Tempered transitions between thimbles

Andrei Alexandru, Gökçe Başar, Paulo F. Bedaque, and Neill C. Warrington

Phys. Rev. D **96**, 034513 — Published 18 August 2017

DOI: 10.1103/PhysRevD.96.034513

Tempered transitions between thimbles

Andrei Alexandru*
Department of Physics, George Washington University, Washington, DC 20052

Gökçe Başar†
Physics Department,University of Illinois at Chicago, Chicago, Il 60607 and
Department of Physics, University of Maryland, College Park, MD 20742

Paulo F. Bedaque‡ and Neill C. Warrington§
Department of Physics, University of Maryland, College Park, MD 20742

Quantum field theories with complex actions cannot be investigated using importance sampling due to the sign problem. One possible solution is to use the holomorphic gradient flow, a method we introduced related to the Lefschetz thimbles idea. In many cases the probability distribution generated by this method is multi-modal and standard Monte-Carlo sampling fails. We propose an algorithm that incorporates tempered proposals to solve this problem. We apply this algorithm to the Thirring $0+1$ dimensional model at finite density for a parameter set where standard sampling fails and show that tempered proposals cure this problem.

I. INTRODUCTION

Non-perturbative results from quantum field theories can be obtained using stochastic sampling, as long as the path integral can be represented as a sum over real and positive contributions. This is no longer the case when one considers phenomenologically interesting problems regarding systems at finite density, like QCD at non-zero baryon density, or questions related to real-time dynamics. In these cases, the integrand is complex and direct Monte-Carlo sampling cannot be applied. The standard work-around, *reweighting*, that samples according to a positive measure and corrects for the difference by introducing a complex fluctuating phase in the observables, fails due to phase oscillations; this is the infamous *sign problem*. Recently a possible solution was proposed by Cristoforetti et al. [1]: We start by embedding the integration domain of the path integral in a complex space. Using the analytical properties of the integrand we then deform the integration manifold without changing the integral value. The proposal by Cristoforetti was to use the a manifold that corresponds to a union of Lefschetz thimbles. This *thimble decomposition* is always possible and it has the advantage that the integrand's complex phase on each thimble is constant. When the integral is dominated by the contribution of one thimble and the residual phase is mildly fluctuating (as in all realistic models), the sign problem is effectively solved [2–6].

In cases where more than one thimble contributes significantly to the integral, the problem is significantly harder. Sampling algorithms have to be able not only to sample each thimble but also be able to dynamically transition between thimbles in order to properly take into account their relative contribution. Additionally, significant analytical work is needed to identify all the thimbles contributing to the integral. In an earlier paper [7] we proposed a method that sidesteps this task. The idea is to use a class of manifolds generated by the *holomorphic gradient flow* and parametrized by the flow time $T_{\rm flow}$ interpolating smoothly between the original integration domain ($T_{\rm flow}=0$) and the thimble decomposition ($T_{\rm flow}=+\infty$). While the value of the integral on all these manifolds is the same, the phase fluctuations become progressively smaller as we increase $T_{\rm flow}$, allowing us to use reweighting. On the other hand as $T_{\rm flow}$ increases the probability distribution becomes multi-modal with growing potential barriers between different modes. For algorithms that rely on small-step updates, such distributions are difficult to sample since the transitions rate between modes becomes very small. For some systems there are values of $T_{\rm flow}$ for which the sign problem is mild and the transition rate between modes is good such that standard small-step algorithms can be used [8, 9].

For systems where the sign problem only becomes manageable when $T_{\rm flow}$ is large and the probability distribution is multi-modal a possible solution is to use the method of *tempered transitions*. The basic idea is to use a set of small-steps to build a large mode-to-mode move [10]. The sequence of steps is constructed using a set of guiding distributions that overlap well sequentially but gradually lower and raise the potential barriers between modes. In this paper we discuss a proposal where the guiding probabilities are generated by changing the $T_{\rm flow}$ and apply it to the

* aalexan@gwu.edu
† gbasar@umd.edu
‡ bedaque@umd.edu
§ ncwarrin@umd.edu

Thirring model in $0+1$ dimensions, a system whose path integral decomposition requires multiple thimbles [5]. The plan of the paper is the following: in Section II we review the relevant details for the holomorphic flow and thimble decomposition, in Section III we review the details of the tempered transitions algorithm as it applies to our problem, and in Section IV we review the relevant details for the $0+1$ Thirring model and present the results of our simulations.

II. HOLOMORPHIC GRADIENT FLOW

Here we show how to deform the domain of path integration in order to ameliorate the sign problem [2, 7]. The starting point is Cauchy's theorem, which allows one to deform the domain of path integration ($\mathbb{R}^N$) into a submanifold $\mathcal{M}$ of complex space ($\mathbb{C}^N \approx \mathbb{R}^{2N}$) without changing the value of the path integral:

$$\langle \mathcal{O} \rangle = \frac{\int_{\mathbb{R}^N} d\zeta_i \; e^{-S[\zeta]} \mathcal{O}[\zeta]}{\int_{\mathbb{R}^N} d\zeta_i \; e^{-S[\zeta]}} = \frac{\int_{\mathcal{M}} d\phi \; e^{-S[\phi]} \mathcal{O}[\phi]}{\int_{\mathcal{M}} d\phi_i \; e^{-S[\phi]}}, \tag{2.1}$$

where ζ_i, $i = 1, \ldots, N$ are real field variables. The sign problem arises because $S[\zeta]$ is not real, leading to rapid phase oscillations in the path integral. The goal is to find a manifold, $\mathcal{M}$, such that Cauchy's theorem applies and $S[\phi]$ does not oscillate as rapidly for $\phi \in \mathcal{M}$ as it does for $\zeta \in \mathbb{R}^N$. One way to construct such a manifold is to identify every field configuration in the original integration domain ($\mathbb{R}^N$) as an initial condition for the following set of first order differential equations known as the *holomorphic gradient flow equations:*

