

Finite-density Monte Carlo calculations on sign-optimized manifolds

Andrei Alexandru,^{1,2,3,*} Paulo F. Bedaque,^{2,†} Henry Lamm,^{2,‡} and Scott Lawrence^{2,§}

¹*Department of Physics, The George Washington University, Washington, DC 20052, USA*

²*Department of Physics, University of Maryland, College Park, Maryland 20742, USA*

³*Albert Einstein Center for Fundamental Physics, Institute for Theoretical Physics, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland*



(Received 5 April 2018; published 31 May 2018)

We present a general technique for addressing sign problems that arise in Monte Carlo simulations of field theories. This method deforms the domain of the path integral to a manifold in complex field space that maximizes the average sign (therefore reducing the sign problem) within a parametrized family of manifolds. We present results for the 1 + 1 dimensional Thirring model with Wilson fermions on lattice sizes up to 40×10 . This method reaches higher μ than previous techniques while substantially decreasing the computational time required.

DOI: [10.1103/PhysRevD.97.094510](https://doi.org/10.1103/PhysRevD.97.094510)

I. INTRODUCTION

Monte Carlo methods are critical to the study of field-theoretical and many-body systems. In particular, they are the only general-purpose approach to address strongly interacting field theories. The basic idea of all Monte Carlo methods is simple: observables are formulated as path integrals which, on a discretized spacetime, become high dimensional integrals. Those are then estimated stochastically by importance sampling. Importance sampling relies on interpreting part of the integrand (typically the exponential of the action) as a probability, which makes sense only if this term is real and non-negative. Unfortunately, many theories, even when formulated in imaginary time (Euclidean space), have a negative or even complex integrand. This so-called “sign problem” is a major roadblock to the understanding of some of the most important systems in physics. Many systems at finite density (including QCD at finite baryon density) and nonrelativistic systems lacking some special symmetry between fermion species (as in the Hubbard model away from half-filling or on nonbipartite lattices) suffer from sign problems. Also, some real-time observables in thermal equilibrium as well as truly nonequilibrium phenomena are not amenable to imaginary time calculations and have a

particularly severe sign problem that renders most Monte Carlo methods nonstarters. A simple, albeit not very effective, way of dealing with the sign problem is to choose a manifestly positive part of the integrand as the statistical weight while moving the part with the fluctuating sign or phase to the observable to be measured. This “reweighting” is effective to the extent that the average sign, that is, the average of the fluctuating sign on the ensemble defined by the positive measure, is not too small. However, in theories with sign problems, the average sign typically decreases exponentially with the volume and the inverse temperature of the system. Many techniques have been proposed in the past to ameliorate the sign problem. Among them are the complex Langevin method [1], the density of states method [2], canonical methods [3,4], reweighting methods [5], series expansion in the chemical potential [6], fermion bags [7], dual variables [8], and analytic continuation from imaginary chemical potentials [9]. Each one has its successes and pitfalls. It is fair to say, however, that the sign problems of field theories remain largely unsolved.

More recently, the “thimble” method was proposed [10,11]. The main idea is to complexify the domain of the path integral. Instead of integrating over real values of the fields, one deforms the manifold of integration from $\mathbb{R}^N \subset \mathbb{C}^N$ to some other N -dimensional manifold, $\mathcal{M} \subset \mathbb{C}^N$. A multidimensional generalization of Cauchy’s theorem of complex analysis guarantees, under some conditions on \mathcal{M} , that the integral over \mathcal{M} and \mathbb{R}^N of any holomorphic integrand is the same. This allows one to compute expectation values of observables \mathcal{O} for which $\mathcal{O}e^{-S}$ is holomorphic, if \mathcal{M} is properly chosen. The key to these methods is that the average sign $\langle e^{-iS_I} \rangle$ is an integral of a nonholomorphic function, and therefore depends upon

* aalexan@gwu.edu

† bedaque@umd.edu

‡ hlammm@umd.edu

§ srl@umd.edu

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Funded by SCOAP³.

the integration manifold, whereas the physical expectation values do not. The manifold \mathcal{M} was originally suggested to be the combination of thimbles, multidimensional generalizations of the steepest descent/constant phase path familiar from complex analysis. This method and its associated algorithmic problems were pursued by several groups [12–17,17–25]. Relevant analytical work, closely connected to the “resurgent transseries” (for a recent review see [26]) was also pursued in [27–31]. Experience with actual simulations made evident some problems with the thimble approach. The first is that thimbles are complicated manifolds that have to be found “on the fly” by the algorithm, and the lack of a local characterization of thimbles makes this computationally expensive. Second, theories where more than one thimble contributes to the path integral significantly [28] are particularly difficult to sample [17,21].

