1,1,..,1,0\}, \\ \mathrm{adj}^{(\mathrm{U})} & = \{1,0,..,0,-1\}, & \mathrm{adj}^{(\mathrm{SU})} = \{2,1,..,1,0\}, \end{array} \tag{5.2.25}$$

where "trv", "fnd", and "adj" stand for trivial, fundamental, and adjoint representations respectively, and the representations with a bar mean the complex conjugate representation.

Using the character expansion, we can rewrite the partition function as

$$Z = \int DU \prod_n f(P_n) = \sum_{\{r(n)\}} \int DU \prod_n \tilde{\beta}_{r(n)} \mathrm{tr}_{r(n)} P_n, \tag{5.2.26}$$

where $r(n)$ is the representation that appears in the expansion of $f(P_n)$. Next, we decompose $\mathrm{tr}_{\mathrm{r(n)}}$ in terms of link variables as

$$\mathrm{tr}_{r(n)} P_n = D^{r(n)}_{\alpha\beta}(U_{n,1}) D^{r(n)}_{\beta\gamma}(U_{n+\hat{1},2}) D^{r(n)}_{\gamma\delta}(U^{\dagger}_{n+\hat{2},1}) D^{r(n)}_{\delta\alpha}(U^{\dagger}_{n,2}). \tag{5.2.27}$$

The matrix indices a, b, c, d in this equation are summed over implicitly. In two dimensions, each link variable only appears twice in the partition function, one as U and the other as $U^{\dagger}$. Thus, we can factorize the partition function as

$$Z = \mathrm{tr} \bigotimes_{n,\mu} z^{(n,\mu)}, \tag{5.2.28}$$

$$(z^{(n,\mu)})_{rs,\alpha\beta\gamma\delta} = \tilde{\beta}_r^{1/4} \tilde{\beta}_s^{1/4} \int dU_{n,\mu} D^r_{\alpha\beta}(U_{n,\mu}) D^r_{\gamma\delta}(U^{\dagger}_{n,\mu}) = \frac{\tilde{\beta}_r^{1/2}}{d_r} \delta_{rs} \delta_{\alpha\delta} \delta_{\beta\gamma}, \tag{5.2.29}$$

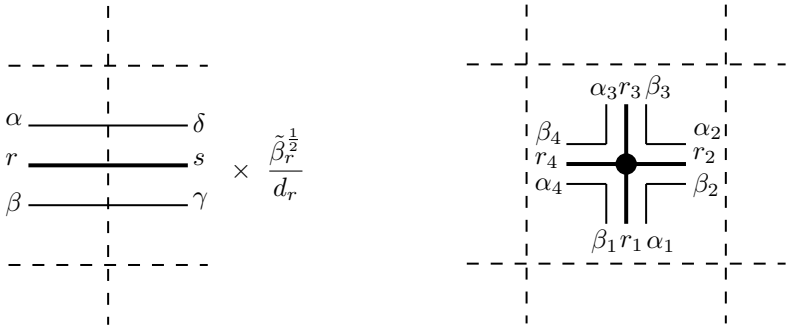

Figure 5.3: Two types of the tensors appearing from character expansion.

where we have used the orthogonality relation

$$\int dU D^r_{ab}(U) D^s_{cd}(U^\dagger) = \frac{1}{d_r} \delta_{rs} \delta_{\alpha\delta} \delta_{\beta\gamma} \tag{5.2.30}$$

with d_r being the dimensionality of the representation r given by (5.2.21).

The tensor (5.2.29) can be considered as a tensor associated to a link variable as depicted as a diagram in Fig. 5.3 (left). Note that these tensors must be contracted in the partition function according to (5.2.27), therefore, we need to also include a tensor

$$\delta_{r_1 r_2 r_3 r_4} \delta_{\alpha_1 \beta_1} \delta_{\alpha_2 \beta_2} \delta_{\alpha_3 \beta_3} \delta_{\alpha_4 \beta_4} \tag{5.2.31}$$

associated to each plaquette as depicted as a diagram Fig. 5.3 (right), where we have defined a 4-way Kronecker delta $\delta_{abcd} = \delta_{ab}\delta_{bc}\delta_{cd}$.

After patching up all these tensors together, one can notice that there is a loop of matrix indices on each lattice site. Contracting this loop gives us a factor of d_r on each site. After appropriately reassigning all the factors in to the tensor associating to the plaquette, we have a reduced tensor

$$T_{pqrs} = \frac{\tilde{\beta}_p}{d_p} \delta_{pqrs} \tag{5.2.32}$$

and the partition function is simply given by the tensor network

$$Z = \mathrm{tr} \bigotimes_{n \in \Lambda} T(n), \tag{5.2.33}$$

which can be depicted as a diagram as in Fig. 5.4. Since the tensor is already 'diagonalized', the contraction can be done straightforwardly, giving

$$Z = \sum_r \left(\frac{\tilde{\beta}_r}{d_r} \right)^{L_1 \times L_2} \tag{5.2.34}$$

where we assume that the lattice is a periodic $L_1 \times L_2$ regtangular Euclidean lattice.

There are a few comments regarding the construction via character expansion in relation to the previous two constructions. Firstly, the fundamental tensor (5.2.32) is in some sense

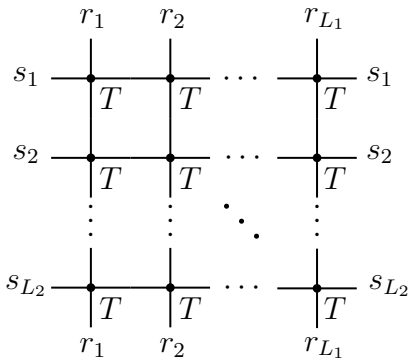

Figure 5.4: The tensor network representation of the partition function (5.2.1).

the 'reduced' version of the tensors from the other construction. This is because we have summed the loop of the matrix indices. In this sense, the singular value profile of this tensor will be different since the 'redundancy' of the singular values arising from the loop is already absorbed into the tensor. This redundancy issue is discussed in Ref. [129]. Secondly, by using character expansion, the tensor becomes very simplified as it is already diagonalized. Although its simplicity may seem discouraging since it undoubtedly trivializes the computation, another way to look at it is we can now easily study large-N gauge theory because the analysis now becomes straightforward, unlike the other two constructions.

Singular-value spectrum

As already mentioned, the fundamental tensor in this construction is already 'diagonalized,' making the singular-value decomposition trivial. Specifically, the singular-value vector of the fundamental tensor is

$$G_r^{(i)} = \sigma_r = \left|\frac{\tilde{\beta}_r}{d_r}\right|. \tag{5.2.35}$$

After performing a single coarse-graining procedure, the tensor becomes

$$T'_{pqrs} = (\sigma_p)^2 \delta_{pqrs}. \tag{5.2.36}$$

After performing this n times, it is easily convinced that the singular value becomes

$$\sigma_r \to \sigma_r^{2^n}. \tag{5.2.37}$$

It is a special property of the 2D gauge theory that the singular value has a certain scaling property like this, instead of mixing among each other during the process.

And finally, using (5.1.7), one can show that the partition function after performing coarse-graining n times, becomes

$$Z = \sum_r \left(\frac{\tilde{\beta}_r}{d_r}\right)^V \tag{5.2.38}$$

where $V = 2^n$, agreeing with the known result in (5.2.34).

5.3 Restricting the number of representations

As described above, the number of representations of U(N) and SU(N) groups are infinite. This also implies that the size of the fundamental tensor (5.2.32) is also infinite. This is not practical, so we need to somehow restrict to only the representations with the largest contributions.

For U(1) and SU(2), this process is rather simple, since the representations are labeled by a singer integer, which is the charge and the angular momentum respectively. But in groups with higher rank, the weight diagram forms a more complicated lattice. In this work, we propose an efficient way to choose only the representations that contribute the most, even at large N.

Let us first consider the SU(N) case. Recall that any representation can be labelled by $N-1$ non-negative integers $\{l_i\}$ where ($l_N = 0$)

$$l_1 \geq l_2 \geq \cdots \geq l_{N-1} \geq 0. \tag{5.3.1}$$

A natural cutoff would be Λ where $l_1 \leq \Lambda$. This cutoff already include the conjugate representation of any included representation since the conjugate representation of $r = \{l_i\}$ is given by $\bar{r} = \{\bar{l}_i\}$ where $\bar{l}_i = l_1 - l_{N+1-i} \leq \Lambda$. It can be easily proven that the number of representations within this cutoff is ${}_{\Lambda+N-1}\mathrm{C}_{\Lambda}$, which grows as Λ^{N-1} with Λ for fixed N.

Alternatively, one can put another cutoff on dimensionality: $d_r \leq \Delta_\Lambda$, where

$$\Delta_\Lambda = \prod_{j=2}^{\infty}\left(1 + \frac{\Lambda}{j-N}\right) = {}_{\Lambda+N-1}\mathrm{C}_{\Lambda}. \tag{5.3.2}$$

This can be shown to be the smallest dimensionality of all the representations with $l_1 = \Lambda$.

After making an appropriate cutoff condition, we then calculate the singular value

$$\sigma_i = \left|\frac{\tilde{\beta}_r}{d_r}\right| \tag{5.3.3}$$

of all the representations and then keep only a specified number of representation to be used in the construction of the fundamental tensor.

As for the U(N) case, we can make use of te relation (5.2.22). We first obtain the set of representation as if the group was SU(N). Then we attach the 'charge' q to these representation. The interval of q is restricted by another cutoff Λ_q, where

$$\left|q + \frac{l_1}{2}\right| + \frac{l_1}{2} \leq \Lambda_q. \tag{5.3.4}$$

This choice of cutoff guarantees that both a representation and its conjugate are included in this cutoff. Also, note that a simple cutoff of $|q| \leq \Lambda_q$ for U(1) is a special case for this since $l_1 = 0$ in this case.

