

This is the accepted manuscript made available via CHORUS. The article has been published as:

## Deep learning beyond Lefschetz thimbles

Andrei Alexandru, Paulo F. Bedaque, Henry Lamm, and Scott Lawrence

Phys. Rev. D **96**, 094505 — Published 10 November 2017

DOI: [10.1103/PhysRevD.96.094505](https://doi.org/10.1103/PhysRevD.96.094505)

# Deep Learning Beyond Lefschetz Thimbles

Andrei Alexandru,<sup>1,2,3,\*</sup> Paulo F. Bedaque,<sup>2,†</sup> Henry Lamm,<sup>2,‡</sup> and Scott Lawrence<sup>2,§</sup>

<sup>1</sup>*Department of Physics, The George Washington University, Washington, D.C. 20052, USA*

<sup>2</sup>*Department of Physics, University of Maryland, College Park, MD 20742, USA*

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

The generalized thimble method to treat field theories with sign problems requires repeatedly solving the computationally-expensive holomorphic flow equations. We present a machine learning technique to bypass this problem. The central idea is to obtain a few field configurations via the flow equations to train a feed-forward neural network. The trained network defines a new manifold of integration which reduces the sign problem and can be rapidly sampled. We present results for the  $1+1$  dimensional Thirring model with Wilson fermions on sizable lattices. In addition to the gain in speed, the parameterization of the integration manifold we use avoids the “trapping” of Monte Carlo chains which plagues large-flow calculations, a considerable shortcoming of the previous attempts.

## I. INTRODUCTION

Monte Carlo methods are widely used in the study of field theoretical and many-body systems. They can be understood as a way of computing a large dimensional integral encoding the physics of the system through the importance sampling of the integrand. In field theories, this integral is a discretized version of the Feynman path integral. The Monte Carlo method is essentially the only general purpose method capable of dealing with strongly interacting field theories. Unfortunately, many systems of great interest are defined by path integrals where the integrand oscillates wildly, making a direct stochastic estimation impossible in practice. This situation is referred to as the “sign problem”. The class of systems with a severe sign problem includes most finite-density systems, among them QCD, condensed matter models (e.g. the Hubbard model away from half-filling), and all real-time calculations. In the context of QCD, many ideas aiming at solving the sign problem have been developed through the years, among them the complex Langevin method [1], the density of states method [2], canonical methods [3, 4], reweighting methods [5], series expansion on the chemical potential [6], fermion bags [7] and analytic continuation from imaginary chemical potentials [8].

In the last few years the thimble approach [9, 10] has received a lot of attention. The main idea of this method is to deform the region of integration of the path integral from the original real fields to some other manifold,  $\mathcal{M}$ , embedded in the space of complexified fields. This deformation, if care is exercised, does not change the value of the integral, thanks to a multidimensional generalization of the Cauchy theorem of complex analysis. If  $\mathcal{M}$  is properly chosen, the sign problem can be solved or, at least, substantially reduced.

In the first attempts,  $\mathcal{M}$  was chosen to be the set of Lefschetz thimbles, which are the multidimensional generalization of the “steepest descent” or “constant phase” paths of complex analysis [11–16]. The phase of the (Euclidean) action, and consequently of the Boltzmann factor  $e^{-S}$ , is constant over one thimble<sup>1</sup>, so the sign problem is ameliorated. Two complications immediately appear. The first is that the thimbles cannot be found analytically in non-trivial models so an algorithm to compute them is necessary to perform the integral. This is a non-trivial task, but a few proposals have been put forward [11, 15, 16]. The second complication is that it is very hard to determine which combination of thimbles equals the original integral over real fields. It has been conjectured that one thimble dominates the path integral in the thermodynamic and/or continuum limits [9] while, at the same time, the importance of multiple thimbles has been demonstrated in finite volume systems [17, 18].

A new method, inspired by the thimble approach, was suggested in [17]. The deformation of the manifold of integration  $\mathcal{M}_T$  is chosen to be the result of evolving every point  $\zeta$  of the original integration region (the set of real fields, identified with  $\mathbb{R}^N \subset \mathbb{C}^N$ ) through the holomorphic flow equations by a “time”  $T$ :

$$\frac{d\phi_i}{dt} = \frac{\partial \bar{S}}{\partial \phi_i}, \quad \phi_i(0) = \zeta_i, \quad (1)$$

where  $S$  is the (Euclidean) action of the model and the bar denotes complex conjugation. The point  $\tilde{\phi}$  in  $\mathcal{M}_T$  corresponding to  $\zeta$  is given by  $\tilde{\phi}(\zeta) = \phi(T)$ . As the flow time  $T$  increases, the set of all points thus obtained ( $\mathcal{M}_T$ ) approaches the right combination of thimbles which equals the original integral over  $\mathbb{R}^N$ . At intermediate values of flow time  $T$ ,  $\mathcal{M}_T$  differs from the thimbles and the sign problem is not completely solved but ameliorated. For this reason, the method introduced in [17] is