$$\frac{d\phi_i}{dt} = \overline{\frac{\partial S}{\partial \phi_i}}, \quad \phi_i(0) = \zeta_i, \tag{2.2}$$

integrated up to a fixed "flow time" T_{flow}. Here the bar on the RHS of Eq. 2.2 denotes complex conjugation. These equations map a particular field configuration, $\zeta \in \mathbb{R}^N$ to a point $\phi(T_{\text{flow}}) \in \mathbb{C}^N$. We will call this motion the "flow". The map defined by the flow $\zeta \mapsto \phi(T_{\text{flow}})$ is one-to-one as Eq. 2.2 is first order. Therefore flowing $\mathbb{R}^N$ generates a N real-dimensional manifold in $\mathcal{M} \subset \mathbb{C}^N$ (i.e. an N real-dimensional manifold, $\mathcal{M}$, embedded in $\mathbb{C}^N \simeq \mathbb{R}^{2N}$).

Having constructed $\mathcal{M}$ we now to establish that (i) Cauchy's theorem applies on $\mathcal{M}$ (so the path integral on $\mathcal{M}$ is equal to the path integral on $\mathbb{R}^N$) and (ii) the phase oscillations on $\mathcal{M}$ are milder than the phase oscillations on $\mathbb{R}^N$, which leads to a milder sign problem. First observe an important property of the flow equations: the real part of the action, S_R, increases monotonically along a flow trajectory, whereas the imaginary part, S_I, stays constant:

$$\frac{dS_R}{dt} = \frac{1}{2}\left(\frac{\partial S}{\partial \phi_i}\frac{d\phi_i}{dt} + \overline{\frac{\partial S}{\partial \phi_i}}\,\overline{\frac{d\phi_i}{dt}}\right) = \left|\frac{\partial S}{\partial \phi_i}\right|^2 > 0 \quad , \quad \frac{dS_I}{dt} = \frac{1}{2i}\left(\frac{\partial S}{\partial \phi_i}\frac{d\phi_i}{dt} - \overline{\frac{\partial S}{\partial \phi_i}}\,\overline{\frac{d\phi_i}{dt}}\right) = 0 \tag{2.3}$$

where in the second equalities we used Eq. 2.2. For (i) to hold, it must be that points do not cross any singularities throughout the continuous deformation process, especially singularities that might arise when the field variables go to infinity. Indeed no singularities are crossed because S_R increases monotonically with the flow Eq. 2.3, which ensures that the modulus of the integrand, $|e^{-S}| = e^{-S_R}$ decreases monotonically, and therefore remains bounded from above and damps the integral exponentially as the fields approach infinity [1].

For property (ii), consider the limit of large flow time, T_{flow}. In this limit, almost all the original field configurations in $\mathbb{R}^N$ will flow into configurations with very large S_R and will contribute practically nothing to the path integral due to the exponential damping factor e^{-S_R}. Therefore in the large flow limit, the main support of the path integral will come from fields that flow to fixed points of Eq. 2.2 which are critical points of the action $\partial S/\partial \phi_i = 0$ (i.e. classical solutions to the complexified equations of motion). Consider a point in $\mathbb{R}^N$ that flows to a critical point. Then an infinitesimal neighborhood around this point will flow to an N-real dimensional manifold $\mathcal{J}$, attached to that critical point. Since S_I remains unchanged with the flow and the variation of the S_I in an infinitesimal neighborhood is infinitesimal, S_I will be approximately constant on $\mathcal{J}$. In the limit $T_{\text{flow}} \to \infty$, S_I will be exactly constant. This N dimensional manifold $\mathcal{J}$ attached to a critical point, over which S_I is constant is known as a "Lefschetz thimble" (the multi-dimensional generalization of the stationary phase/steepest descent contour familiar from complex analysis). In the limit $T_{\text{flow}} \to \infty$, $\mathcal{M}$ will be a particular combination of thimbles, and S_I will be piecewise constant on $\mathcal{M}$. For sufficiently large, but finite T_{flow}, S_I will not exactly be piecewise constant, but will be approximately piecewise constant. Consequently, by tuning T_{flow} we can continuously soften the severity of the phase oscillations caused by e^{iS_I}. In other words, the manifolds defined by different T_{flow} interpolate between $\mathbb{R}^N$ ($T_{\text{flow}} = 0$) where S_I varies rapidly and the associated thimble decomposition of the path integral ($T_{\text{flow}} \to \infty$) where S_I is piecewise constant.

---

[1] We assumed that the path integral was convergent to begin with. This is indeed the case when the lattice spacing is finite. The standard renormalization procedure has to be followed to approach the continuum limit.

It is desirable, from the point of view of the sign problem, to integrate over a highly flowed manifold because the only regions with appreciable support have nearly constant S_I. However there are costs to this procedure. First, it is numerically expensive to do so. More fundamentally, however, regions of support on the parameterization manifold (where the Monte Carlo takes place) of a highly flowed surface are separated in field space by extended regions with very large S_R. Consequently a Monte Carlo simulation must sample from a multi-modal distribution. The situation is worse when there are fermions involved, because in these cases, thimbles have *boundaries*: $N-1$ dimensional sub-manifolds where the fermion determinant vanishes and $S_R \to \infty$. As an illustration of this phenomenon we shown in Fig. 1 the action on a submanifold of $\mathcal{M}$ for various flow times. One way to avoid this situation is, instead of approaching the $T_{\rm flow} \to \infty$ limit, to tune $T_{\rm flow}$ to a finite value where the phase oscillations are under control yet the action barriers are not so high that the Monte Carlo gets trapped. In a variety of examples this is indeed possible [6, 8, 11]. However, depending on the model and, in particular at large volumes, it may not be possible to solve the sign problem and the isolated minima problem simultaneously by tuning the flow time. Thus we come to the central point of this paper: how to tackle multimodal distributions created by the holomorphic gradient flow.