5.4 Large-N insights via TRG analysis

One question one may ask is whether the theory has any undesirable behavior in the large-N limit. To answer that, we investigate singular values of both the U(N) and SU(N) theories at large N. We will separate the discussions between the strong coupling $\lambda > 2$ and the weak coupling region $\lambda < 2$ discussions since they are separated by the Gross-Witten-Wadia third-order phase transition [134, 135].

Let us first define what it means as a large-N limit of a certain representation. We will only focus on U(N) since SU(N) representations can be directly obtained from U(N) representations by subtracting all the l'_i with l'_N. Simple representations in (5.2.25) can easily be defined in large-N by adding zeros in the middle of the sequence $\{l'_i\}$. We can generalize this to any representation, namely by adding zeros in the middle. Although this generalization obviously introduces some ambiguity, such as when a single representation can possibly belong to multiple sequences, this is not a problem at large N. Therefore, in the discussion below, it should be understood that the symbol $r^{(\mathrm{U})}$ will refer to a large-N sequence of a representation obtained from adding zeros in the middle of $\{l_i\}$. These obviously do not exhaust all U(N) representations. Specifically, we will refer to these representations as having zero 'U(1) charge,' $q = 0$. We can also defined another related sequence to $r^{(\mathrm{U})}$ but with nonzero charges by the symbol $(r^{(\mathrm{U})}, q)$.

The conclusion of our observation is that the singular value profile for these theory have well-defined limit at large N:

$$\log \sigma_r = \begin{cases} C^{(0)}N^2 + C_r^{(1)} + C_r^{(2)}N^{-2} + \mathcal{O}(N^{-3}) & ; \quad \mathrm{SU}(N) \text{ any } \lambda, \\ C^{(0)}N^2 + C_r^{(1)} + \mathcal{O}(N^{-1}) & ; \quad \mathrm{U}(N)\ \lambda < 2, \\ C^{(0)}N^2 + C_r^{(1)} + a_r N + b_r \chi_N + \mathcal{O}(N^{-1}) & ; \quad \mathrm{U}(N)\ \lambda > 2. \end{cases} \tag{5.4.1}$$

where χ_N, which grows like $\mathcal{O}(\log N)$, is defined in (5.5.11). The reason for this specific asymptotic form of the U(N) cases will be explained in later section. It should also be noted that the coefficient of the leading term $C^{(0)}$ for a given λ are all common for every representation. This will have an important implication as will be discussed after this. An example at $\lambda = 3$ and $\lambda = 1.5$ are shown in Fig. 5.5.

Another important point for U(N) is the fact that, for $\lambda > 2$, the coefficient a_r actually only depends on the U(1) part of the representation r, which we will simply refer to as the 'U(1) charge'. We can directly look at the quantity $(\log \sigma_r - C^{(0)}N^2)/N$, which is shown in Fig. 5.6. The significance of this is that, since a_r are all negative except for $q = 0$ as shown in the figure, the singular values with nonzero charges all exponentially suppressed at large N. And since we have the relation (derived from (5.2.18))

$$\sigma_{r^{(\mathrm{SU})}} = \sum_{q=-\infty}^{+\infty} \sigma_{(r^{(\mathrm{U})},q)} \tag{5.4.2}$$

we arrive at the fact that the only term on the right-hand side that dominates at large N and $\lambda > 2$ is $\sigma_{(r^{(\mathrm{U})},0)}$. In other words, we have a large-N equivalence between singular values of U(N) and SU(N) theories at large coupling.

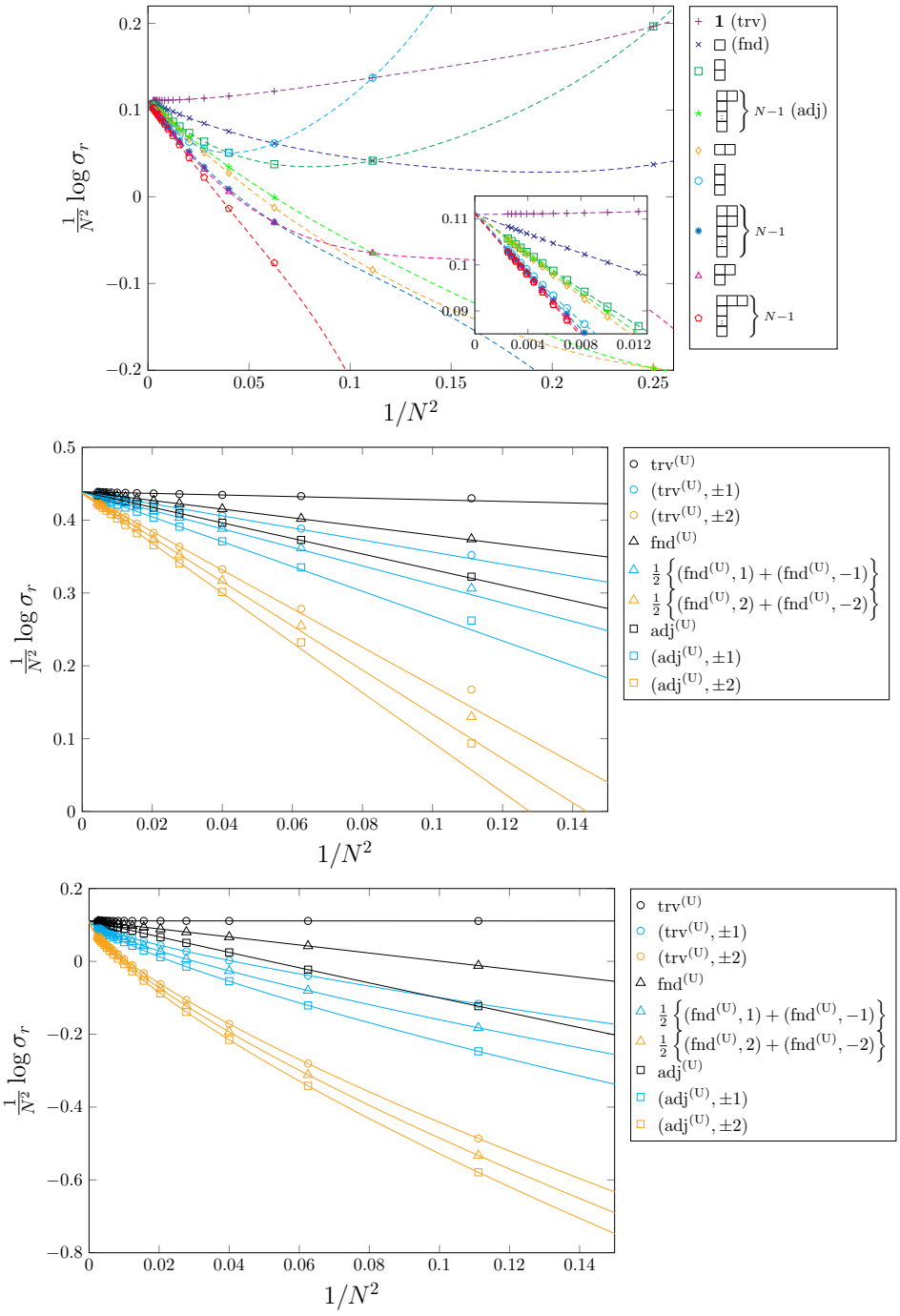

Figure 5.5: Singular values of different theories in the large-N limit. (Top) SU(N) theory with $\lambda = 3$, (Middle) U(N) theory with $\lambda = 1.5$, and (Bottom) U(N) theory with $\lambda = 3$. All of them are fitted with (5.4.1). The notation $(r^{(\mathrm{U})}, q)$ refers to the representation $\{l'_i + q\}$ when $r^{(\mathrm{U})} = \{l'_i\}$.

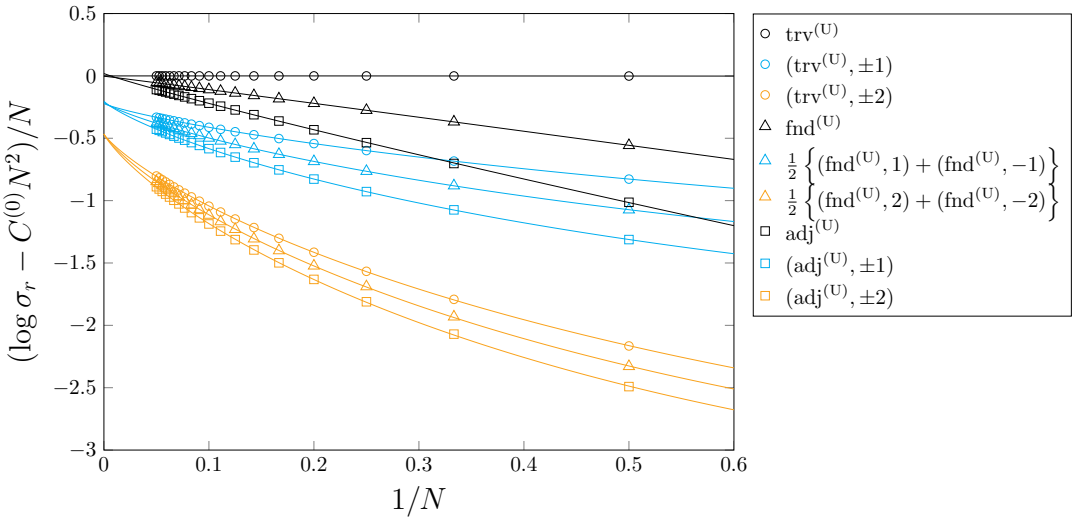

Figure 5.6: The quantity $(\log \sigma_r - C^{(0)}N^2)/N$ for $\mathrm{U}(N)$ theory with $\lambda = 3$. The y-intercept corresponds to the coefficient a_r in (5.4.1). Note that representations with the same charges (or the absolute values) have the same coefficient a_r.