This led to some modifications of the method. In the generalized thimble method [20,23,24] the manifold of integration \mathcal{M} is chosen to be the deformation of \mathbb{R}^N by the holomorphic flow defined by the action. If \mathbb{R}^N is deformed by the flow by an infinite amount of flow time, \mathcal{M} approaches the right combination of thimbles equivalent to the original integration domain. If the flow is stopped at some finite flow time, \mathcal{M} is close, but not identical, to the sum of appropriate thimbles. It is, however, a legitimate manifold of integration in the sense that it gives exactly the same result as the original manifold \mathbb{R}^N . The advantage of the manifold \mathcal{M} over the thimbles is that (1) less flow corresponds to smaller computational cost and (2) \mathcal{M} can be algorithmically constructed during the simulation by solving the flow equations, while finding the thimbles and determining which ones contribute to the integral is a difficult task in all but the simplest field theories.

This is not to say that the generalized method does not have its own problems. Large flow times can improve the sign problem but generate multimodal distributions difficult to sample. Shorter flow times avoid the multimodality but improve the sign problem less, so the flow time has to be carefully chosen; in fact, there is no guarantee that a “middle ground” flow time can be found. (Multimodality can also be dealt with by more sophisticated sampling algorithms [25,32].) In addition, the computation of the Jacobian arising from parametrizing \mathcal{M} by the initial point of the flow in \mathbb{R}^N is expensive. The proposal presented in this paper drastically reduces the cost of the Jacobian. At the same time, it provides more flexibility in the choice of \mathcal{M} while systematically improving the sign problem.

One step towards speeding up the costly calculations involved in the generalized thimble method was given in [22]. A feed-forward neural network was trained to interpolate points in \mathcal{M} obtained by the more expensive holomorphic flow. The neural net was then used to quickly generate more points in \mathcal{M} . In the present paper we go one step further and completely bypass the need to generate

points by flowing. Instead, we seek to flow directly toward a manifold of maximum average sign, albeit in a restricted family of manifolds \mathcal{M}_λ which are parametrized by a finite number of parameters λ . A similar proposal based on maximizing the approximate average sign was pursued in [33,34]. Our method produces a manifold \mathcal{M} that can be sampled as rapidly as \mathbb{R}^N via

$$\begin{aligned} \langle \mathcal{O} \rangle &= \frac{\int_{\mathcal{M}} \mathcal{D}\tilde{\phi} \mathcal{O}(\tilde{\phi}) e^{-S(\tilde{\phi})}}{\int_{\mathcal{M}} \mathcal{D}\tilde{\phi} e^{-S(\tilde{\phi})}} \\ &= \frac{\int_{\mathbb{R}^N} \mathcal{D}\phi \mathcal{O}[\tilde{\phi}(\phi)] e^{-S[\tilde{\phi}(\phi)]} \det J(\phi)}{\int_{\mathbb{R}^N} \mathcal{D}\phi e^{-S[\tilde{\phi}(\phi)]} \det J(\phi)} \\ &\equiv \frac{\int_{\mathbb{R}^N} \mathcal{D}\phi \mathcal{O}[\tilde{\phi}(\phi)] e^{-S_{\text{eff}}(\phi)}}{\int_{\mathbb{R}^N} \mathcal{D}\phi e^{-S_{\text{eff}}(\phi)}}, \end{aligned} \quad (1)$$

where a point $\tilde{\phi}$ in \mathcal{M} is parametrized by a point ϕ in \mathbb{R}^N . We present in Sec. II how the algorithm can be implemented, with special emphasis on the gradient ascent method we use to obtain the local maximum value of average sign. Further, it is shown that the derivative of the sign problem with respect to λ can be efficiently calculated despite a potentially small sign.

The method of determining an optimal manifold for integration, as well as the procedure for integrating along that manifold, is detailed in Sec. II. In Sec. III, we define the physical model we study with this algorithm, the Thirring model. In Sec. IV, we present our results, and conclusions are summarized in Sec. V.

II. THE METHOD

We start by specifying a family \mathcal{M}_λ of submanifolds of \mathbb{C}^N , parametrized by λ . The choice of this family is guided by the ease of computation of the Jacobian and some experience acquired with the generalized thimble method. We then proceed to maximize the average sign among this family of manifolds using a simple gradient ascent technique. On a manifold of integration \mathcal{M}_λ , the average sign is

$$\langle \sigma \rangle_\lambda = \frac{\int_{\mathbb{R}^N} \mathcal{D}\phi e^{-S_{\text{eff}}[\phi;\lambda]}}{\int_{\mathbb{R}^N} \mathcal{D}\phi e^{-\text{Re}S_{\text{eff}}[\phi;\lambda]}}, \quad (2)$$

where ϕ are the fields in the theory and $S_{\text{eff}} \equiv S - \ln \det J$ is the effective action. On this manifold, we compute a vector proportional to the gradient of the magnitude of the average sign, and then proceed to change λ by a small amount along this vector,

$$\lambda_s - \lambda_{s-1} \propto \eta \nabla_\lambda |\langle \sigma \rangle_\lambda|. \quad (3)$$

Here the step size η determines how large each step along the computed gradient should be. We initialize \mathcal{M}_{λ_0} to be \mathbb{R}^N . After a large number of steps, and if the step size η is

small enough, we should arrive at a (local) maximum of the average sign. Critically, the computation of the direction of the gradient has no sign problem.