---

\* aalexan@gwu.edu

† bedaque@umd.edu

‡ hlamm@umd.edu

§ srl@umd.edu

---

<sup>1</sup> An additional contribution to the integral from the curvature of the thimble can be nonzero, but this “residual phase” is typically found to be small.

sometimes called the “generalized thimble” approach [19]. This approach has the advantage that neither *a priori* information about the correct combination of thimbles equivalent to the integration over real fields, nor their location and shape, is required. This method was demonstrated in two dimensional models both in Euclidean time [20] and in real time (Minkowski space) [20]. However, as the method was applied to larger systems two complications became apparent. They are a consequence of the fact that, in some models and parameter values, the sign problem is alleviated only after considerable flow. The first problem is that large flow times are computationally expensive. The second is that the probability distribution over  $\mathbb{R}^N$  — induced by the probability distribution over  $\mathcal{M}_T$  and the parameterization  $\phi(\zeta)$  — typically becomes strongly multimodal, with wide barriers of low probability separating isolated regions of high probability. Such a distribution is difficult to properly sample through local updates. Although methods to deal with the multimodality have been proposed [21, 22], they significantly increase the computational cost of the calculation.

This paper introduces a substantial add-on to the generalized thimble method which addresses the two shortcomings described above. The main idea is to use a parameterized form of the manifold  $\mathcal{M}_T$  obtained by interpolating between some (complex) fields obtained from evolving real configurations by Eq. (1). This multidimensional interpolation is a complex non-linear regression problem, a very non-trivial task. We approach it using machine learning techniques. More precisely, we will use a feed-forward network which inputs a real configuration  $\phi_R$  and outputs the corresponding imaginary part on the integration manifold:

$$\phi_R \rightarrow \tilde{\phi} = \phi_R + i\tilde{f}(\phi_R), \quad (2)$$

where the function  $\tilde{f}$  is implemented using the feed-forward network. The network is “trained” in such a way that the set of complex fields  $\tilde{\phi}$  obtained by running all real  $\phi$ , the “learnifold”  $\mathcal{L}_T$ , approximates the flowed manifold  $\mathcal{M}_T$ . The advantage of using the network to generate the configurations is that it bypasses the computationally expensive, repeated solution of Eq. (1) (and the even more expensive computation of a Jacobian, see below). In addition, the parameterization  $\tilde{\phi} = \phi_R + i\tilde{f}(\phi_R)$  has better properties compared to the one used previously regarding the multimodality problem, as explained below.

We will review the generalized thimble method in Sec. II and the use of feed-forward network methods in Sec. III. The specifics of the learnifold will be covered in Sec. IV. In Sec. V we discuss the 1+1 dimensional Thirring model, which we will use to test and demonstrate our method. The details of the simulations are presented in Sec. VI and results are presented in Sec. VII. Conclusions are summarized in Sec. VIII.

## II. GENERALIZED THIMBLE METHOD

The expectation value of an observable in field theory can be cast in the form of a path integral

$$\langle \mathcal{O} \rangle = \frac{\int_{\mathbb{R}^N} D\phi e^{-S[\phi]} \mathcal{O}[\phi]}{\int_{\mathbb{R}^N} D\phi e^{-S[\phi]}}, \quad (3)$$

where  $S[\phi]$  is the (Euclidean) action. The stochastic evaluation of this ratio is accomplished by approximating it by

$$\langle \mathcal{O} \rangle \approx \frac{1}{N_s} \sum_{\phi^{(s)}} \mathcal{O}[\phi^{(s)}], \quad (4)$$

where the configurations  $\phi^{(s)}$  are sampled randomly from the distribution  $p[\phi] \sim e^{-S[\phi]}$ . The exact value is approached as the number of configurations  $N_s$  grows.