One important consequence of this equivalence is a new understanding of the volume independence of 2D gauge theories, known as the Eguchi-Kawai reduction [136]. To see this, let us write the free energy explicitly with the large-N expansion (5.4.1)

$$F \equiv \frac{1}{N^2 V} \log Z = C^{(0)} + \frac{1}{VN^2} \log \left(\sum_r e^{V C_r^{(1)}} \right) + \cdots \tag{5.4.3}$$

This quantity is independent of volume in the large-N limit. And sice most other important quantities such as Wilson loops are derivatives of the free energy, they are also volume independent.

In Fig. 5.7, we plot the free energy of different theories as a function of $1/N^2$. All of these are obtained with $\lambda = 3$ and $D_{\mathrm{cut}} = 64$. We have clearly demonstrated that the free energy at large N is independent of volumes.

It can be understood from this model that the Eguchi-Kawai reduction is a result of a simple volume scaling of the singular values, namely

$$\sigma_r(V) = \sigma_r(1)^V. \tag{5.4.4}$$

This property is not expected to be true in general. However, it is possible that such a nice scaling behavior only appears if the volume is sufficiently large, which is expected to be what happens in gauge theories in higher dimensions.

5.5 θ term

In this section, we discuss how the TRG works when we have a θ term. The partition function with nonzero can also be exactly solved with nonzero θ. And the derivation for

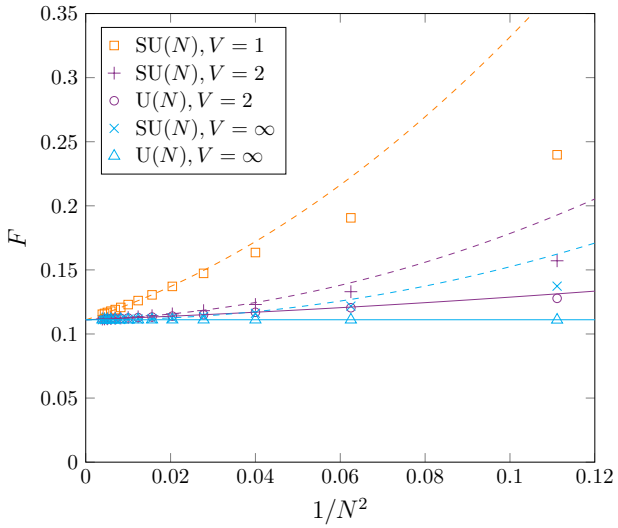

Figure 5.7: Free energy from different theories and volumes as a function of $1/N^2$ at $\lambda = 3$. They all approach the same value in the large-N limit.

the fundamental tensor also follows in the same way. In the continuum limit, the θ term reads

$$S_\theta = -i\theta Q; \tag{5.5.1}$$

$$Q = \frac{1}{4\pi}\int d^2x \epsilon_{\mu\nu} \mathrm{tr} F_{\mu\nu}. \tag{5.5.2}$$

Although it is possible to use any definition of the topological charge, we will use the 'log definition' of the topological charge when we discretize it on a lattice:

$$Q = \frac{1}{2\pi i}\sum_n \log\det P_n. \tag{5.5.3}$$

Adding this to the action, one obtains the partition function

$$Z(\theta) = \sum_r \left(\frac{\tilde{\gamma}_r(\theta)}{d_r}\right)^{L_1 L_2}; \tag{5.5.4}$$

$$\tilde{\gamma}_r(\theta) = \det \mathcal{M}_r(\theta), \tag{5.5.5}$$

$$(\mathcal{M}_r(\theta))_{ij} = \int_{-\pi}^{+\pi} \frac{d\phi}{2\pi} \cos\{(l_j + i - j + \tfrac{\theta}{2\pi})\phi\} \exp(\tfrac{2N}{\lambda}\cos\phi). \tag{5.5.6}$$

The corresponding singular values are

$$\sigma_r(\theta) = \left|\frac{\tilde{\gamma}_r(\theta)}{d_r}\right|. \tag{5.5.7}$$

From this expression, one can easily show the properties

$$\sigma_{(r^{(\mathrm{U})},q)}(\theta + 2\pi) = \sigma_{(r^{(\mathrm{U})},q+1)}(\theta), \tag{5.5.8}$$

$$\sigma_{r^{(\mathrm{U})}}(-\theta) = \sigma_{\bar{r}^{(\mathrm{U})}}(\theta). \tag{5.5.9}$$

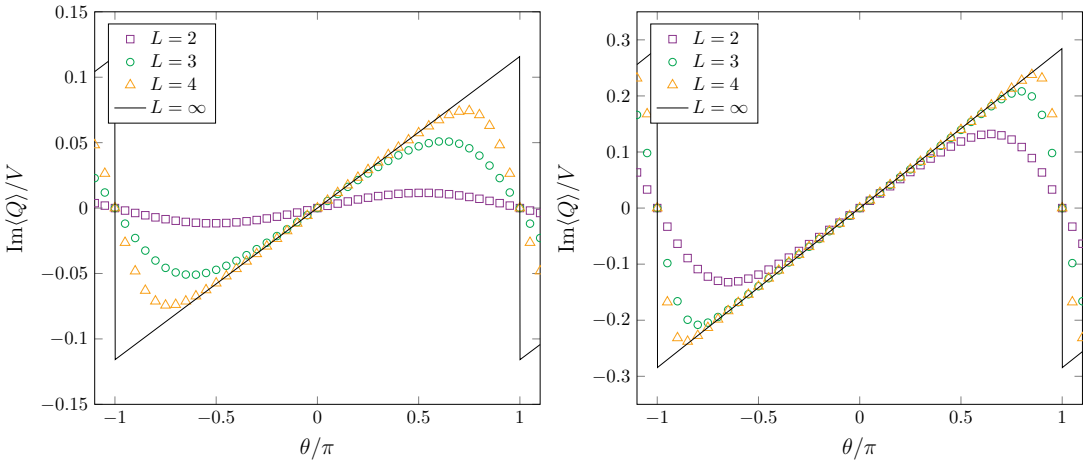

Figure 5.8: The imaginary part of the topological charge for U(3) gauge theory with $\lambda = 1.5$ (Left) and $\lambda = 3$ (Right). The gap is developed at $\theta = \pm\pi$ as we increase the system dimension L.

Namely, the singular values are not guaranteed to be neither periodic in θ and invariant under $\theta \to -\theta$.

One physically interesting point about gauge theory with a θ term is the existence of the first-order phase transition at $\theta = \pi$. An important observable is the topological charge density:

$$\frac{1}{V}\mathrm{Im}\langle Q\rangle = -\frac{1}{V}\frac{\partial}{\partial\theta}\log Z(\theta) \tag{5.5.10}$$

which develops a gap at $\theta \sim \pm\pi$ as the volume V increases.

The next question is, what will happen to this gap in the large N limit? To answer this, note that the topological susceptibility at $\theta = 0$ can be exactly evaluated at large N:

$$\chi_N = \begin{cases} -\dfrac{1}{4\pi^2}\log\left(1-\dfrac{\lambda}{2}\right) & ;\lambda < 2, \\ \dfrac{1}{2\pi^2}\left\{\log N + \log\left(1-\dfrac{2}{\lambda}\right) + \gamma_{\mathrm{E}} + 1\right\} & ;\lambda > 2, \end{cases} \tag{5.5.11}$$

where $\gamma_{\mathrm{E}} \approx 0.5772$ is the Euler constant. This is nothing but the slope at $\theta = 0$ in Fig. 5.8 in the large-N limit. But as one can see that χ_N grows like $\mathcal{O}(\log N)$ in the strong coupling region, it is instructive to consider the quantity $\mathrm{Im}\langle Q\rangle/(V\chi_N)$ instead. The result is shown in Fig. 5.9. Interestingly, the behavior with $\lambda < 2$ and $\lambda > 2$ are totally different. For $\lambda < 2$, the shape of the plot already approaches the large-N limit. And the gaps at $\theta = \pm\pi$ do not develop. However, for $\lambda > 2$, the large-N limit even at $L = 2$, is also approaching the infinite volume limit. Namely, the gaps are developed at $\theta = \pm\pi$. This is a new kind of large-N volume independence that is only particular to the strong coupling region.

To understand this behaviour, it is intructive to look at the singular values at different θ individually. Let us first list what we know so far about the effect of the θ term on some of the singular values. Since it is known that the trivial representation dominates all other

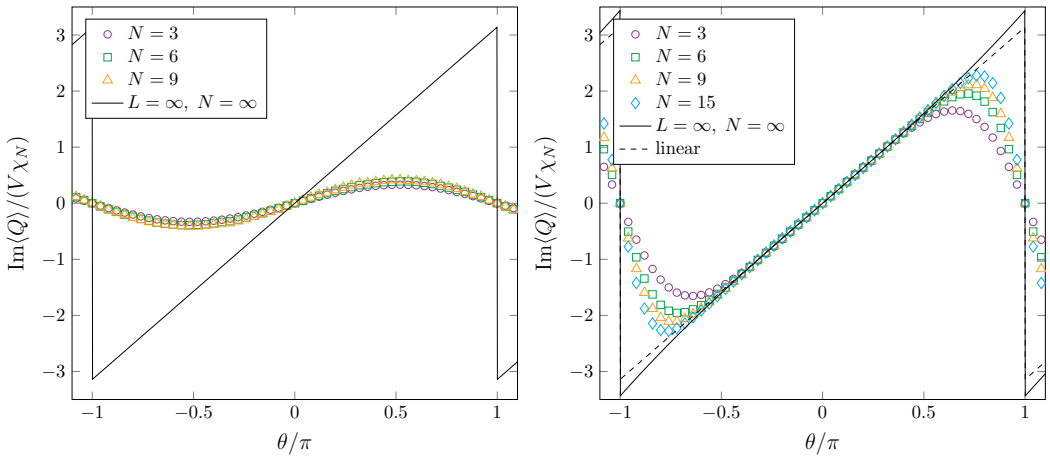

Figure 5.9: The imaginary part of the topological charge for different gauge theories at $L = 2$ with $\lambda = 1.5$ (Left) and $\lambda = 3$ (Right). The limit $L \to \infty$ and $N \to \infty$ is also shown as a solid line.