We now show how to compute the direction of the gradient. The numerator of Eq. (2), being the integral of a holomorphic function e^{-S} along \mathcal{M} , does not depend on λ . In contrast, since the integral of $e^{-\text{Re}S_{\text{eff}}}$ cannot be written as an integral of a holomorphic function, the denominator will vary with λ . The gradient of the magnitude $|\langle\sigma\rangle_\lambda|$ with respect to the manifold parameters λ , then, is given by

$$\begin{aligned} \nabla_\lambda |\langle\sigma\rangle_\lambda| &= -|\langle\sigma\rangle_\lambda| \frac{\nabla_\lambda \int_{\mathbb{R}^N} \mathcal{D}\phi e^{-\text{Re}S_{\text{eff}}[\phi;\lambda]}}{\int_{\mathbb{R}^N} \mathcal{D}\phi e^{-\text{Re}S_{\text{eff}}[\phi;\lambda]}} \\ &= |\langle\sigma\rangle_\lambda| \frac{\int_{\mathbb{R}^N} \mathcal{D}\phi e^{-\text{Re}S_{\text{eff}}[\phi;\lambda]} [\nabla_\lambda S_R - \text{ReTr}J^{-1}\nabla_\lambda J]}{\int_{\mathbb{R}^N} \mathcal{D}\phi e^{-\text{Re}S_{\text{eff}}[\phi;\lambda]}}. \end{aligned} \quad (4)$$

From this, we see that the gradient factorizes into two pieces: the average sign on \mathcal{M}_λ , and an expectation value of an operator on that manifold. The second factor is an expectation value with respect to $e^{-\text{Re}S_{\text{eff}}}$ and therefore is sign-problem free; the first is a scalar which does not affect the direction. This allows us to compute (up to that overall scalar) the gradient on a manifold reliably by a short Monte Carlo simulation. For a gradient ascent method, an overall magnitude like $|\langle\sigma\rangle_\lambda|$ (even varying with λ) can be safely neglected: it does not change the direction that the gradient points in λ -space. This allows our method to be efficient even when the average sign is statistically indistinguishable from zero. Therefore, at each step s of the gradient ascent, we update the manifold parameters λ_s according to

$$\lambda_s - \lambda_{s-1} = \eta \langle \nabla_\lambda S_R - \text{ReTr}J^{-1}\nabla_\lambda J \rangle_{\text{Re}S_{\text{eff}}}. \quad (5)$$

In principle, one might use a more sophisticated stochastic gradient ascent algorithm to both speed up the calculation and avoid finding a suboptimal local maximum. For this work, we found that a naive gradient ascent converges adequately swiftly, and the stochastic nature of the Monte Carlo simulation used to compute the gradient helps to explore parameter space.

In a gradient ascent method, the parameter η must be chosen to be small enough to avoid overshooting a maximum, but not much smaller; otherwise it will oscillate around the maximum but never converge. For our purposes, there is one additional practical consideration restricting the size of η . In calculating the expectation value of Eq. (5), we would like to avoid needing to completely rethermalize the Markov chain after every gradient ascent step. To this end, we set the step size η to be sufficiently small that $S[\tilde{\phi}(\phi)]$, for any fixed ϕ , changes only slowly with s . The value of ϕ at the end of one Monte Carlo run can then be used to seed

the next run on the new manifold, minimizing the necessary thermalization time.

It should be stressed that lack of care in this process, or in any other detail of the sign maximization process, may reduce the average sign of the manifold ultimately found and increase the computational time, but it does not affect the correctness of physical observables on that manifold. The ‘‘real-plane’’ integral is calculated as an integral over compact variables, that is, an integral over $\mathbb{T}^N = (S^1)^N$. The manifold \mathcal{M}_λ is a submanifold of the complexified N -torus $(S^1 \times \mathbb{R})^N$. Cauchy’s integral theorem guarantees that, provided the domain of integration is compact (as \mathbb{T}^N is), the integral over \mathcal{M}_λ will equal that over \mathbb{T}^N if the manifold \mathbb{T}^N is continuously deformable to \mathcal{M}_λ . For our purposes, this is guaranteed by making the family \mathcal{M}_λ continuous in the parameters λ , and letting $\mathcal{M}_\lambda = \mathbb{T}^N$.