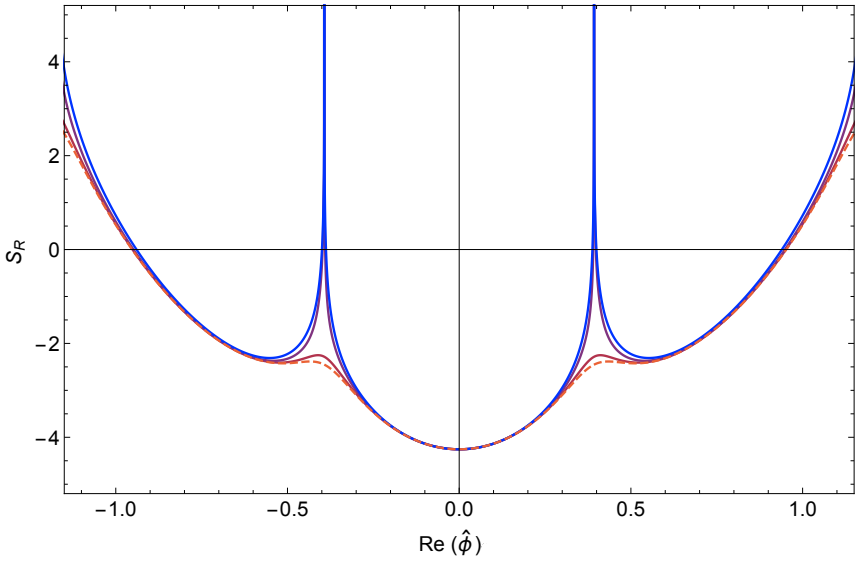

FIG. 1: Plot of the real part of the action on the one dimensional subspace of $\mathcal{M}$ obtained by projecting onto constant fields [7]. The red dotted line is the real part of the action along the tangent space of the thimble while the progressively bluer lines are obtained by flowing the tangent space by increasing amounts. The action barriers quickly diverge to infinity and are the reason for fields getting trapped to a single thimble.

Before concluding this section we outline the Monte Carlo computation that we will employ later on. Using the fact that the flow defines a one-to-one mapping between the initial field, $\phi_i(0) = \zeta_i$ and the flowed field $\phi_i(T_{\rm flow})$ we can parameterize the path integral over $\mathcal{M}$ using real variables ζ_i

$$Z = \int_{\mathcal{M}} d\phi_i \; e^{-S[\phi]} = \int_{\mathbb{R}^N} d\zeta_i \; \det\left(\frac{\partial \phi_i}{\partial \zeta_j}\right) e^{-S[\phi(\zeta)]}. \tag{2.4}$$

Notice that this a re-parameterization of $\mathcal{M}$ and is distinct from the contour deformation from $\mathbb{R}^N$ to $\mathcal{M}$ that we discussed earlier. Parameterizing $\mathcal{M}$ with real fields allows us to perform the Metropolis updates on $\mathbb{R}^N$. The Jacobian, $J_{ij} = \partial\phi_i/\partial\zeta_j$, associated with this change of variables also satisfies a flow equation

$$\frac{dJ_{ij}}{dt} = \overline{H_{ik} J_{kj}}, \quad H_{ij} \equiv \frac{\partial^2 S}{\partial\phi_i \partial\phi_k}, \quad J_{ij}(0) = \delta_{ij}, \tag{2.5}$$

which transports the local tangent space at ζ_i to the flowed point $\phi_i(T_{\rm flow})$ along the flow trajectory followed by ζ_i. The determinant of this Jacobian is a complex number which we combine with the action to define an effective action, $S_{\rm eff}[\zeta] = S[\phi(\zeta)] - \log\det J$. In our Monte Carlo simulations, the configurations are sampled according to the real part of the effective action, $\mathrm{Re}\, S_{\rm eff}[\zeta] = S_R[\phi(\zeta)] - \log|\det J|$. The phase of the Jacobian, along with the phase of the action, $\varphi(\zeta) \equiv \mathrm{Im}\, S_{\rm eff}[\zeta] = S_I[\phi(\zeta)] - \mathrm{Im}\det J$, is included via "reweighing":

$$\langle \mathcal{O} \rangle = \frac{\int d\zeta_i \;\; \mathcal{O} \det J e^{-S[\phi(\zeta)]}}{\int d\zeta_i \;\; \det J e^{-S[\phi(\zeta)]}} = \frac{\int d\zeta_i \;\; \mathcal{O} e^{-i\varphi(\zeta)} e^{-\mathrm{Re}\, S_{\rm eff}[\zeta]}}{\int d\zeta_i \;\; e^{-\mathrm{Re}\, S_{\rm eff}[\zeta]}} \frac{\int d\zeta_i \;\; e^{-\mathrm{Re}\, S_{\rm eff}[\zeta]}}{\int d\zeta_i \;\; e^{-i\varphi(\zeta)} e^{-\mathrm{Re}\, S_{\rm eff}[\zeta]}} = \frac{\left\langle \mathcal{O} e^{-i\varphi(\zeta)} \right\rangle_{\mathrm{Re}\, S_{\rm eff}}}{\left\langle e^{-i\varphi(\zeta)} \right\rangle_{\mathrm{Re}\, S_{\rm eff}}}. \tag{2.6}$$

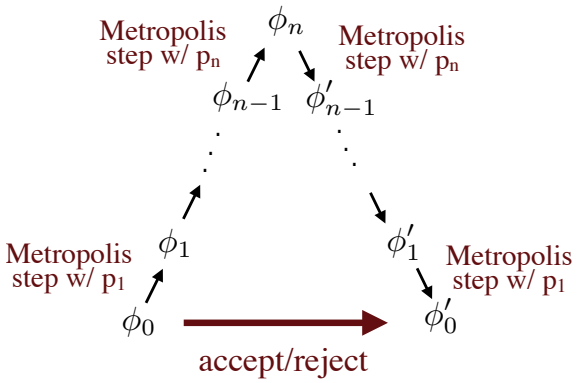

FIG. 2: The tempered transition Monte Carlo step.