This method works if  $S[\phi]$  is real; otherwise the sampling can be done with respect to  $\sim e^{-\text{Re } S[\phi]}$  and the phase of the integrand included when computing observables by reweighting

$$\begin{aligned} \langle \mathcal{O} \rangle &= \frac{\int_{\mathbb{R}^N} D\phi e^{-S[\phi]} \mathcal{O}[\phi]}{\int_{\mathbb{R}^N} D\phi e^{-S[\phi]}} \\ &= \frac{\int_{\mathbb{R}^N} D\phi e^{-S_R[\phi]} \mathcal{O}[\phi] e^{-iS_I[\phi]}}{\int_{\mathbb{R}^N} D\phi e^{-S_R[\phi]}} \frac{\int_{\mathbb{R}^N} D\phi e^{-S_R[\phi]} e^{-iS_I[\phi]}}{\int_{\mathbb{R}^N} D\phi e^{-S_R[\phi]} e^{-iS_I[\phi]}} \\ &= \frac{\langle \mathcal{O} e^{-iS_I} \rangle_{S_R}}{\langle e^{-iS_I} \rangle_{S_R}}. \end{aligned} \quad (5)$$

with  $S_R = \text{Re } S$ ,  $S_I = \text{Im } S$ . This procedure is practical only if the average phase  $\langle e^{-iS_I} \rangle_{S_R}$  is large. Otherwise the calculated expectation value will result from a ratio of small numbers, each resulting from detailed cancellations among configurations. In most theories the average sign is expected to vanish exponentially as the volume increases and/or the temperature decreases. This is the sign problem.

The idea of the generalized thimble method is to deform the domain of integration from  $\mathbb{R}^N$  to  $\mathcal{M}_T$ , a submanifold of  $\mathbb{C}^N$  of (real) dimension  $N$ :

$$\begin{aligned} \langle \mathcal{O} \rangle &= \frac{\int_{\mathcal{M}_T} D\tilde{\phi} e^{-S[\tilde{\phi}]} \mathcal{O}[\tilde{\phi}]}{\int_{\mathcal{M}_T} D\tilde{\phi} e^{-S[\tilde{\phi}]}} \\ &= \frac{\int_{\mathbb{R}^N} D\zeta \det J(\zeta) e^{-S[\phi(\zeta)]} \mathcal{O}[\phi(\zeta)]}{\int_{\mathbb{R}^N} D\zeta \det J(\zeta) e^{-S[\tilde{\phi}]} } \end{aligned} \quad (6)$$

where  $S[\tilde{\phi}]$  is the analytic continuation of the action to complex values of the field,  $\tilde{\phi}_i(\zeta_j)$  is a parameterization of the manifold  $\mathcal{M}_T$  by the  $N$  real variables  $\zeta_j$  and  $J_{ij} \equiv (\partial \tilde{\phi}_i / \partial \zeta_j)$  is the Jacobian relating  $\tilde{\phi}_i$  to  $\zeta_j$ .

For our purposes, there are three conditions sufficient to ensure the equality of Eqs. (5) and (6). First, the integrand should be holomorphic (i.e., complex analytic with no poles in the complexified domain), which is the

case for most observables in quantum field theories. Second, the original and final manifolds must be homotopic — that is, there should be a continuous family of manifolds connecting them. Lastly, the domain of integration should be compact<sup>2</sup>. In the particular case of the Thirring model, all degrees of freedom are periodic, so that the original domain of integration is not  $\mathbb{R}^N$  but  $(S^1)^N = T^N$  (an  $N$ -torus) which, upon complexification, becomes  $(S^1 \times \mathbb{R})^N = T^N \times \mathbb{R}^N$ .

The calculation of the integral over  $\mathcal{M}_T$  requires a parameterization of the manifold. Previously the pre-image  $\zeta$  of the point  $\phi$  on  $\mathcal{M}_T$  was used [20]. In this case, the corresponding Jacobian can be calculated by solving

$$\frac{dJ_{ij}}{dt} = \frac{\partial^2 S}{\partial \zeta_i \partial \zeta_k} \bar{J}_{kj}, \quad J(0) = \mathbb{1}. \quad (7)$$