The determinant of the Jacobian J —which must be computed during a Monte Carlo on \mathcal{M} —is a potentially expensive operation, with a cost approximately cubic in the number of degrees of freedom. To avoid this, we will choose an ansatz family \mathcal{M}_λ for which the Jacobian is diagonal. In particular, we write $\tilde{\phi}_i(\phi) = \phi_i + i f_i(\phi_i)$, so that $J_{ij} = \delta_{ij}(1 + f'_i(\phi_i))$, which is the most general ansatz possible satisfying our constraints. Relaxing this constraint to a nondiagonal Jacobian should improve the sign problem by allowing nonlocal correlations in the imaginary components of ϕ , but this will come at a computational expense and will be left to future work.

III. THIRRING MODEL

In order to make the ideas more concrete, we will phrase our discussion in terms of a specific field theory model, the 1 + 1D massive Thirring model with Wilson fermions. The lattice action is given by

$$S = \sum_{x,\nu} \frac{N_F}{g^2} (1 - \cos A_\nu(x)) + \sum_{x,y} \bar{\psi}^a(x) D_{xy}^W(A) \psi^a(y) \quad (6)$$

with

$$\begin{aligned} D_{xy}^W &= \delta_{xy} - \kappa \sum_{\nu=0,1} [(1 - \gamma_\nu) e^{iA_\nu(x) + \mu\delta_{\nu,0}} \delta_{x+\nu,y} \\ &\quad + (1 + \gamma_\nu) e^{-iA_\nu(x) - \mu\delta_{\nu,0}} \delta_{x,y+\nu}], \end{aligned}$$

where ψ is a two-component Dirac spinor with the flavor indices a taking values from 1, ..., N_F ; g is the coupling; μ is the fermion chemical potential; and $\kappa = 1/(2m + 4)$, where m is the bare mass of the fermions. Standard universality arguments applied to this asymptotically free theory indicate that, in the continuum limit, this action is equivalent to the continuum action

$$S = \int d^2x \left[\bar{\psi}^a (\not{\partial} + \mu\gamma_0 + m)\psi^a + \frac{g^2}{2N_F} \bar{\psi}^a \gamma_\mu \psi^a \bar{\psi}^b \gamma_\mu \psi^b \right]. \quad (7)$$

The four-fermion interaction is generated when the bosonic $A_\mu(x)$ auxiliary field is integrated over. The Thirring model was chosen since other similar methods have been applied to it; thus, it serves as a useful benchmark for our method.

The integration over the fermion fields results in the action

$$S_{\text{eff}} = N_F \left(\frac{1}{g^2} \sum_{x,\nu} (1 - \cos A_\nu(x)) - \log \det D(A) \right). \quad (8)$$

In this work we take $N_F = 2$. For finite chemical potential $\mu \neq 0$, the determinant $\det D(A)$ is not strictly real, and we must address a sign problem.

In applying the method described in Sec. II, we enforce three additional constraints on f , all coming from symmetries of the action Eq. (8). The action is 2π -periodic in the fields A_0, A_1 and an even function; therefore, we require the same of f . Finally, the action is invariant under translations of the lattice. The lattice degrees of freedom are divided into timelike links A_0 and spatial links A_1 . Translational invariance of the f_i implies that the form of f_i can depend only on whether the index i refers to an A_0 field or an A_1 field.

Consistent with these demands, we use a simple two-parameter family with $f_0(\phi) = \lambda_0 + \lambda_1 \cos \phi$ and $f_1(\phi) = 0$, so that the manifold \mathcal{M}_λ is defined by

$$\begin{aligned} \tilde{A}_0(A_0, A_1) &= A_0 + i(\lambda_0 + \lambda_1 \cos A_0) \\ \tilde{A}_1(A_0, A_1) &= A_1. \end{aligned}$$

As discussed above, Cauchy's theorem guarantees that expectation values computed on \mathcal{M}_λ are equal to those computed on \mathbb{T}^N , provided that one manifold may be continuously deformed to the other. To see that this is so, note that $\mathcal{M}_0 = \mathbb{T}^N$ and that \tilde{A} is a continuous function of λ . One might consider using a larger class of manifolds. We have investigated including a $\cos^2 A_0$ term in $\tilde{A}_0(A_0, A_1)$ and a $\cos A_1$ in the $\tilde{A}_1(A_0, A_1)$, but in all cases, we found negligible improvement in the average sign computed on the resulting manifolds.

Once the manifold has been selected by a suitably long gradient ascent, we perform a Monte Carlo calculation to determine observables of interest via Eq. (1). The imaginary part of both the action and the log of the Jacobian determinant must be included in the reweighting. Since we chose a manifold of integration for which the Jacobian is diagonal, the Monte Carlo sampling proceeds as quickly as it would for a standard Metropolis running on \mathbb{R}^N . There

are no constraints on the observables computed, aside from the requirement that $\mathcal{O}e^{-S}$ be holomorphic.