Finally, even though we are performing a path integral over $\mathbb{R}^N$ due to the way we parameterize $\mathcal{M}$, the action and all the operators are still evaluated on $\mathcal{M}$. In particular the fluctuations of $S_I[\phi(\zeta)]$ which enter in are drastically milder than those over the original domain given by $S_I[\zeta]$.

III. TEMPERED TRANSITIONS

The method of tempered transitions was designed to perform Monte Carlo calculations in situations where the desired probability distribution is multimodal, that is, has more than one, well separated peaks [10]. Multimodal distributions are challenging for Monte Carlo methods because, with most algorithms, the Monte Carlo chain ends up being trapped in one of the modes for a very large number of steps making it nearly impossible to sample properly.

We now quickly describe the method of tempered transitions in general terms, before applying it to a specific model. Suppose the distribution of interest is $p(\phi)$. Then a standard importance sampling technique is to make symmetric proposals $\phi \to \phi'$ in the sampling space and accept such proposals with probability $\min\{1, \frac{p(\phi')}{p(\phi)}\}$. In order to achieve a reasonable acceptance rate, the proposed ϕ' is chosen close to ϕ. For a multimodal distribution this leads to the trapping alluded above. In the tempered transitions methods one makes a more sophisticated (and computationally expensive) proposal that has a fair likelihood of being on another mode and also of being accepted. This is achieved by considering a sequence of $n+1$ progressively flatter probability distributions $p_i(\phi)$, with $p_0(\phi) = p(\phi)$ being the desired distribution to sample from. For each of these distributions $p_i(\phi)$ consider a transition probability $\hat{T}_{i+1}(\hat{\phi}_i \to \hat{\phi}_{i+1})$ satisfying detailed balance with respect to $p_i(\phi)$:

$$p_i(\phi_i)\hat{T}_{i+1}(\hat{\phi}_i \to \hat{\phi}_{i+1}) = p_i(\phi_{i+1})\hat{T}_{i+1}(\hat{\phi}_{i+1} \to \hat{\phi}_i). \tag{3.1}$$

The transition probabilities $\hat{T}_i(\hat{\phi}_{i-1} \to \hat{\phi}_i)$ can be chosen to be, for instance, Metropolis steps.

The proposed configuration ϕ' is obtained from $\phi = \phi_0$ by evolving ϕ through a series of updates with the transition probabilities $\hat{T}_i(\hat{\phi}_i \to \hat{\phi}_{i+1})$ all the way up to $\hat{T}_n(\hat{\phi}_{n-1} \to \hat{\phi}_n)$ and then again, in reverse order, down to $\hat{T}_1(\hat{\phi}'_1 \to \hat{\phi}'_0)$ (see Fig. 2). More precisely, the probability of proposing $\phi' = \phi'_0$ starting from $\phi = \phi_0$ along the path $\phi_0 \to \phi_1 \to ... \to \phi_1{}' \to \phi_0{}'$ is given by:

$$\begin{aligned}\mathcal{P}(\phi_0 \to \phi'_0) = \hat{T}_1(\hat{\phi_0} \to \hat{\phi_1})\hat{T}_2(\hat{\phi_1} \to \hat{\phi_2}) \cdots \hat{T}_n(\hat{\phi}_{n-1} \to \hat{\phi_n}) \\ \hat{T}_n(\hat{\phi_n} \to \hat{\phi}_{n-1}) \cdots \hat{T}_2(\hat{\phi}'_1 \to \hat{\phi}'_2)\hat{T}_1(\hat{\phi}'_1 \to \hat{\phi}'_0),\end{aligned} \tag{3.2}$$

The final configuration ϕ'_0 is the proposed configuration that is accepted or not according to the acceptance probability

$$\mathcal{A}(\phi_0 \to \phi'_0) = \min\left(1, \frac{p_1(\phi_0)}{p_0(\phi_0)} \cdots \frac{p_n(\phi_{n-1})}{p_{n-1}(\phi_{n-1})} \frac{p_{n-1}(\phi'_{n-1})}{p_n(\phi'_{n-1})} \cdots \frac{p_0(\phi'_0)}{p_1(\phi'_0)}\right). \tag{3.3}$$

It is straightforward to verify that the combination of a tempered transition and the accept/reject step satisfy detailed balance and thus samples the true distribution $p(\phi)$:

$$p(\phi_0)\mathcal{P}(\phi_0 \to \phi'_0)\mathcal{A}(\phi_0 \to \phi'_0) = p(\phi'_0)\mathcal{P}(\phi'_0 \to \phi_0)\mathcal{A}(\phi'_0 \to \phi_0). \tag{3.4}$$

While Eq. 3.4 guarantees the correctness of the method, its usefulness relies on its ability to generate proposals transitioning between modes with a high probability of acceptance. A heuristic discussion of how to choose the intermediate probabilities $p_i(\phi)$ is presented in [10]. To summarize the conclusions reached in [10], it is necessary that the probabilities $p_i(\phi)$ at the "top" of the ladder in Fig. 2 *do not* have well separated modes. This lack of separated modes is what allows the otherwise trapped Monte Carlo chain to travel between them. Moreover, it is necessary that the interpolating probability distributions ($p_i(\phi)$ and $p_{i+1}(\phi)$) be close enough that the probability of accepting a tempered proposal $\mathcal{A}(\phi_0 \to \phi_0')$ is high. Optimizations of the interpolating distributions $p_i(\phi)$ are discussed in [10] for a simple case, however an optimal interpolating procedure for general $p_0(\phi)$ and $p_n(\phi)$ is not obvious.