Evolving  $J$  according to Eq. (7) is the most computationally expensive part of the whole method. Since  $\det J$  and  $S$  over  $\mathcal{M}_T$  are complex, reweighting is done using:

$$\begin{aligned} \langle \mathcal{O} \rangle &= \frac{\int_{\mathbb{R}^N} D\zeta \det J(\zeta) e^{-S} \mathcal{O}}{\int_{\mathbb{R}^N} D\zeta \det J(\zeta) e^{-S}} \\ &= \frac{\int_{\mathbb{R}^N} D\zeta e^{-S_R + \text{Re log}(\det J)} e^{-iS_I + i \text{Im log}(\det J)} \mathcal{O}}{\int_{\mathbb{R}^N} D\zeta e^{-S_R + \text{Re log}(\det J)}} \\ &\quad \times \frac{\int_{\mathbb{R}^N} D\zeta e^{-S_R + \text{Re log}(\det J)}}{\int_{\mathbb{R}^N} D\zeta e^{-S_R + \text{Re log}(\det J)} e^{-iS_I + i \text{Im log}(\det J)}} \\ &= \frac{\langle \mathcal{O} e^{-iS_I + i \text{Im log}(\det J)} \rangle_{S_{\text{eff}}}}{\langle e^{-iS_I + i \text{Im log}(\det J)} \rangle_{S_{\text{eff}}}}, \end{aligned} \quad (8)$$

where  $S_{\text{eff}} = S_R - \text{Re log}(\det J)$ . The phase  $e^{-iS_I}$  is constant over each thimble so it typically fluctuates little over  $\mathcal{M}_T$  for large enough flow time. Experience shows that  $e^{i \text{Im log}(\det J)}$  also fluctuates little for problems of interest.

The parameterization of a point  $\tilde{\phi}(\zeta)$  of  $\mathcal{M}_T$  by its pre-image  $\zeta$  is problematic. This is because the regions in  $\zeta$ -space with large probability shrink with increasing flow time while the distance between their centers stay fixed. The resulting probability distribution is strongly multimodal and difficult to sample via a Monte Carlo chain with small updates. The difficulty arises because a single Metropolis proposal is unlikely to “tunnel” between such regions and therefore is “trapped”.

### III. FEED-FORWARD NETWORKS

In this section we summarize the use of feed-forward networks for interpolation purposes. In the next one

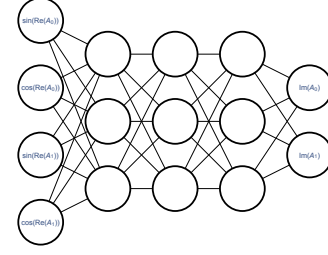


FIG. 1. Graphical representation of a feed-forward network as used in the paper. The input layer is on the left, the output layer on the right.

we will detail the application of it to constructing the “learnifold”. Feed-forward neural networks provide a family of functions particularly amenable to non-linear regression, and we use them to represent an approximation to the flowed manifold  $\mathcal{M}_T$ . A feed-forward network may be thought of as a directed graph organized into several layers which we sketch in Fig. 1. All edges point from a node in one layer to a node in the next layer. The first layer, termed the “input layer”, has exactly as many nodes as the function has inputs. Similarly, each node of the last layer, termed the “output layer”, corresponds to a different degree of freedom in the output of the function. There are no restrictions on the number of nodes in the intermediate (“hidden”) layers, although for simplicity, the networks we use in this paper have the same number of nodes in each hidden layer. To the edge from node  $i$  to node  $j$  is assigned the weight  $w_{ij}$ , and to each non-input node  $j$  is assigned a bias  $b_j$ . For a fixed topology — number of layers and nodes — the weights and biases parameterize a family of functions. It is these weights and biases that will be adjusted when performing the non-linear regression.

Given a fixed assignment of weights and biases, the network represents a function  $\vec{f}(\vec{x})$ , where  $\vec{x}$  has as many components as the network has input nodes, and  $\vec{f}$  has as many components as the network has output nodes. This function is evaluated in the following manner. The input values  $\vec{x}$  are fed into the network at the input nodes. Each node  $j$  in the first hidden layer computes a linear combination of these values, weighted by  $w_{ij}$  and shifted by the bias  $b_j$ , and then applies a certain nonlinear function  $\sigma$ . The result becomes the output of that particular node that is then taken as the input in the nodes of the next layer. This process is repeated for each hidden layer and then the output layer, as values are “fed forward” through the network. Thus the value at node  $j$  is given by

$$v_j = \sigma \left( b_j + \sum_i w_{ij} v_i \right) \quad (9)$$

where the sum is taken over all nodes  $i$  that have an edge leading to node  $j$  (that is, all nodes in the previous layer). The values at the output nodes are taken to be

<sup>2</sup> This condition can be relaxed for non-periodic variables if the deformation is such that the asymptotic behavior of the integrand guarantees the existence of the integral at all intermediate steps of the deformation.

the outputs of the function  $\vec{f}(\vec{x})$ . The computation time required for  $\vec{f}(\vec{x})$  is linear in the number of nodes in the network.