IV. RESULTS

We choose bare parameters g and m of the action so that the renormalized particle masses lie below the lattice cutoff scale. We measure two particle masses—a fermion mass am_f and a boson mass am_b —by fitting the large-time behavior of correlators $\langle \mathcal{O}_\alpha(t) \mathcal{O}_\alpha(0)^\dagger \rangle$, where $\mathcal{O}_f = \psi_1$ and $\mathcal{O}_b = \bar{\psi}_i \gamma_5 (\tau_3)_{ij} \psi_j$ (the fermion subscripts denote flavor). For simulations in this paper, we take $g = 1.0$ and $m = -0.25$, leading to renormalized masses of $am_f = 0.30(1)$ and $am_b = 0.44(1)$. We then have $m_b/m_f = 1.5(2)$, corresponding to a strongly coupled theory since the binding energy of the boson is comparable to the rest mass of the constituent fermions.

We perform calculations on two lattice sizes: $N_t \times N_x = 20 \times 10$ and 40×10 . The maximization of the sign average is done using a step size $\eta = 10^{-4}$. This step size was determined by starting with a large η , where the optimization process exhibited oscillatory behavior, and then reducing it until the process becomes smooth. We only tuned it on the most demanding ensemble, the 40×10 ensemble with the largest chemical potential, and used the same value for all other ensembles. The optimization process is stopped when the λ parameters converge, that is, when the gradient in Eq. (5) becomes too small. The manifold parameters determined by the optimization procedure are shown in Fig. 1. For both lattice sizes, the chosen parameters appear to be nearly continuous functions of μ/m_f . This suggests a simple optimization going forward: perform the gradient ascent at a small number of values of μ , and interpolate to determine the manifold of integration for all other desired chemical potentials. Another option is to use, as a starting point for the optimization process, the values determined for a “nearby” ensemble, one with similar chemical potential.

The fact that the interpolated values of λ_0, λ_1 are approximate and not strictly optimal affects only the efficiency of the algorithm, not its correctness. For more elaborate families of manifolds, with more parameters, the gradient ascent phase becomes more time-consuming. This optimization could be computationally expensive in such cases, but optimizations along the lines suggested above are likely to be available. We note that the discontinuities in Fig. 1 are due to an early exit from the optimization loop. We decided to keep these parameters to show that this discontinuity is not reflected in the observables, as a further check of the method.

With the parameters determined above, we performed a Monte Carlo calculation generating of the order of 2000 to 10 000 independent configurations (except for a few points discussed below). The average sign and measurements of the average fermion density (per flavor) $\langle n \rangle$

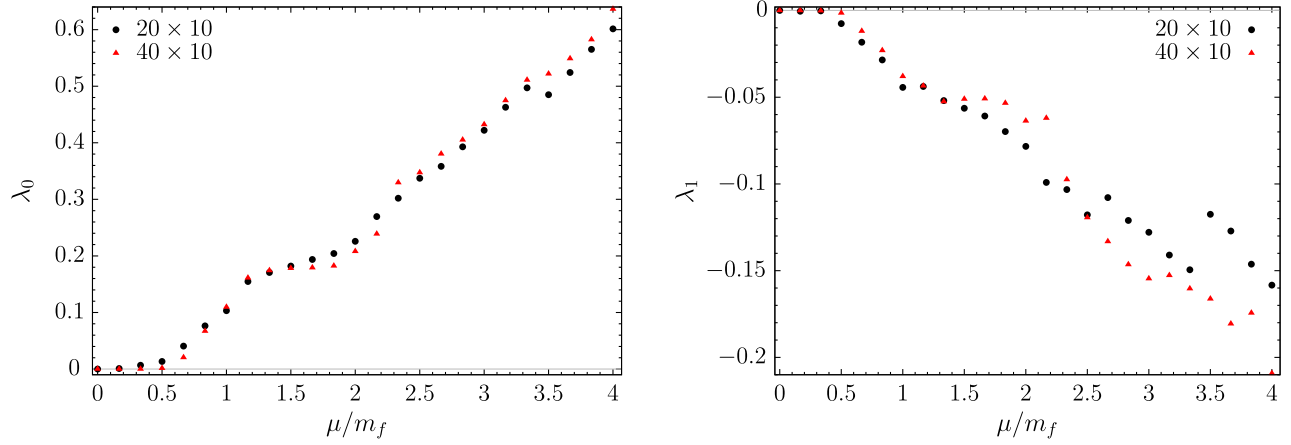


FIG. 1. Manifold parameters as a function of chemical potential μ/m_f , for both 20×10 and 40×10 lattices, with bare parameters $m = -0.25$ and $g = 1$.