In our implementation of tempered transitions, we modulate the probability distributions up the ladder by adjusting the amount of flow we subject the parameterization manifold to. It can be seen from Fig. 1 that at $T_{\text{flow}} = 0.0$, the action barriers on $\mathcal{M}$ are mild while for large T_{flow} the action barriers are high. Therefore, we choose $p_0(\phi)$ to be the probability distribution generated with a large enough T_{flow} to tame the sign problem and we choose $p_n(\phi)$ to be the probability distribution generated when $T_{\text{flow}} = 0.0$ where fields are mobile. There is much latitude in how to choose the probability distributions between $p_0(\phi)$ and $p_n(\phi)$, but not all choices perform equally well. Motivated by [10], we interpolate between $T_{\text{flow}} = 0.0$ and T_{max} linearly. Additionally, $\hat{T}_i$ consists of several Metropolis steps at each flow time. We find that this gives the ladder a chance to "equilibrate" locally and increases the probability of accepting a tempered proposal. For good acceptance thousands of ladder steps are required. This also has the effect of allowing fields to travel far.

IV. RESULTS

We now apply the method of tempered transitions on $0+1$ Thirring model at finite density. For non-zero chemical potential the determinant of the Dirac matrix is complex and the theory suffers from a sign problem. We discretize the Euclidean time direction using staggered fermions. The lattice partition function for this theory is

$$Z = \left[\prod_{t=1}^{N} \int_0^{2\pi} \frac{d\phi_t}{2\pi}\right] \det D e^{-\frac{1}{2g^2}\sum_{t=1}^{N}(1-\cos\phi_t)} \equiv \left[\prod_{t=1}^{N} \int_0^{2\pi} \frac{d\phi_t}{2\pi}\right] e^{-S[\phi]}, \tag{4.1}$$

where the effective action and the Dirac matrix are explicitly given by

$$S[\phi] = \frac{1}{2g^2}\sum_{t=1}^{N}(1-\cos\phi_t) - \log\det D \tag{4.2}$$

$$D_{t,t'} = \frac{1}{2}\left(e^{\mu+i\phi_t}\delta_{t+1,t'} - e^{-\mu-i\phi_{t'}}\delta_{t-1,t'} + e^{-\mu-i\phi_{t'}}\delta_{t,1}\delta_{t',N} - e^{-\mu-i\phi_t}\delta_{t,N}\delta_{t',1}\right) + m\,\delta_{t,t'}\,. \tag{4.3}$$

Here N is an even number equal to to the number of lattice sites and all dimensionful quantities are measured in units of the lattice spacing a which we set to one. This discretized model is exactly solvable [12]; an observable of interest is the chiral condensate, which is given by:

$$\langle\bar{\psi}\psi\rangle = \frac{1}{N}\frac{\partial}{\partial m}\log Z = \frac{(1+m^2)^{-1/2} I_0^N(\frac{1}{2g^2})\sinh(N\sinh^{-1}(m))}{I_1^N(\frac{1}{2g^2})\cosh(N\mu) + I_0^N(\alpha)\cosh(N\sinh^{-1}(m))} \tag{4.4}$$

where $I_n(z)$ denotes the modified Bessel function of the first kind of order n. We will use this exact solution to compare against.

As an example we take an $N = 16$ lattice with $m = 1.0$ and $g^2 = \frac{1}{6}$ over a range of chemical potentials. First, we reduce the sign problem for these parameters by shifting the domain of integration from the real hyperplane to the tangent plane of the thimble fixed to the global minimum of the action [7] where the phase fluctuations are smaller than on the real plane. The remaining phase fluctuations can be tamed by flowing the tangent plane. Calculations performed with flow times in the range $0.2 \lesssim T_{\text{flow}} \lesssim 0.4$ shows that at least three thimbles contribute significantly (see Fig. 3). For larger values of flow time, calculations performed with the Metropolis algorithm remain trapped in the region with $\langle\phi\rangle \approx 0$ for a very large number of steps (we have followed the Monte Carlo chain up to 10 million steps without seeing a transition to other thimbles). For some of the parameters explored, in particular when the chiral condensate drops quickly around the mass of the fermion ψ, computing observables when trapped to the global minimal thimble yields statistically incorrect results, see Fig. 4. The need to integrate over several thimbles at points of sharp variation in thermodynamic functions has been stressed before in fermionic models [13].