representation in the partition function at $\theta = 0$ and large N (see Fig. 5.6), we can estimate the large-N partition function by $\sigma^V_{\mathrm{trv}}(\theta = 0)$, thus giving the topological charge at $\theta = 0$

$$\frac{1}{V}\mathrm{Im}\langle Q\rangle\Big|_{\theta=0} = -\frac{1}{V}\frac{\partial}{\partial\theta}\log Z(\theta)\Big|_{\theta=0} \sim -\frac{\partial}{\partial\theta}\log\sigma_{\mathrm{trv}}(\theta)\Big|_{\theta=0} \tag{5.5.12}$$

which is known to be zero, and the topological susceptibility

$$\chi = -\frac{1}{V}\frac{\partial^2}{\partial\theta^2}\log Z\Big|_{\theta=0} \sim \frac{1}{N^2}\frac{\partial^2}{\partial\theta^2}\log\sigma_{\mathrm{trv}}(\theta)\Big|_{\theta=0}, \tag{5.5.13}$$

which is known to be χ_N. With this information, we can write the general large-N asymptotic form of the trivial representation as a function of θ as

$$\log\sigma_{\mathrm{trv}}(\theta) = \log\sigma_{\mathrm{trv}}(0) + \chi_N f(\theta) + \mathcal{O}(1/N), \tag{5.5.14}$$

where $f(\theta) \sim -\frac{1}{2}\theta^2$ at small θ. We plot the quantity $\log\{\sigma_r(\theta)/\sigma_r(0)\}/\chi_N$ in Fig. 5.10 (top row) for different representations in both weak and strong coupling regions. Perhaps surprisingly, the function $f(\theta)$ is common for all representation. Although the function for both coupling regions are totally different. In Fig. 5.10 (bottom row), we show the large-N limit of the function $f(\theta)$ of the trivial representation. In the weak coupling region, the function converge to $-\frac{1}{2}\theta^2$, while in the strong coupling region, it converges to some function that goes to negative infinity at $\theta = 2q\pi$ for $q \in \mathbb{Z}/\{0\}$. This gives us the incentive to propose the general large-N asymptotic form of the singular value

$$\log\sigma_{r^{(\mathrm{U})}}(\theta) = \log\sigma_{r^{(\mathrm{U})}}(0) + \chi_N f(\theta) + \mathcal{O}(1/N). \tag{5.5.15}$$

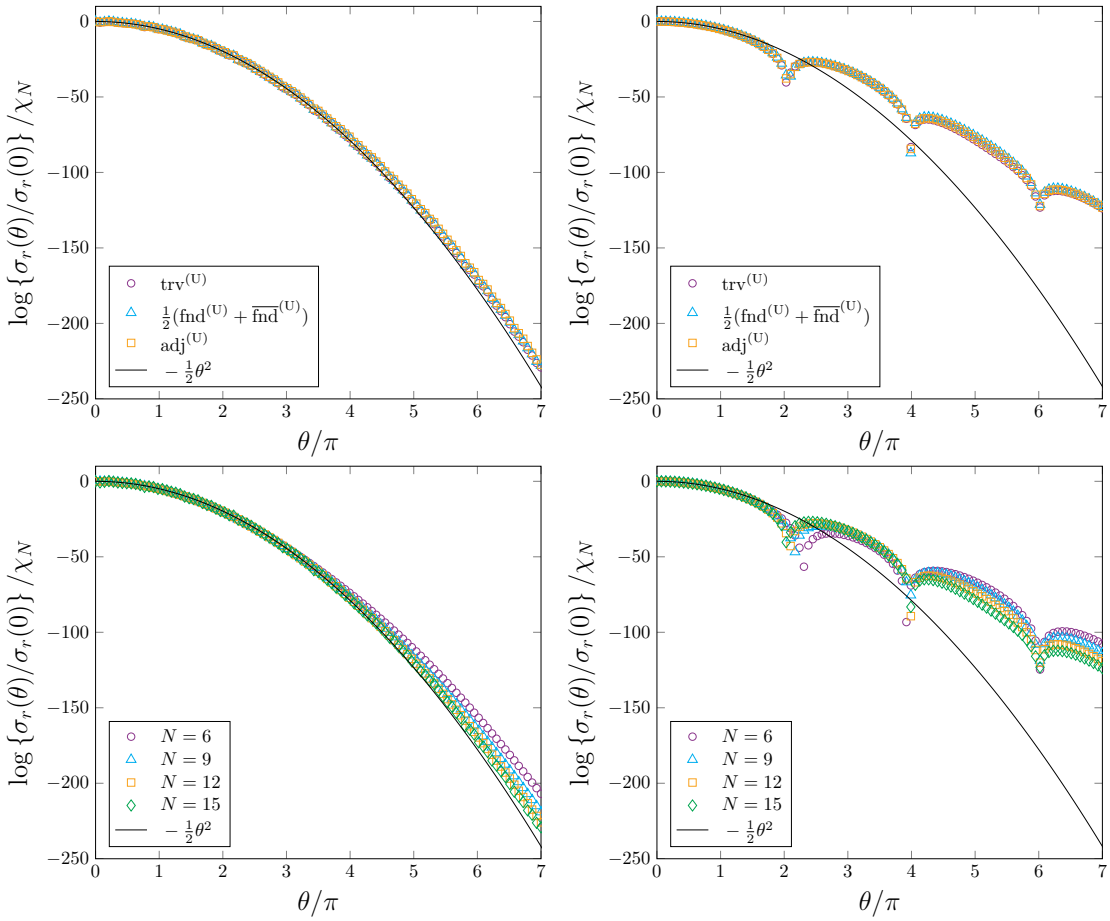

Figure 5.10: (Top row) The quantity $\log\{\sigma_r(\theta)/\sigma_r(0)\}/\chi_N$ is plotted against θ for different representations. The curve $-\frac{1}{2}\theta^2$. This suggests that this quantity is independent of the representation r in both the weak coupling (Left) and strong coupling region (Right). (Bottom row) The same plot for the trivial representation with increasing N. A clear large-N scaling behaviour is seen, which can also be similarly observed for other representations.

There are several things that can be inferred from this fact. Firstly, if $f(\theta)$ is negatively large, it means that the singular values will vanish at that point. Recall that the singular value with $\theta = 2\pi n$ is related to the one with $\theta = 0$ via

$$\sigma_{(r^{(\mathrm{U})},0)}(2q\pi) = \sigma_{(r^{(\mathrm{U})},q)}(0). \tag{5.5.16}$$

This relation explains why the asymptotic form for $\mathrm{U}(N)$ singular values are given by (5.4.1). Also, this suggests that $\mathrm{U}(N)$ singular values with nonzero charge q will all vanish in the strong coupling region and $\theta = 0$.

Another important consequence of this is the fact that we have a new volume independence in the strong coupling region as observed in Fig. 5.8. To see this, let us rewrite (5.5.15) as

$$\sigma_{r^{(\mathrm{U})}}(\theta) = \sigma_{r^{(\mathrm{U})}}(0) e^{\chi_N f(\theta) + \mathcal{O}(1/N)}. \tag{5.5.17}$$

Since $f(\theta) < 0$, $f(\theta) > f(\theta + 2q\pi)$ for $q \neq 0$ and $|\theta| < \pi$, and χ_N grows like $\mathcal{O}(\log N)$ for $\lambda > 2$, we can conclude that only the representations with $q = 0$ dominate other charges in the region $|\theta| < \pi$. Therefore, we have

$$\frac{1}{V}\log Z = C^{(0)}N^2 + \chi_N f(\theta) + \frac{1}{V}\log\left(\sum_{r^{(\mathrm{U})}} e^{VC^{(1)}_{r^{(\mathrm{U})}}}\right) + \cdots. \tag{5.5.18}$$

It should be stressed once again that the term $\chi_N\varphi(\theta)$ can only be pulled out of the logarithm only because the representations with $q = 0$ dominates. Taking the derivative of this quantity with respect to θ gives the topolgical charge

$$\frac{1}{V}\mathrm{Im}\langle Q\rangle = -\chi_N \frac{d}{d\theta} f(\theta) + \cdots, \tag{5.5.19}$$

which is independent of volume at large N.

5.6 Brief summary

In this chapter, I explain the tensor renormalization group method and its application to 2D gauge theories. I first give an overview of a general TRG method in section 5.1. I then explain three methods to construct the initial tensor in the case of 2D gauge theories in section 5.2. In our work, we study the character expansion construction, which requires an appropriate way to strict the number of representations. This process is described in 5.3. In section 5.4, I discuss the behavior of singular values of the tensor, especially in the large-N limit. A new way to understand volume independence through nice volume-scaling of singular values is also given. In section 5.5, I explain how the θ term affects the singular values and what happens in the large-N limit. Intriguingly, we find after analyzing the singular values at finite θ that there is a new kind of volume independence at strong couplings.

Chapter 6

Summary and Outlook

In Chapter 2, we briefly review two of the systems that suffer from the sign problem. The first of the two is the gauge theory with a θ term, which is interesting for 4d SU(2) theory since there are contradicting predictions regarding the point at which CP symmetry is restored at $\theta = \pi$. But since the sign problem for 4D and 2D are similar, we are particularly interested in the 2D gauge theories as the starting point. The second system is the real-time path integral formulation of quantum mechanics. In particular, the tunneling problem from the viewpoint of Picard-Lefschetz theory. Starting from the instanton description by Coleman, we discussed some important works which addressed the dominant path in the real-time path integral, some of which arrive at different conclusions.