for $0 < \mu/m_f < 4$ on a 20×10 lattice are shown in Fig. 2. The real plane (\mathbb{R}^N) calculations are shown in black; data points for which the average sign could not be distinguished from 0 at 2σ (indicating that no measured observable will be meaningful) are grayed out. Calculations on the tangent plane of the dominant Lefschetz thimble $A_0(x) + i\mathbf{A}$ are shown in red, and those of the machine-learned learnifold \mathcal{L}_T from Ref. [22] are shown in blue. Finally, we present calculations done on the sign-optimized manifold \mathcal{M}_S in green. We see that the sign-optimized manifold finds an average sign problem as good as or better than the learnifold does, with the added benefits of being computationally faster and simpler to implement. These improvements allow for us to compute the density with reduced uncertainty and even reach higher values of μ/m_f . As a further check of our results, we show the result for noninteracting fermions (with the same renormalized mass) as a dotted line.

Similarly, results for a lattice size of 40×10 are shown in Fig. 3. On this larger lattice, the relative performance of

the sign-optimized manifold is moderately improved. For $\mu/m_f > 1.83$, neither the real plane calculation nor the learnifold could resolve the sign problem. The sign-optimized manifold has sufficiently large average signs to allow us to measure the density up to $\mu/m_f = 2.50$. Furthermore, other methods had problems computing the density near $\mu/m_f \approx 1.00$, while using the sign-optimized manifold, we compute the density at this point easily.

At both lattice sizes, we demonstrate that the sign-optimized manifold method is capable of reproducing the ‘‘Silver Blaze’’ phenomenon [35]: the μ -independence of observables below the threshold chemical potential $\mu \approx m_f$.

To estimate the speedup given by the optimized manifold over a naive calculation on \mathbb{R}^N , we performed two tests. First, we used 72 000 decorrelated measurements at $\mu/m_f = 3.33$ on a 20×10 lattice. This number of measurements is not enough to resolve the sign average from zero, so we obtain only a lower bound on the speedup attributable to using \mathcal{M}_S . We find that the real plane has an average sign of 0.002 ± 0.003 , whereas the \mathcal{M}_S has an

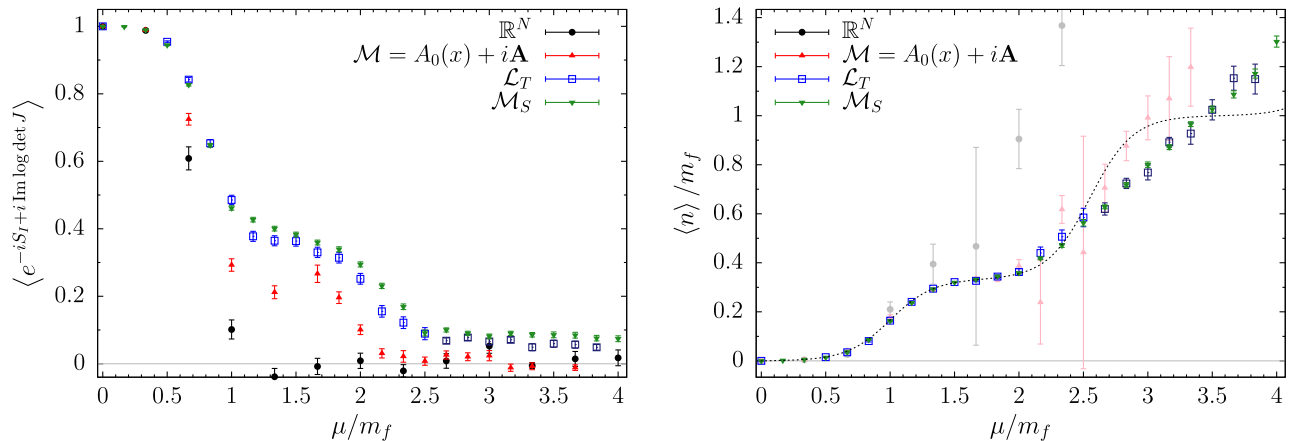


FIG. 2. $\langle e^{-iS_f + i\text{Im} \log \det J} \rangle$ and $\langle n \rangle / m_f$ as a function of μ/m_f for Wilson fermions on lattices of size 20×10 . The dashed curve represents the free fermion gas with the same mass.