As the discrepancy between the trapped numerical result and the exact result is largest at $\mu = 1.0$, we restrict our further analysis to this value of the chemical potential. We apply the method of tempered transitions using a ladder

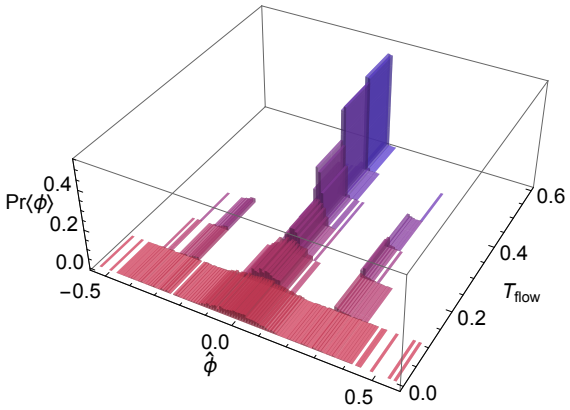

FIG. 3: A demonstration of the trapping process. We show histograms of the time average of the field $\hat{\phi} = \frac{1}{N}\sum_{k=1}^{N}\phi_k$ for simulations performed in a range of six flow times ($T_{\text{flow}} = 0.0, 0.1, ..., 0.5$). The widest histogram corresponds to a Monte Carlo run at zero flow on the tangent plane of the global minimal thimble and the narrowest histogram corresponds to a Monte Carlo run at $T_{\text{flow}} = 0.5$. Beyond $T_{\text{flow}} = 0.3$, the Monte Carlo is unable to tunnel through to neighboring thimbles and misses their significant contributions. We have demonstrated that at the $T_{\text{flow}} = 0.5$ this trapping is for all practical purposes indefinite. We will use the value $T_{\text{flow}} = 0.5$ to illustrate how the tempered transition algorithm allows for the proper sampling of the space even in this case.

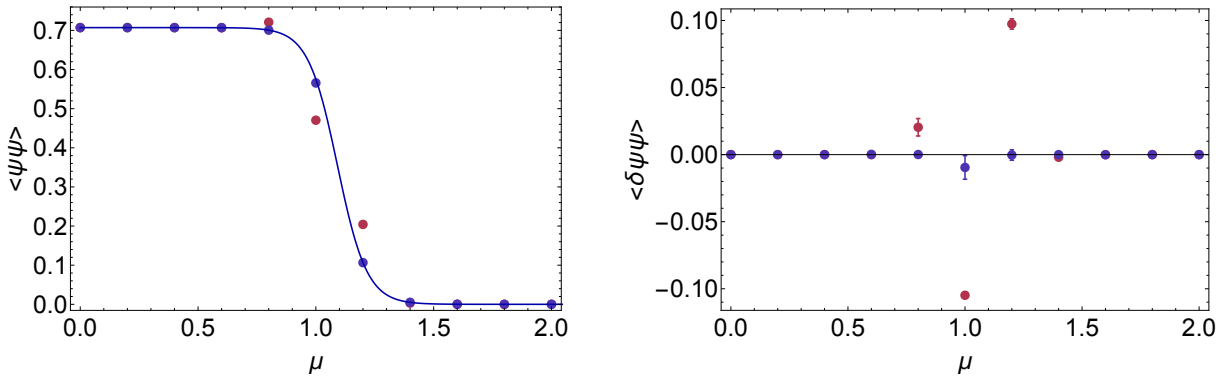

FIG. 4: In the left panel we have the numerical computation of the chiral condensate $\langle\bar{\psi}\psi\rangle$ at $T_{\text{flow}} = 0.0$ (shown in blue) and $T_{\text{flow}} = 0.5$ (shown in red) for the parameters $N_t = 16$, $g^2 = \frac{1}{6}$ and $m = 1.0$. The exact result is shown as the solid line. In the right panel we plot the difference between the exact result and the numerical result with the same color coding. The discrepancy between the exact result and the numerical result is easily visible for chemical potentials near the transition.

with 1600 steps in each direction. At the bottom of the ladder the flow is 0.5 where the fields are trapped to a single thimble, and at the top the flow is 0.0 were the Monte Carlo chain explores all thimbles. We interpolate linearly in the flow as sub-transitions are applied. At each step of the ladder, we use a transition operator $T_k(\phi, \phi')$ which applies 10 standard Metropolis proposals at fixed flow time.

There are two time scales in the equilibration of the Monte Carlo chain: a fast one for the equilibration within one mode and a slow one for the equilibration between modes. These separated time scales exist for the simple reason that only small proposals are required to explore an individual mode while a large proposal is required to transition between them. This is a generic property of distribution with multiple modes. In light of this observation, we construct our Monte Carlo as follows: between two tempered transitions we perform 1000 standard Metropolis steps. Such a division is convenient because tempered transitions are expensive compared to standard ones. In spite of this cost, tempered proposals induce transitions at a substantially higher rate than standard Metropolis proposals. In this study, a tempered transition is composed of 10×3200 individual Metropolis steps. We find that roughly every 10^{th} proposal between the main thimble and the shoulder thimbles is accepted, so it takes the computational effort of $\sim 3 \times 10^5$ Metropolis steps to induce a transition. Therefore tempered proposals are at least 1–2 orders of magnitude more efficient than standard Metropolis proposals in this study.

Method	$\langle\bar{\psi}\psi\rangle$
Exact	0.575
with tempered transition	0.602(20)
w/o tempered transitions	0.470(02)

TABLE I: Results for the chiral condensate at $\mu = 1.0$ for the exact solution, a Monte Carlo trapped to the global minimum thimble and a Monte Carlo utilizing tempered transitions.