Chapter 3 is devoted to the application of the CLM to the 2D U(1) gauge theory. We firstly discuss the justification of the method. Then we show that naive implementation will not work due to the large-drift problem and the topology freezing, which are actually a trade-off. We then demonstrate that these two problems can be simultaneously solved by introducing a topological defect on the lattice. Despite the fact that this defect heavily destroys the unitarity of link variables, we show that it is not of any physical consequence.

In Chapter 4, we discuss how to implement the HMC algorithm for thimble simulation. Recent developments of the algorithm are reviewed. We also discuss the new flow equation whose flowing modes are modified to flow at similar rates via a preconditioner matrix. And finally, some results of the simulation for real-time quantum mechanics are presented. We show that the time evolution of the wavefunction can be computed without any problem, even with large degrees of freedom. We also discuss the double-well problem and show that the dominant trajectories for real time (and also for any other Wick rotation) are actually regular and not a complex spiral trajectory, at least for a low barrier.

And finally, we discuss the TRG method in Chapter 5. Three of the ways to construct the fundamental tensor for 2D gauge theories are discussed. Of our particular interest is the character expansion construction. We present a way to efficiently cut off the number of representations during the construction of the tensor. Such construction is expected to be useful in general theories. We then discuss some properties of singular values at large N and finite θ. One of the remarkable results is that the large-N volume independence can be understood as a consequence of a nice scaling of singular values.

In this thesis, I discuss three of the most popular methods to handle the infamous sign problem, which has been hindering us from studying many interesting problems. To that end, the developments of different tools are necessary. By applying the methods to simpler problems, we can understand the advantages and limitations of each of the techniques in more detail.

Firstly, the complex Langevin method can be used as the first method to try solving a system with complex action. Its implementation is very simple and requires a minimal number of fine-tunings. As long as the computational resources allow, one can also approach the continuum or thermodynamic limits without much worry. Although we cannot hope for the method to explore all parameter regions due to the wrong convergence problem, we can use it to gain important insights from these limited parameter regions (See, e.g., Ref. [50–54] for QCD and Ref. [137, 138] for the type IIB matrix model).

On the other hand, the generalized Lefschetz thimble method can be useful in scenarios where the CLM breaks down. It should be noted, however, that various modifications to the algorithm (e.g., the worldvolume approach, time-dependent HMC mass, or preconditioning) and in some cases a very careful fine-tuning (e.g., W-potential and HMC parameters) might be necessary in order for the simulation to be possible. Because big progress in the field only happened relatively recently, there have not been many implementations of the method toward realistic systems. Our work shows that it is at least possible to study one-dimensional real-time quantum mechanics. This system suffers from a severe sign problem due to the pure imaginary action, which even the complex Langevin method cannot handle. It is still important to note that the method can still suffer from the residue sign problem if the phase of the Jacobian becomes highly fluctuating. This is expected to be the case when the number of degrees of freedom becomes large, such as in the continuum or thermodynamic limits. If one chooses the system carefully, it is possible to investigate interesting results before such a problem happens.

And finally, the tensor renormalization group method originally proved to be useful in condensed matter systems, which are usually of lower dimensions. The method is completely free of the sign problem by construction and can access large volumes very easily. If the system of interest has a nice singular-value profile, the TRG method is undoubtedly the best method. For example, we can easily study the first-order phase transition at $\theta = \pi$ and can observe its critical behavior as we approach infinite-volume limit [128]. This is, however, impossible with the CLM because we need to modify the topology of the lattice. Such manipulation spoils θ-periodicity and thus renders the first-order phase transition at $\theta = \pi$ invisible. Another big advantage of the method is that it can handle the fermionic degrees of freedom without the need to integrate them out first. This not only simplifies the process, but also reduces the computational cost. One smaller concern is that the technique is not as flexible as the other two methods since the construction of the tensor network is problem-specific and it is not straightforward to compute observables other than the free energy and its derivatives, albeit possible. However, the biggest concern of the method is that the cost of performing coarse graining can grow exponentially with the number of dimensions. Recent studies show that it is possible to study simple gauge theories in higher dimensions [130, 139]. It is thus still promising to find an optimized version of TRG that

let us study more realistic systems.

Our studies of the three methods so far show that the sign problem is not an insurmountable obstacle that we must give up. With the right tools, it is possible to extract important information that is otherwise inaccessible in non-perturbative physics. I am looking forward to applying these tools to more realistic problems in the future.

Acknowledgement

This thesis is based on collaborative research with G. Fujisawa, K. Hatakeyama, M. Hirasawa, M. Honda, Y. Ito, A. Matsumoto, J. Nishimura, and K. Sakai. Without insightful discussions with them, I would not be able to learn such a wide range of knowledge. I am especially thankful to my supervisor, J. Nishimura, for invaluable support and advice during my time as a Ph.D. student at SOKENDAI.

I would like to thank Y. Asano, M. Fukuma, S. Hashimoto, S. Iso, D. Kadoh, R. Kitano, T. Mori, and N. Yamada, for valuable discussions. I also would like to express my appreciation to my family, my girlfriend Nuntabhorn, my friends who always keep in touch, and S. Hoshimachi, for giving me mental support.

The computations in these works were carried out on the PC clusters in KEK Computing Research Center and KEK Theory Center. I am supported by the Japanese government under the MEXT scholarship.

Appendix A

Derivation of the Fokker-Planck equation for the Langevin process

In this section, we will show that the configuration governed by the Langevin equation has a probability distribution satisfying the Fokker-Planck equation. Let us start with the discretized Langevin equation

$$x(n+1) = x(n) - \epsilon \frac{\partial S}{\partial x}(n) + \sqrt{\epsilon}\eta_\epsilon(n). \tag{A.0.1}$$

where ϵ is the time interval and $n = t/\epsilon$. Note that the Gaussian noise η_ϵ is now normalized by the correlation

$$\langle \eta_\epsilon(n)\eta_\epsilon(m)\rangle = 2\delta_{n,m} \tag{A.0.2}$$

which is related to the continuum definition via

$$\eta_\epsilon(n) = \sqrt{\epsilon}\eta(t). \tag{A.0.3}$$

Two important identities that will be used in the derivation are

$$\frac{1}{\epsilon}\langle x(n+1) - x(n)\rangle_{\eta_\epsilon(n)} = -\frac{\partial S}{\partial x}(n), \tag{A.0.4}$$

$$\frac{1}{\epsilon}\left\langle (x(n+1) - x(n))^2 \right\rangle_{\eta_\epsilon(n)} = 2, \tag{A.0.5}$$

where $\langle \cdot \rangle_{\eta_\epsilon(n)}$ stands for an expectation value with respect to the noise $\eta_\epsilon(n)$. Higher order moments are proportional to $\sqrt{\epsilon}$ or smaller and will not be considered.

Now to the derivation. Consider the following integral

$$\int d^N x h(x)\dot{P}(x;n) \tag{A.0.6}$$

where $h(x)$ is some arbitrary function, and $\dot{P}$ is the time derivative of P. Using the Chapman-Kolmogorov equation 1.2.4, this quantity can be rewritten as

$$\begin{aligned}\int d^N x h(x)\dot{P}(x;n) &= \int d^N x h(x)\frac{P(x;n+1) - P(x;n)}{\epsilon} \\ &= \frac{1}{\epsilon}\left\{ \int d^N x h(x) \int d^N y P(x|y)P(y;n) - \int d^N x h(x)P(x;n)\right\}. \end{aligned} \tag{A.0.7}$$

Interchanging the order of the integral of the first term and swap the variable $x \leftrightarrow y$, as well as using the identity $\int d^N y P(y|x) = 1$ gives

$$\int d^N x h(x) \dot{P}(x;n) = \frac{1}{\epsilon} \int d^N x P(x;n) \int d^N y \left\{ P(y|x)(h(y) - h(x)) \right\}. \tag{A.0.8}$$

Since $P(y|x)$ is highly localized near the point $y = x$, we can perform Taylor expansion[1] of y around x;

$$\int d^N x h(x) \dot{P}(x;n) = \frac{1}{\epsilon} \int d^N x P(x;n) \int d^N y \left\{ P(y|x) \sum_{k=1}^{\infty} \frac{1}{k!} (y-x)^k h^{(k)}(x) \right\}. \tag{A.0.9}$$

Next, we define the jump moments as

$$\mu_{(k)}(x) = \frac{1}{k!\epsilon} \int d^N y P(y|x)(y-x)^k = \frac{1}{k!\epsilon} \langle (x(n+1) - x(n))^k \rangle_{\eta_\epsilon}, \tag{A.0.10}$$

which we already know the first two:

$$\mu_{(1)}(x) = -\frac{\partial S}{\partial x}, \tag{A.0.11}$$

$$\mu_{(2)}(x) = 1, \tag{A.0.12}$$

and the rest vanishes with $\epsilon \to 0$. We then have

$$\int d^N x h(x) \dot{P}(x;n) = \int d^N x P(x;n) \sum_{k=1}^{\infty} \mu_{(k)}(x) h^{(k)}(x). \tag{A.0.13}$$

Moving everything to the same side and performing partial integration gives

$$0 = \int d^N x h(x) \left\{ \dot{P}(x;n) + \sum_{k=1}^{\infty} (-)^{k-1} \partial^k (P(x;n) \mu_{(k)}(x)) \right\}. \tag{A.0.14}$$

Since $h(x)$ is an arbitrary function, we can set the quantity in the curly bracket to zero;

$$\dot{P}(x;n) = \sum_{k=1}^{\infty} (-)^k \partial^k (P(x;n) \mu_{(k)}(x)) = \frac{\partial}{\partial x} \left(P \frac{\partial S}{\partial x} + \frac{\partial P}{\partial x} \right). \tag{A.0.15}$$

—Q.E.D.—

[1] Remember that since $h(x)$ is a multivariate function, $(y-x)^k h^{(k)}(x)$ actually means $(y-x)_{i_1} \cdots (y-x)_{i_k} (\partial_{i_1} \cdots \partial_{i_k} h(x))$.