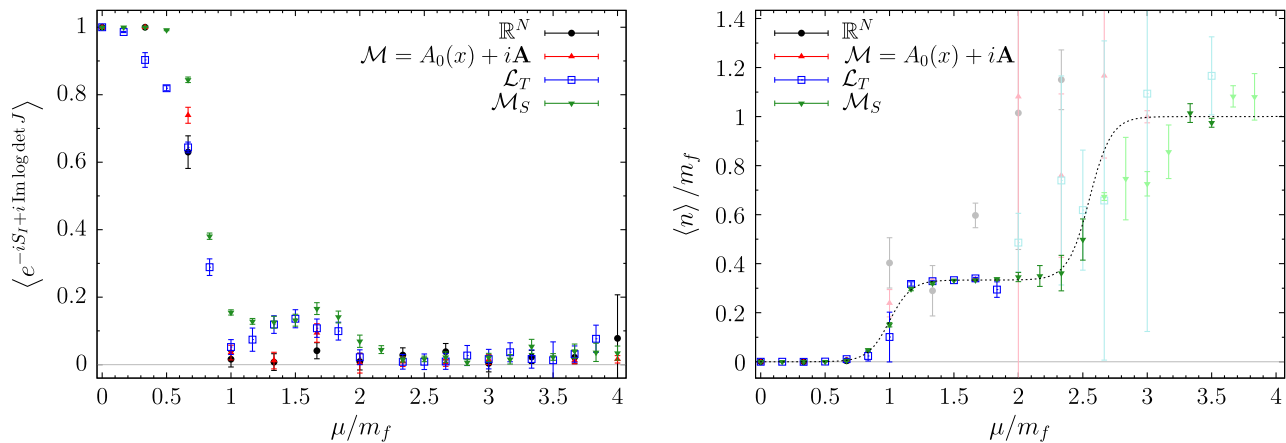


FIG. 3. $\langle e^{-iS_f + i \text{Im} \log \det J} \rangle$ and $\langle n \rangle / m_f$ as a function of μ / m_f for Wilson fermions on lattices of size 40×10 . The dashed curve represents the free fermion gas with the same mass.

average sign of 0.086 ± 0.007 , which is larger by at least a factor of 16. The number of measurements required to obtain a fixed precision is proportional $\langle \sigma \rangle^{-2}$; therefore, this corresponds to a speedup greater than 250. The second test computed 10 000 measurements on \mathbb{R}^N at $\mu / m_f = 1.00$ on a 40×10 lattice (this value is of interest because it corresponds to the first particle threshold). The real plane was found to have an average sign of 0.005 ± 0.005 , but an average sign of 0.155 ± 0.007 on the optimized manifold, which is larger by at least a factor of 15, giving a speedup of 225.

The speedup given by this algorithm over the learnifold procedure is more difficult to estimate; however, the learnifold procedure requires evolving the holomorphic flow equations many times to achieve at best the same average sign as \mathcal{M}_S . According to Ref. [22], generating the training set and training the neural network took 114 CPU hours for a 20×10 lattice at $\mu / m_f = 3.83$. The algorithm described in this paper replaces that step with a gradient ascent routine, which took approximately 24 CPU hours, and is amenable to further optimization.

V. DISCUSSION AND PROSPECTS

We have exhibited an efficient method for reducing the sign problem of the finite density Thirring model in $1 + 1$ dimensions. Our method works with a predetermined family of manifolds, seeking the manifold in that family which has the largest average sign. Once such a manifold has been found, a standard Metropolis calculation, with reweighting, is performed on that manifold. Using this method, we have increased the μ / m_f range that can reliably be computed. From the two lattice sizes computed here, it is not possible to determine the large-volume scaling of this method. It is important to stress that comparisons with other methods of dealing with the sign problem must take into account that the computational cost of the method presented here has both a fixed cost (independent of the

number of measurements made) and a variable one (that is proportional to the number of measurements). The variable cost compares very favorably with other methods, especially the generalized thimble method. Therefore, even where the optimization step is found to be expensive, one may determine the parameters of \mathcal{M} roughly, so the average sign is distinguishable from zero but not necessarily particularly close to one. Then, a high number of measurements can be made cheaply to reduce the error bars.

This method is closely related to previous approaches based on the complexification of lattice degrees of freedom, but works without evolving a differential equation to determine the manifold of integration. This makes it faster as long as the parameters defining the manifold can be determined quickly. The method has the drawback that it requires the construction of a model-specific family of manifolds, so physical insight is required. Nevertheless, given such an ansatz, the method is very advantageous. This suggests that theoretical effort should be put into generating ansätze applicable to more interesting physical theories, like gauge theories—where the set of reasonable integration contours is dramatically constrained by symmetry considerations—and real-time (Minkowski space) calculations of other models.

ACKNOWLEDGMENTS

A. A. is supported in part by the National Science Foundation CAREER Grant No. PHY-1151648 and by U.S. Department of Energy Grant No. DE-FG02-95ER40907. A. A. gratefully acknowledges the hospitality of the Physics Departments at the Universities of Maryland and Kentucky, and the Albert Einstein Center at the University of Bern where part of this work was carried out. P. F. B., H. L., and S. L. are supported by the U.S. Department of Energy under Contract No. DE-FG02-93ER-40762.