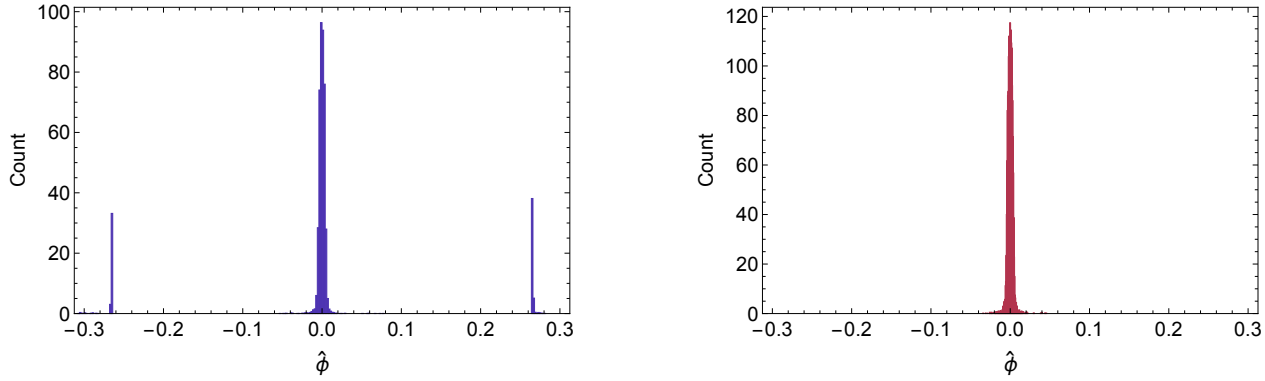

FIG. 5: Here we show the distribution over the course of a simulation of the time average in the tempered (left) and non-tempered (right) case. The two Monte Carlos are of equal length and sample from the same sharply peaked probability distribution with isolated regions of support. The simulation without tempered steps misses the two side peaks.

The results of a Monte Carlo using 2000 tempered transitions is shown in Fig. 5. We find that the combination of tempered transitions with 1600 ladder steps and regular Metropolis steps is sufficient to transition regularly between thimbles. In addition, we find that the inclusion of the neighboring thimbles reproduces the exact result up to statistical errors as can be seen in Table I. Roughly 10% of tempered transitions from the central thimble to the shoulder thimbles were accepted. Note however that this low acceptance rate is not due to poor proposals but because the shoulder thimbles carry roughly 10% of the weight of the path integral. In fact, the tempered transitions make proposals to the next-to-nearest shoulder thimbles frequently. This can be seen in Fig. 6 where the tempered proposals are plotted as a function of Monte Carlo step. These transitions to the next-to-nearest shoulder thimbles are, however, not accepted because these thimbles contribute very little to the path integral. It is expected, then, that in a problem where many thimbles carry substantial weight, tempered transitions may provide a natural means to explore many thimbles without any *a priori* knowledge of their location.

To summarize the algorithm consists of the following steps:

1. 1000 Metropolis updates with accept/reject based on the Boltzmann factor $\exp\left(-\operatorname{Re} S_{\text{eff}}[T_{\text{flow}} = T_{\max}]\right)$.

2. One tempered step:

   (a) For $i = 1, \ldots, 1600$ set $T_{\text{flow}} = t_i \equiv T_{\max}(1 - \frac{i}{1600})$.
   10 Metropolis updates with accept/reject based on the Boltzmann factor $\exp\left(-\operatorname{Re} S_{\text{eff}}[T_{\text{flow}} = t_i]\right)$.

   (b) For $i = 1600, \ldots, 1$ set $T_{\text{flow}} = t_i$.
   10 Metropolis updates with accept/reject based on the Boltzmann factor $\exp\left(-\operatorname{Re} S_{\text{eff}}[T_{\text{flow}} = t_i]\right)$.

   (c) accept/reject using the acceptance probability in Eq. 3.3.

3. Make a measurement, including the calculation of the phase $\operatorname{Im} S_{\text{eff}}$, and repeat from step 1.

The stepsize for these proposals is adjusted such that the change in the field variables on the integration manifold remains roughly the same. This is accomplished by the scaling of the stepsize (in the parametrization manifold) described in reference [5]. As the flow time is decreased, the stepsize increases.

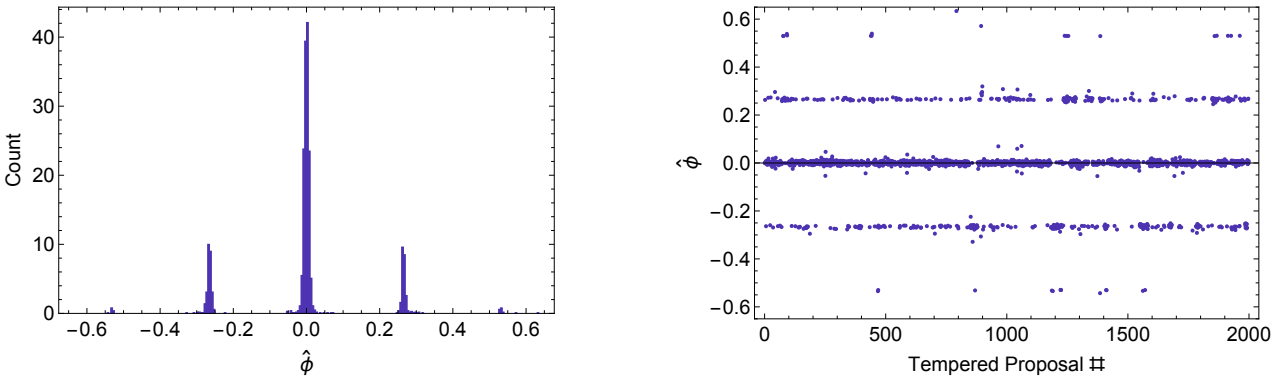

FIG. 6: Proposals made by tempered steps (projected on the direction) as a function of Monte Carlo step (right panel) and its histogram (left panel). Notice that proposals are made mainly to five different regions, corresponding to the five dominating thimbles. The two next-to-nearest thimbles to the central one are not accepted, in accordance to their small statistical weight.

V. CONCLUSION

The holomorphic gradient flow approach to the computation of path integrals with a sign problem frequently leads to multimodal probability distributions. Each of the modes are related to a thimble contributing to the integral, as described by Picard-Lefschetz theory. This poses a challenge to numerical computations as Monte Carlo chains tend to get "stuck" in one of the modes for exponentially long times. We applied the method of tempered transitions to this problem. We take the tempering parameter, which controls the steepness of the probability landscape in each ladder of the tempering process, to be the flow time [14] by which the real space is transported by the holomorphic gradient flow. We demonstrated in a simple example that this procedure is feasible and that it allows for the proper sampling of the field space even in circumstances where a simpler Metropolis algorithm fails. It was found that a combination of tempered steps interspersed with regular Metropolis steps was the most efficient choice.