Appendix B

Justification of the preconditioned flow equation

It is important that the modification of the flow equation does not invalidate the equivalence between the original integral and the integral on the flowed contour. For that, we have to show that the flow contour will not cross any poles of the Boltzmann weight. Equivalently, we want to show that the flow vector $(\bar{H}H)^{-\frac{1}{2}}\overline{\partial S}$ always points away from the pole $z_\star$:

$$\lim_{z\to z_\star} \mathrm{Re}\Big[(z-z_\star)\cdot(H\bar{H})^{-\frac{1}{2}}\cdot\partial S\Big] \geq 0. \tag{B.0.1}$$

First, recall that Cauchy's integral formula for multivariate function is

$$f(\vec{u}) = \frac{1}{(2\pi i)^n}\int_{\mathcal{M}} d^n z f(\vec{z})\prod_a (z-u)_a^{-1}. \tag{B.0.2}$$

where $\mathcal{M}$ is a direct product of counter-clockwise circles around each component of u. This implies that pole expansion of a multivariate function is of the form

$$f(\vec{z}) \sim f_\star \prod_a (z-z_\star)_a^{-1}. \tag{B.0.3}$$

Thusly, we rewrite the Boltzmann weight around its pole as

$$\exp\left[-S(z)\right] \approx r_\star \prod_a (z-z_\star)_a^{-1} \tag{B.0.4}$$

where $r_\star$ is the residue of the Boltzmann weight at $z_\star$. This implies that

$$S(z) \approx -\log r_\star + \sum_a \log(z-z_\star)_a, \tag{B.0.5}$$

$$\partial_a S(z) \approx (z-z_\star)_a^{-1}, \tag{B.0.6}$$

$$\partial_a\partial_b S(z) \approx -(z-z_\star)_a^{-2}\delta_{ab}, \tag{B.0.7}$$

$$(H\bar{H})_{ab}^{-\frac{1}{2}} \approx |(z-z_\star)_a|^2\delta_{ab}. \tag{B.0.8}$$

Therefore, we have

$$\begin{aligned}
&\lim_{z\to z_\star} \mathrm{Re}\Big[(z-z_\star)\cdot(H\bar{H})^{-\frac{1}{2}}\cdot\partial S\Big] \\
&\quad = \lim_{z\to z_\star} \mathrm{Re}\sum_{a,b}\Big[(z-z_\star)_a\cdot\Big(|(z-z_\star)_a|^2\delta_{ab}\Big)\cdot(z-z_\star)_b^{-1}\Big] \\
&\quad = \lim_{z\to z_\star} \mathrm{Re}|z-z_\star|^2 \geq 0.
\end{aligned}$$

—Q.E.D.—

The same conclusion can also be reached even for the case of rational approximation

$$A = c_0\mathbb{1} + \sum_{q=1}^{Q} c_q(\bar{H}H + m_q\mathbb{1})^{-1} \tag{B.0.9}$$

because H is diagonal, making $\bar{H}H$, and therefore A, a positive-definite diagonal matrix.

Bibliography

[1] D. Hanneke, S.F. Hoogerheide and G. Gabrielse, *Cavity Control of a Single-Electron Quantum Cyclotron: Measuring the Electron Magnetic Moment*, *Phys. Rev. A* **83** (2011) 052122 [1009.4831].

[2] T. Aoyama, M. Hayakawa, T. Kinoshita and M. Nio, *Tenth-Order Electron Anomalous Magnetic Moment — Contribution of Diagrams without Closed Lepton Loops*, *Phys. Rev. D* **91** (2015) 033006 [1412.8284].

[3] B. Lucini, M. Teper and U. Wenger, *Glueballs and k-strings in SU(N) gauge theories: Calculations with improved operators*, *JHEP* **06** (2004) 012 [hep-lat/0404008].

[4] Y. Chen et al., *Glueball spectrum and matrix elements on anisotropic lattices*, *Phys. Rev. D* **73** (2006) 014516 [hep-lat/0510074].

[5] E. Witten, *String theory dynamics in various dimensions*, *Nucl. Phys. B* **443** (1995) 85 [hep-th/9503124].

[6] J. Polchinski, *Dirichlet Branes and Ramond-Ramond charges*, *Phys. Rev. Lett.* **75** (1995) 4724 [hep-th/9510017].

[7] P. Horava and E. Witten, *Heterotic and type I string dynamics from eleven-dimensions*, *Nucl. Phys. B* **460** (1996) 506 [hep-th/9510209].

[8] J. Polchinski, *M theory and the light cone*, *Prog. Theor. Phys. Suppl.* **134** (1999) 158 [hep-th/9903165].

[9] T. Banks, W. Fischler, S.H. Shenker and L. Susskind, *M theory as a matrix model: A Conjecture*, *Phys. Rev. D* **55** (1997) 5112 [hep-th/9610043].

[10] D.E. Berenstein, J.M. Maldacena and H.S. Nastase, *Strings in flat space and pp waves from N=4 superYang-Mills*, *JHEP* **04** (2002) 013 [hep-th/0202021].

[11] N. Ishibashi, H. Kawai, Y. Kitazawa and A. Tsuchiya, *A Large N reduced model as superstring*, *Nucl. Phys. B* **498** (1997) 467 [hep-th/9612115].

[12] R. Healey, *Gauge Theory and the Theta Vacuum*, (2007) .

[13] E. Witten, *Topological Quantum Field Theory*, *Commun. Math. Phys.* **117** (1988) 353.

[14] J. Davighi, B. Gripaios and O. Randal-Williams, *Differential cohomology and topological actions in physics*, 2011.05768.

[15] F.J. Wegner, *Duality in Generalized Ising Models and Phase Transitions Without Local Order Parameters*, *J. Math. Phys.* **12** (1971) 2259.

[16] K.G. Wilson, *Confinement of Quarks*, *Phys. Rev. D* **10** (1974) 2445.

[17] M. Cardoso, N. Cardoso and P. Bicudo, *Lattice QCD computation of the colour fields for the static hybrid quark-gluon-antiquark system, and microscopic study of the Casimir scaling*, *Phys. Rev. D* **81** (2010) 034504 [0912.3181].

[18] J.B. Kogut, *An Introduction to Lattice Gauge Theory and Spin Systems*, *Rev. Mod. Phys.* **51** (1979) 659.

[19] A. Joseph, *Review of Lattice Supersymmetry and Gauge-Gravity Duality*, *Int. J. Mod. Phys. A* **30** (2015) 1530054 [1509.01440].

[20] M. Hanada, *What lattice theorists can do for superstring/M-theory*, *Int. J. Mod. Phys. A* **31** (2016) 1643006 [1604.05421].

[21] V.G. Filev and D. O'Connor, *The BFSS model on the lattice*, *JHEP* **05** (2016) 167 [1506.01366].

[22] M. Hanada and P. Romatschke, *Lattice Simulations of 10d Yang-Mills toroidally compactified to 1d, 2d and 4d*, *Phys. Rev. D* **96** (2017) 094502 [1612.06395].

[23] H. Gharibyan, M. Hanada, M. Honda and J. Liu, *Toward simulating superstring/M-theory on a quantum computer*, *JHEP* **07** (2021) 140 [2011.06573].

[24] MCSMC collaboration, *Confinement/deconfinement transition in the D0-brane matrix model – A signature of M-theory?*, *JHEP* **05** (2022) 096 [2110.01312].

[25] C. Robert and G. Casella, *Monte Carlo Statistical Methods (2004)*, .

[26] K. Nagata, J. Nishimura and S. Shimasaki, *Justification of the complex Langevin method with the gauge cooling procedure*, *PTEP* **2016** (2016) 013B01 [1508.02377].

[27] W.K. Hastings, *Monte Carlo sampling methods using Markov chains and their applications*, *Biometrika* **57** (1970) 97 [https://academic.oup.com/biomet/article-pdf/57/1/97/23940249/57-1-97.pdf].

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

[29] X.-Q. Luo, E.B. Gregory, S.-H. Guo and H. Kroger, *QCD at finite density*, in *International Workshop on Nonperturbative Methods and Lattice QCD*, pp. 138–149, 11, 2000, DOI [hep-ph/0011120].

[30] K.N. Anagnostopoulos, M. Hanada, J. Nishimura and S. Takeuchi, *Monte Carlo studies of supersymmetric matrix quantum mechanics with sixteen supercharges at finite temperature*, *Phys. Rev. Lett.* **100** (2008) 021601 [0707.4454].

[31] S. Catterall and T. Wiseman, *Black hole thermodynamics from simulations of lattice Yang-Mills theory*, *Phys. Rev. D* **78** (2008) 041502 [0803.4273].

[32] M. Hanada, Y. Hyakutake, J. Nishimura and S. Takeuchi, *Higher derivative corrections to black hole thermodynamics from supersymmetric matrix quantum mechanics*, *Phys. Rev. Lett.* **102** (2009) 191602 [0811.3102].

[33] M. Hanada, Y. Matsuo and N. Yamamoto, *Sign problem and phase quenching in finite-density QCD: models, holography, and lattice*, *Phys. Rev. D* **86** (2012) 074510 [1205.1030].

[34] D. Kadoh and S. Kamata, *Gauge/gravity duality and lattice simulations of one dimensional SYM with sixteen supercharges*, 1503.08499.

[35] E. Berkowitz, E. Rinaldi, M. Hanada, G. Ishiki, S. Shimasaki and P. Vranas, *Precision lattice test of the gauge/gravity duality at large-N*, *Phys. Rev. D* **94** (2016) 094501 [1606.04951].

[36] T.D.. Cohen, *Functional integrals for QCD at nonzero chemical potential and zero density*, *Phys. Rev. Lett.* **91** (2003) 222001 [hep-ph/0307089].