- [1] G. Aarts and I.-O. Stamatescu, *J. High Energy Phys.* **09** (2008) 018.
- [2] K. Langfeld and B. Lucini, *Acta Phys. Pol. B Proc. Suppl.* **9**, 503 (2016).
- [3] A. Alexandru, M. Faber, I. Horvath, and K.-F. Liu, *Phys. Rev. D* **72**, 114513 (2005).
- [4] P. de Forcrand and S. Kratochvila, *Nucl. Phys. B Proc. Suppl.* **153**, 62 (2006).
- [5] Z. Fodor and S. D. Katz, *Phys. Lett. B* **534**, 87 (2002).
- [6] C. R. Allton, S. Ejiri, S. J. Hands, O. Kaczmarek, F. Karsch, E. Laermann, C. Schmidt, and L. Scorzato, *Phys. Rev. D* **66**, 074507 (2002).
- [7] S. Chandrasekharan, *Eur. Phys. J. A* **49**, 90 (2013).
- [8] C. Gattringer and A. Schmidt, *Phys. Rev. D* **86**, 094506 (2012).
- [9] P. de Forcrand and O. Philipsen, *J. High Energy Phys.* **01** (2007) 077.
- [10] M. Cristoforetti, F. Di Renzo, and L. Scorzato (Aurora Science), *Phys. Rev. D* **86**, 074506 (2012).
- [11] M. Cristoforetti, F. Di Renzo, A. Mukherjee, and L. Scorzato, *Proc. Sci., LATTICE2013* (**2014**) 197.
- [12] M. Cristoforetti, F. Di Renzo, G. Erucci, A. Mukherjee, C. Schmidt, L. Scorzato, and C. Torrero, *Phys. Rev. D* **89**, 114505 (2014).
- [13] F. Di Renzo and G. Erucci, *Phys. Rev. D* **92**, 085030 (2015).
- [14] A. Mukherjee, M. Cristoforetti, and L. Scorzato, *Phys. Rev. D* **88**, 051502 (2013).
- [15] H. Fujii, S. Kamata, and Y. Kikukawa, *J. High Energy Phys.* **12** (2015) 125; **09** (2016) 172(E).
- [16] K. Fukushima and Y. Tanizaki, *Prog. Theor. Exp. Phys.* (2015) 111A01.
- [17] A. Alexandru, G. Basar, and P. Bedaque, *Phys. Rev. D* **93**, 014504 (2016).
- [18] A. Alexandru, G. Basar, P. F. Bedaque, G. W. Ridgway, and N. C. Warrington, *Phys. Rev. D* **93**, 094514 (2016).
- [19] A. Alexandru, G. Basar, P. F. Bedaque, S. Vartak, and N. C. Warrington, *Phys. Rev. Lett.* **117**, 081602 (2016).
- [20] A. Alexandru, G. Basar, P. F. Bedaque, G. W. Ridgway, and N. C. Warrington, *J. High Energy Phys.* **05** (2016) 053.
- [21] A. Alexandru, G. Basar, P. Bedaque, G. W. Ridgway, and N. C. Warrington, *Phys. Rev. D* **94**, 045017 (2016).
- [22] A. Alexandru, P. F. Bedaque, H. Lamm, and S. Lawrence, *Phys. Rev. D* **96**, 094505 (2017).
- [23] A. Alexandru, G. Basar, P. F. Bedaque, G. W. Ridgway, and N. C. Warrington, *Phys. Rev. D* **95**, 014502 (2017).
- [24] A. Alexandru, G. Basar, P. F. Bedaque, and G. W. Ridgway, *Phys. Rev. D* **95**, 114501 (2017).
- [25] A. Alexandru, G. Basar, P. F. Bedaque, and N. C. Warrington, *Phys. Rev. D* **96**, 034513 (2017).
- [26] I. Aniceto, G. Başar, and R. Schiappa, *arXiv:1802.10441*.
- [27] Y. Tanizaki and M. Tachibana, *J. High Energy Phys.* **02** (2017) 081.
- [28] Y. Tanizaki, Y. Hidaka, and T. Hayata, *Proc. Sci., LATTICE2016* (**2016**) 030.
- [29] Y. Tanizaki, Y. Hidaka, and T. Hayata, *New J. Phys.* **18**, 033002 (2016).
- [30] T. Kanazawa and Y. Tanizaki, *J. High Energy Phys.* **03** (2015) 044.
- [31] C. Schmidt and F. Ziesché, *Proc. Sci., LATTICE2016* (**2017**) 076.
- [32] M. Fukuma and N. Umeda, *Prog. Theor. Exp. Phys.* (2017) 073B01.
- [33] Y. Mori, K. Kashiwa, and A. Ohnishi, *Prog. Theor. Exp. Phys.* (2018) 023B04.
- [34] Y. Mori, K. Kashiwa, and A. Ohnishi, *Phys. Rev. D* **96**, 111501 (2017).
- [35] T. D. Cohen, *Phys. Rev. Lett.* **91**, 222001 (2003).