There is considerable freedom in the choice of the nonlinear function  $\sigma$ . We adopt the common choice of the “SoftPlus” function:

$$\sigma(x) = \log(1 + e^x) \quad (10)$$

which asymptotically behaves like the integral of the step function, but is smoothly differentiable (making training easier). Because this function is bounded from below, but the function we wish to approximate may not be, we do not apply any nonlinear function at the output layer. The output values are a simple (shifted) linear combination of the values at the last hidden layer.

Our goal is to use the feed-forward network to interpolate a “training set”, that is, a set of vectors  $x^{(h)}$  and their corresponding  $\vec{y}^{(h)} = \vec{f}(\vec{x}^{(h)})$  that are assumed to be known. For that we minimize a “cost function”

$$C(w, b) = \frac{1}{N} \sum_{h=1}^N \left| \vec{f}_{w,b}(\vec{x}^{(h)}) - \vec{y}^{(h)} \right|. \quad (11)$$

in relation to the biases and weights ( $N$  is the size of the training set). The minimization procedure is a simple gradient descent algorithm. For a network with  $N_I$  inputs,  $K$  hidden layers,  $N_H$  nodes in each hidden layer, and  $N_O$  outputs, there are  $N_H(N_I + N_O) + (K-1)N_H^2 + KN_H + N_O$  weights and biases to be adjusted, and therefore that many dimensions to be explored during gradient descent. The gradient of the cost function is efficiently calculated through repeated use of the chain rule. Starting at the *output* end of the network, we compute the gradient of the cost function with respect to the output of each node. This step is known as backpropagation.

$$\frac{\partial C}{\partial v_i} = \sum_j \frac{\partial C}{\partial v_j} w_{ij} \sigma'(v_j) \quad (12)$$

Once backpropagation is complete, the gradient of the cost function with respect to the weights and biases is immediately determined:

$$\frac{\partial C}{\partial w_{ij}} = \frac{\partial C}{\partial v_j} v_i \sigma'(v_j) \quad \text{and} \quad \frac{\partial C}{\partial b_j} = \frac{\partial C}{\partial v_j} \sigma'(v_j). \quad (13)$$

Like the evaluation of the function itself, the determination of the gradient of the cost function (once the gradient with respect to the values of the output nodes is known) is linear in the number of nodes in the network.

The minimization of  $C(w, b)$  is a tricky problem due to the existence of many local minima and an extensive literature is dedicated to this problem (a good review of modern gradient-descent methods is given in [23]). We use the Adaptive Moment Estimate algorithm (ADAM) [24], which was found to perform best among the methods tried. In ADAM, the weights and biases are repeatedly updated according to the descent rule

$$w_{s+1} = w_s - \frac{\tilde{\eta}_s}{1 - \alpha^s} (\tilde{\nabla} C)_s, \quad (14)$$

where  $w_s$  collectively denotes the weights and biases at iteration  $s$  of the algorithm,  $\tilde{\nabla} C$  is a modified gradient of the cost function with respect to  $w$ , and  $\tilde{\eta}$  is the dynamical learning rate that determines how far along the gradient to progress. The difference in  $\tilde{\nabla} C$  is the inclusion of “momentum” by a decaying average of previous gradients which decreases the steps needed to reach a minima by encouraging the descent to follow the largest gradient:

$$(\tilde{\nabla} C)_s = [(1 - \alpha)(\nabla C)_s + \alpha(\tilde{\nabla} C)_{s-1}]. \quad (15)$$

We set the weighting between the current and previous gradient to be  $\alpha = 0.9$  following the suggestion of [23]. The prefactor of  $(1 - \alpha^s)^{-1}$  in Eq. (14) corrects for the bias where since  $w_0$  is initialized to 0,  $w_s$  is biased towards remaining there. Once  $s$  is sufficiently large, this term goes to 1 and has negligible effect. At long times, stochastic gradient methods can oscillate around a minima, so it is useful to decrease the learning rate to reach the minima. To decrease the learning rate, we use a dynamical learning rate is defined as

$$\tilde{\eta}_s = \frac{\eta}{\sqrt{\frac{\tilde{v}_s}{(1 - \beta^s)} + \epsilon}} \quad (16)$$

where  $\eta = 5 \times 10^{-4}$  is a base learning rate, and  $\epsilon = 10^{-8}$  is a regulator to prevent numerical instability. Further dynamical improvement comes from using  $\tilde{v}_s$  which is the variance of  $\nabla C$ , correcting for bias and including a momentum:

$$\tilde{v}_s = (1 - \beta) |\nabla C|_s^2 + \beta v_{s-1} \quad (17)$$

where the weighting between current and previous terms is  $\beta = 0.999$  and we have again introduced into Eq. (16) a bias-correction factor  $(1 - \beta^s)^{-1}$ . Thus, at iteration  $s$  of ADAM, we compute the gradient of the cost function  $\nabla_w C$ , update the estimates  $(\tilde{\nabla} C)_s$  and  $v_s$ , and finally update the weights and biases. We perform  $10^6$  iterations to train a learnifold.