[37] G. Parisi, *On complex probabilities*, *Phys. Lett. B* **131** (1983) 393.

[38] J.R. Klauder, *Coherent State Langevin Equations for Canonical Quantum Systems With Applications to the Quantized Hall Effect*, *Phys. Rev. A* **29** (1984) 2036.

[39] J.R. Klauder and W.P. Peterson, *Spectrum of certain non-self-adjoint operators and solutions of Langevin equations with complex drift*, *J Stat Phys* **39** (1985) 53.

[40] E. Gozzi, *Langevin Simulation in Minkowski Space*, *Phys. Lett. B* **150** (1985) 119.

[41] H.W. Hamber and H.-c. Ren, *Complex Probabilities and the Langevin Equation*, *Phys. Lett. B* **159** (1985) 330.

[42] J. Ambjorn, M. Flensburg and C. Peterson, *Langevin Simulations of Configurations With Static Charges*, *Phys. Lett. B* **159** (1985) 335.

[43] G. Aarts, E. Seiler and I.-O. Stamatescu, *The Complex Langevin method: When can it be trusted?*, *Phys. Rev. D* **81** (2010) 054508 [0912.3360].

[44] G. Aarts, F.A. James, E. Seiler and I.-O. Stamatescu, *Complex Langevin: Etiology and Diagnostics of its Main Problem*, *Eur. Phys. J. C* **71** (2011) 1756 [1101.3270].

[45] M. Scherzer, E. Seiler, D. Sexty and I.O. Stamatescu, *Controlling Complex Langevin simulations of lattice models by boundary term analysis*, *Phys. Rev. D* **101** (2020) 014501 [1910.09427].

[46] M. Scherzer, E. Seiler, D. Sexty and I.-O. Stamatescu, *Complex Langevin and boundary terms*, *Phys. Rev. D* **99** (2019) 014512 [1808.05187].

[47] K. Nagata, J. Nishimura and S. Shimasaki, *Argument for justification of the complex Langevin method and the condition for correct convergence*, *Phys. Rev. D* **94** (2016) 114515 [1606.07627].

[48] E. Seiler, D. Sexty and I.-O. Stamatescu, *Gauge cooling in complex Langevin for QCD with heavy quarks*, *Phys. Lett. B* **723** (2013) 213 [1211.3709].

[49] Y. Ito and J. Nishimura, *The complex Langevin analysis of spontaneous symmetry breaking induced by complex fermion determinant*, *JHEP* **12** (2016) 009 [1609.04501].

[50] F. Attanasio, B. Jäger and F.P.G. Ziegler, *QCD equation of state via the complex Langevin method*, 2203.13144.

[51] Y. Namekawa, Y. Asano, Y. Ito, T. Kaneko, H. Matsufuru, J. Nishimura et al., *Flavor number dependence of QCD at finite density by the complex Langevin method*, *PoS* **LATTICE2021** (2022) 623 [2112.00150].

[52] S. Tsutsui, Y. Asano, Y. Ito, H. Matsufuru, Y. Namekawa, J. Nishimura et al., *Color superconductivity in a small box: a complex Langevin study*, *PoS* **LATTICE2021** (2022) 533 [2111.15095].

[53] Y. Ito, H. Matsufuru, Y. Namekawa, J. Nishimura, S. Shimasaki, A. Tsuchiya et al., *Complex Langevin calculations in QCD at finite density*, *JHEP* **10** (2020) 144 [2007.08778].

[54] M. Scherzer, D. Sexty and I.O. Stamatescu, *Deconfinement transition line with the complex Langevin equation up to $\mu/T \sim 5$*, *Phys. Rev. D* **102** (2020) 014515 [2004.05372].

[55] E. Picard and G. Simart, *Theorie des fonctions algebriques de deux variables independantes. Tome I*, (1897) .

[56] S. Lefschetz, *L'analysis situs et la geometrie algebrique*, (1924) .

[57] E. Witten, *Analytic Continuation Of Chern-Simons Theory*, *AMS/IP Stud. Adv. Math.* **50** (2011) 347 [1001.2933].

[58] E. Witten, *A New Look At The Path Integral Of Quantum Mechanics*, 1009.6032.

[59] F. Di Renzo and K. Zambello, *Solution of the Thirring model in thimble regularization*, *Phys. Rev. D* **105** (2022) 054501 [2109.02511].

[60] K. Zambello and F. Di Renzo, *On the Lefschetz thimbles structure of the Thirring model*, *PoS* **LATTICE2019** (2020) 211 [1912.11380].

[61] Y. Tanizaki and T. Koike, *Real-time Feynman path integral with Picard–Lefschetz theory and its applications to quantum tunneling*, *Annals Phys.* **351** (2014) 250 [1406.2386].

[62] W.-Y. Ai, B. Garbrecht and C. Tamarit, *Functional methods for false vacuum decay in real time*, *JHEP* **12** (2019) 095 [1905.04236].

[63] Z.-G. Mou, P.M. Saffin and A. Tranberg, *Quantum tunnelling, real-time dynamics and Picard-Lefschetz thimbles*, *JHEP* **11** (2019) 135 [1909.02488].

[64] D. Jia, *Complex, Lorentzian, and Euclidean simplicial quantum gravity: numerical methods and physical prospects*, *Class. Quant. Grav.* **39** (2022) 065002 [2110.05953].

[65] M. Levin and C.P. Nave, *Tensor renormalization group approach to 2D classical lattice models*, *Phys. Rev. Lett.* **99** (2007) 120601.

[66] Y. Shimizu and Y. Kuramashi, *Grassmann tensor renormalization group approach to one-flavor lattice Schwinger model*, *Phys. Rev. D* **90** (2014) 014508 [1403.0642].

[67] Y. Shimizu and Y. Kuramashi, *Critical behavior of the lattice Schwinger model with a topological term at $\theta = \pi$ using the Grassmann tensor renormalization group*, *Phys. Rev. D* **90** (2014) 074503 [1408.0897].

[68] S. Takeda and Y. Yoshimura, *Grassmann tensor renormalization group for the one-flavor lattice Gross–Neveu model with finite chemical potential*, *PTEP* **2015** (2015) 043B01 [1412.7855].

[69] R. Sakai, S. Takeda and Y. Yoshimura, *Higher order tensor renormalization group for relativistic fermion systems*, *PTEP* **2017** (2017) 063B07 [1705.07764].

[70] Y. Shimizu and Y. Kuramashi, *Berezinskii-Kosterlitz-Thouless transition in lattice Schwinger model with one flavor of Wilson fermion*, *Phys. Rev. D* **97** (2018) 034502 [1712.07808].

[71] D. Kadoh, Y. Kuramashi, Y. Nakamura, R. Sakai, S. Takeda and Y. Yoshimura, *Tensor network formulation for two-dimensional lattice $\mathcal{N} = 1$ Wess-Zumino model*, *JHEP* **03** (2018) 141 [1801.04183].

[72] Y. Yoshimura, Y. Kuramashi, Y. Nakamura, S. Takeda and R. Sakai, *Calculation of fermionic Green functions with Grassmann higher-order tensor renormalization group*, *Phys. Rev. D* **97** (2018) 054511 [1711.08121].

[73] N. Butt, S. Catterall, Y. Meurice, R. Sakai and J. Unmuth-Yockey, *Tensor network formulation of the massless Schwinger model with staggered fermions*, *Phys. Rev. D* **101** (2020) 094509 [1911.01285].

[74] S. Akiyama and D. Kadoh, *More about the Grassmann tensor renormalization group*, *JHEP* **10** (2021) 188 [2005.07570].

[75] S. Akiyama, Y. Kuramashi, T. Yamashita and Y. Yoshimura, *Restoration of chiral symmetry in cold and dense Nambu–Jona-Lasinio model with tensor renormalization group*, *JHEP* **01** (2021) 121 [2009.11583].

[76] S. Akiyama and Y. Kuramashi, *Tensor renormalization group approach to (1+1)-dimensional Hubbard model*, *Phys. Rev. D* **104** (2021) 014504 [2105.00372].

[77] Z.Y. Xie, J. Chen, M.P. Qin, J.W. Zhu, L.P. Yang and T. Xiang, *Coarse-graining renormalization by higher-order singular value decomposition*, *Phys. Rev. B* **86** (2012) 045139.

[78] D. Adachi, T. Okubo and S. Todo, *Anisotropic Tensor Renormalization Group*, *Phys. Rev. B* **102** (2020) 054432 [1906.02007].

[79] D. Kadoh and K. Nakayama, *Renormalization group on a triad network*, 1912.02414.

[80] U.J. Wiese, *Numerical Simulation of Lattice θ Vacua: The 2-d U(1) Gauge Theory as a Test Case*, *Nucl. Phys. B* **318** (1989) 153.

[81] B.E. Rusakov, *Loop averages and partition functions in U(N) gauge theory on two-dimensional manifolds*, *Mod. Phys. Lett. A* **5** (1990) 693.

[82] C. Bonati and P. Rossi, *Topological susceptibility of two-dimensional $U(N)$ gauge theories*, *Phys. Rev. D* **99** (2019) 054503 [1901.09830].

[83] H. Kragh, *Physics and the Totalitarian Principle*, 1907.04623.

[84] C.A. Baker et al., *An Improved experimental limit on the electric dipole moment of the neutron*, *Phys. Rev. Lett.* **97** (2006) 131801 [hep-ex/0602020].

[85] R.D. Peccei and H.R. Quinn, *CP Conservation in the Presence of Instantons*, *Phys. Rev. Lett.* **38** (1977) 1440.

[86] R.D. Peccei and H.R. Quinn, *Constraints Imposed by CP Conservation in the Presence of Instantons*, *Phys. Rev. D* **16** (1977) 1791.

[87] S. Weinberg, *A New Light Boson?*, *Phys. Rev. Lett.* **40** (1978) 223.