The method is not without drawbacks. The main one is the need, even in the simple model considered here, of a very large number of ladders steps during a tempered proposal and the associated large computational cost. This is the main difficulty that limited us in this paper to fairly small toy models. Our experience in scaling up the number of degrees of freedom is that the number of ladder steps required scales roughly linearly with the number of degrees of freedom. By itself this is not such a steep scaling but it should be kept in mind that other, steeper increases in cost are caused by the computation of the jacobian, increase in required flow time and, in models with more spacetime dimensions, increase on the size of the Dirac matrix (in the $0+1$ dimensional model discussed here the determinant has a closed form). This is not to say that significant improvements are not possible by adjusting some parameters in our simulations. For instance, there is tremendous latitude in choosing the values of the intermediate flows. Most applications of the tempered transitions method use the temperature as the parameter changing along the tempered ladder, leading to an exponential flattening of the probabilities distributions [10]. A similar exponential flattening of probabilities can be obtained by choosing the intermediate flows to be equally spaced along the ladder. It could well be that a different choice is significantly better.

As we finished the present paper a study of the same model we discuss using a different tempering method appeared [14]. A direct comparison of the efficacy of the two methods is hindered by some small differences between the two calculations that are unrelated to the tempering method. For instance, in [14], the manifolds of integration are obtained by flowing the real manifold, not the main tangent manifold as is done here. It is clear, however, that either method will be extremely costly when applied to realistic field theories.

ACKNOWLEDGMENTS

A.A. is supported in part by the National Science Foundation CAREER grant PHY-1151648 and by U.S. DOE Grant No. DE-FG02-95ER40907. A.A. gratefully acknowledges the hospitality of the Physics Department at the University of Maryland where part of this work was carried out.

G.B. is supported by the U.S. Department of Energy under Contract No. DE-FG02-01ER41195.

P.B. and N.C.W. are supported by U.S. Department of Energy under Contract No. DE-FG02-93ER-40762.

---

[1] **AuroraScience** Collaboration, M. Cristoforetti, F. Di Renzo, and L. Scorzato, *New approach to the sign problem in quantum field theories: High density QCD on a Lefschetz thimble*, *Phys. Rev.* **D86** (2012) 074506, [arXiv:1205.3996].

[2] M. Cristoforetti, F. Di Renzo, A. Mukherjee, and L. Scorzato, *Monte Carlo simulations on the Lefschetz thimble: Taming the sign problem*, *Phys. Rev.* **D88** (2013), no. 5 051501, [arXiv:1303.7204].

[3] A. Mukherjee, M. Cristoforetti, and L. Scorzato, *Metropolis Monte Carlo integration on the Lefschetz thimble: Application to a one-plaquette model*, *Phys. Rev.* **D88** (2013), no. 5 051502, [arXiv:1308.0233].

[4] H. Fujii, D. Honda, M. Kato, Y. Kikukawa, S. Komatsu, and T. Sano, *Hybrid Monte Carlo on Lefschetz thimbles - A study of the residual sign problem*, *JHEP* **10** (2013) 147, [arXiv:1309.4371].

[5] A. Alexandru, G. Basar, and P. Bedaque, *Monte Carlo algorithm for simulating fermions on Lefschetz thimbles*, *Phys. Rev.* **D93** (2016), no. 1 014504, [arXiv:1510.03258].

[6] A. Alexandru, G. Basar, P. Bedaque, G. W. Ridgway, and N. C. Warrington, *Study of symmetry breaking in a relativistic Bose gas using the contraction algorithm*, *Phys. Rev.* **D94** (2016), no. 4 045017, [arXiv:1606.02742].

[7] A. Alexandru, G. Basar, P. F. Bedaque, G. W. Ridgway, and N. C. Warrington, *Sign problem and Monte Carlo calculations beyond Lefschetz thimbles*, *JHEP* **05** (2016) 053, [arXiv:1512.08764].

[8] A. Alexandru, G. Basar, P. F. Bedaque, S. Vartak, and N. C. Warrington, *Monte Carlo study of real time dynamics*, arXiv:1605.08040.

[9] A. Alexandru, G. Basar, P. F. Bedaque, G. W. Ridgway, and N. C. Warrington, *Monte Carlo calculations of the finite density Thirring model*, *Phys. Rev.* **D95** (2017), no. 1 014502, [arXiv:1609.01730].

[10] R. M. Neal, *Sampling from multimodal distributions using tempered transitions*, *Statistics and Computing* **6** (1996), no. 4 353–366.

[11] C. Schmidt and F. Ziesché, *Simulating low dimensional QCD with Lefschetz thimbles*, in *Proceedings, 34th International Symposium on Lattice Field Theory (Lattice 2016): Southampton, UK, July 24-30, 2016*, 2017. arXiv:1701.08959.

[12] J. M. Pawlowski and C. Zielinski, *Thirring model at finite density in 0+1 dimensions with stochastic quantization: Crosscheck with an exact solution*, *Phys. Rev.* **D87** (2013), no. 9 094503, [arXiv:1302.1622].

[13] Y. Tanizaki, Y. Hidaka, and T. Hayata, *Lefschetz-thimble approach to the Silver Blaze problem of one-site fermion model*, *PoS* **LATTICE2016** (2016) 030, [arXiv:1610.00393].

[14] M. Fukuma and N. Umeda, *Parallel tempering algorithm for the integration over Lefschetz thimbles*, arXiv:1703.00861.