[88] F. Wilczek, *Problem of Strong P and T Invariance in the Presence of Instantons*, *Phys. Rev. Lett.* **40** (1978) 279.

[89] S. Coleman, *Aspects of Symmetry: Selected Erice Lectures*, Cambridge University Press (1985), 10.1017/CBO9780511565045.

[90] M. Hirasawa, A. Matsumoto, J. Nishimura and A. Yosprakob, *Tensor renormalization group and the volume independence in 2D U(N) and SU(N) gauge theories*, *JHEP* **12** (2021) 011 [2110.05800].

[91] R. Kitano, N. Yamada and M. Yamazaki, *Is $N = 2$ Large?*, *JHEP* **02** (2021) 073 [2010.08810].

[92] R. Kitano, R. Matsudo, N. Yamada and M. Yamazaki, *Peeking into the θ vacuum*, *Phys. Lett. B* **822** (2021) 136657 [2102.08784].

[93] E. Witten, *Large N Chiral Dynamics*, *Annals Phys.* **128** (1980) 363.

[94] G. 't Hooft, *Topology of the Gauge Condition and New Confinement Phases in Nonabelian Gauge Theories*, *Nucl. Phys. B* **190** (1981) 455.

[95] E. Witten, *Theta dependence in the large N limit of four-dimensional gauge theories*, *Phys. Rev. Lett.* **81** (1998) 2862 [hep-th/9807109].

[96] D.J. Gross, R.D. Pisarski and L.G. Yaffe, *QCD and Instantons at Finite Temperature*, *Rev. Mod. Phys.* **53** (1981) 43.

[97] S. Chen, K. Fukushima, H. Nishimura and Y. Tanizaki, *Deconfinement and $\mathcal{CP}$ breaking at $\theta = \pi$ in Yang-Mills theories and a novel phase for SU(2)*, *Phys. Rev. D* **102** (2020) 034020 [2006.01487].

[98] A. Parnachev and A.R. Zhitnitsky, *Phase Transitions, theta Behavior and Instantons in QCD and its Holographic Model*, *Phys. Rev. D* **78** (2008) 125002.

[99] S. Dubovsky, A. Lawrence and M.M. Roberts, *Axion monodromy in a model of holographic gluodynamics*, *JHEP* **02** (2012) 053.

[100] F. Bigazzi, A.L. Cotrone and R. Sisca, *Notes on Theta Dependence in Holographic Yang-Mills*, *JHEP* **08** (2015) 090.

[101] D. Arean, I. Iatrakis, M. Jarvinen and E. Kiritsis, *CP-odd sector and θ dynamics in holographic QCD*, *Phys. Rev. D* **96** (2017) 026001.

[102] D. Gaiotto, A. Kapustin, Z. Komargodski and N. Seiberg, *Theta, Time Reversal, and Temperature*, *JHEP* **05** (2017) 091.

[103] R. Kitano, T. Suyama and N. Yamada, *$\theta = \pi$ in $SU(N)/\mathbb{Z}_N$ gauge theories*, *JHEP* **09** (2017) 137.

[104] N. Kan, R. Kitano, S. Yankielowicz and R. Yokokura, *From 3d dualities to hadron physics*, *Phys. Rev. D* **102** (2020) 125034.

[105] M. Hirasawa, A. Matsumoto, J. Nishimura and A. Yosprakob, *Complex Langevin analysis of 2D U(1) gauge theory on a torus with a θ term*, *JHEP* **09** (2020) 023.

[106] J.C. Plefka and S. Samuel, *Monte Carlo studies of two-dimensional systems with a theta term*, *Phys. Rev. D* **56** (1997) 44 [hep-lat/9704016].

[107] C. Alexandrou, A. Athenodorou, K. Cichy, A. Dromard, E. Garcia-Ramos, K. Jansen et al., *Comparison of topological charge definitions in Lattice QCD*, *Eur. Phys. J. C* **80** (2020) 424 [1708.00696].

[108] R.P. Feynman, *Space-time approach to nonrelativistic quantum mechanics*, *Rev. Mod. Phys.* **20** (1948) 367.

[109] S. Abel and M. Spannowsky, *Observing the fate of the false vacuum with a quantum laboratory*, *PRX Quantum* **2** (2021) 010349 [2006.06003].

[110] S. Coleman, *Fate of the false vacuum: Semiclassical theory*, *Phys. Rev. D* **15** (1977) 2929.

[111] C.G. Callan and S. Coleman, *Fate of the false vacuum. ii. first quantum corrections*, *Phys. Rev. D* **16** (1977) 1762.

[112] T. Markkanen, A. Rajantie and S. Stopyra, *Cosmological Aspects of Higgs Vacuum Metastability*, *Front. Astron. Space Sci.* **5** (2018) 40 [1809.06923].

[113] M. Stone, *The Lifetime and Decay of Excited Vacuum States of a Field Theory Associated with Nonabsolute Minima of Its Effective Potential*, *Phys. Rev. D* **14** (1976) 3568.

[114] P.H. Frampton, *Vacuum Instability and Higgs Scalar Mass*, *Phys. Rev. Lett.* **37** (1976) 1378.

[115] M. Stone, *Semiclassical Methods for Unstable States*, *Phys. Lett. B* **67** (1977) 186.

[116] P.H. Frampton, *Consequences of Vacuum Instability in Quantum Field Theory*, *Phys. Rev. D* **15** (1977) 2922.

[117] A. Cherman and M. Unsal, *Real-Time Feynman Path Integral Realization of Instantons*, 1408.0012.

[118] A. Alexandru, G. Basar, P.F. Bedaque and N.C. Warrington, *Tempered transitions between thimbles*, *Phys. Rev. D* **96** (2017) 034513 [1703.02414].

[119] M. Fukuma and N. Umeda, *Parallel tempering algorithm for integration over Lefschetz thimbles*, *PTEP* **2017** (2017) 073B01 [1703.00861].

[120] M. Fukuma, N. Matsumoto and N. Umeda, *Applying the tempered Lefschetz thimble method to the Hubbard model away from half-filling*, *Phys. Rev. D* **100** (2019) 114510 [1906.04243].

[121] M. Fukuma and N. Matsumoto, *Worldvolume approach to the tempered Lefschetz thimble method*, *PTEP* **2021** (2021) 023B08.

[122] G. Fujisawa, J. Nishimura, K. Sakai and A. Yosprakob, *Backpropagating Hybrid Monte Carlo algorithm for fast Lefschetz thimble calculations*, *JHEP* **04** (2022) 179 [2112.10519].

[123] A.D. Kennedy, I. Horvath and S. Sint, *A New exact method for dynamical fermion computations with nonlocal actions*, *Nucl. Phys. B Proc. Suppl.* **73** (1999) 834 [hep-lat/9809092].

[124] M.A. Clark, A.D. Kennedy and Z. Sroczynski, *Exact 2+1 flavour RHMC simulations*, *Nucl. Phys. B Proc. Suppl.* **140** (2005) 835 [hep-lat/0409133].

[125] M.A. Clark, *The Rational Hybrid Monte Carlo Algorithm*, *PoS* **LAT2006** (2006) 004 [hep-lat/0610048].

[126] M.A. Clark and A.D. Kennedy, *Accelerating Staggered Fermion Dynamics with the Rational Hybrid Monte Carlo (RHMC) Algorithm*, *Phys. Rev. D* **75** (2007) 011502 [hep-lat/0610047].

[127] B. Jegerlehner, *Krylov space solvers for shifted linear systems*, hep-lat/9612014.

[128] Y. Kuramashi and Y. Yoshimura, *Tensor renormalization group study of two-dimensional U(1) lattice gauge theory with a θ term*, *JHEP* **04** (2020) 089.

[129] M. Fukuma, D. Kadoh and N. Matsumoto, *Tensor network approach to 2D Yang-Mills theories*, .

[130] T. Kuwahara and A. Tsuchiya, *Tensor renormalization group study of three-dimensional SU(2) gauge theory*, 2205.08883.

[131] I. Bars and F. Green, *Complete Integration of U (N) Lattice Gauge Theory in a Large N Limit*, *Phys. Rev. D* **20** (1979) 3311.

[132] I. Bars, *U(N) Integral for Generating Functional in Lattice Gauge Theory*, *J. Math. Phys.* **21** (1980) 2678.

[133] S. Samuel, *u(n) Integrals, 1/n, and the Dewit-'t Hooft Anomalies*, *J. Math. Phys.* **21** (1980) 2695.

[134] D.J. Gross and E. Witten, *Possible Third Order Phase Transition in the Large N Lattice Gauge Theory*, *Phys. Rev. D* **21** (1980) 446.

[135] S.R. Wadia, *N = Infinity Phase Transition in a Class of Exactly Soluble Model Lattice Gauge Theories*, *Phys. Lett. B* **93** (1980) 403.

[136] T. Eguchi and H. Kawai, *Reduction of Dynamical Degrees of Freedom in the Large N Gauge Theory*, *Phys. Rev. Lett.* **48** (1982) 1063.

[137] K. Hatakeyama, K. Anagnostopoulos, T. Azuma, M. Hirasawa, Y. Ito, J. Nishimura et al., *Relationship between the Euclidean and Lorentzian versions of the type IIB matrix model*, *PoS* **LATTICE2021** (2022) 341 [2112.15368].

[138] M. Hirasawa, K. Anagnostopoulos, T. Azuma, K. Hatakeyama, Y. Ito, J. Nishimura et al., *A new phase in the Lorentzian type IIB matrix model and the emergence of continuous space-time*, *PoS* **LATTICE2021** (2022) 428 [2112.15390].

[139] S. Akiyama and Y. Kuramashi, *Tensor renormalization group study of (3+1)-dimensional $\mathbb{Z}_2$ gauge-Higgs model at finite density*, *JHEP* **05** (2022) 102 [2202.10051].