Computing the gradient of the cost function is computationally expensive due to the size of the training set. For example 100 configurations of  $10 \times 10$  lattices generate by translation a set of  $10^4$  configurations (see the discussion below). Instead, we use a stochastic gradient descent: we approximate the gradient at each step by a sum over a small, randomly-selected batch of the configurations. We use a batch size of 25 to begin the gradient descent, and then switch to a batch size of 200 after  $2 \times 10^5$  steps. This increase in batch size decreases the amount of stochastic noise, and the second half of the gradient descent is able to perform more fine-tuned optimizations.

#### IV. THE LEARNIFOLD

In order to avoid having to solve Eq. (1) and Eq. (7) at every step of the Monte Carlo chain, we will find another manifold  $\mathcal{L}_T$  (the “learnifold”) that approximates  $\mathcal{M}_T$ ,



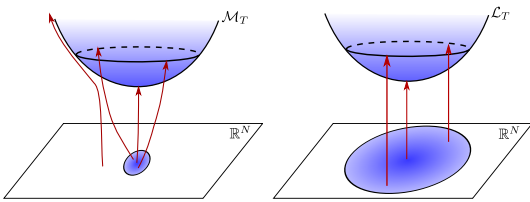


FIG. 2. Pictorial representation of parameterizations of manifolds. The standard generalized thimble method parameterizes (left) the manifold  $\mathcal{M}_T$  by its preimage on the real plane. This results in a large Jacobian because most regions flow into the singularities and a small region stretches. The learnifold  $\mathcal{L}_T$  is parameterized (right) by its real part, so the region on the real plane is larger and the barrier between thimbles are narrower.

but can be more readily computed (using a feed-forward network). Points  $\tilde{\phi}$  on the learnifold are parameterized by points  $\phi$  on the real plane:

$$\tilde{\phi}_i(\phi) = \phi_i + i\tilde{f}_i(\phi) \quad (18)$$

where the function  $\tilde{f}$  will be constructed using the kernel function  $f$  represented by a feed-forward network (see below).

Besides the gain in speed from the use of a feed-forward network in place of evolution of Eq. (1), our method differs from the one in [16, 17, 20, 22, 25, 26] by the use of a different manifold parameterization: a point in the manifold of integration is parameterized by the real part of its coordinates instead of its pre-image under the flow in Eq. (1). This new parameterization is portrayed in Fig. 2: the left-hand panel shows the parameterization arrived at from a pure-flow algorithm, and the right panel shows the parameterization from the learnifold.

This parameterization choice suffers from one drawback. Since a point  $\tilde{\phi}$  on the learnifold is parameterized by the real part  $\phi \equiv \text{Re } \tilde{\phi}$ , the learnifold will necessarily have exactly one point with any given real part. In other words, the function  $\tilde{f}$  defining the  $\mathcal{L}_T$  is single-valued. This is a restriction on the set of manifolds that can be represented by this scheme: if the flowed manifold is such that multiple points share the same real coordinates, the class of learnifolds described here may not contain a good approximation. In practice, we find that the parts of the  $\mathcal{M}_T$  that are of interest (those parts with comparatively low actions) do not behave in this way.

Despite this mild caveat, the advantages are substantial. Firstly, the parameterization of  $\mathcal{M}_T$  by the pre-image of the flow causes a small region of parameter space (shaded blue in the figure) to map to a large region of the manifold. This results in large fluctuations of  $|\det J|$ , which is expensive to compute. The parameterization of  $\mathcal{L}_T$  by the real coordinates should not lead to large stretchings. In fact, we find that  $\det J \approx 1$  in practice, so that this contribution may be accounted for after-the-fact in reweighting. The second advantage is that the parameterization reduces the multimodality problem. In  $\phi$ -space, regions of large probability do not shrink with

the flow, and so no large gaps are created between regions that contribute significantly to the integral. A Monte Carlo chain can therefore more easily explore the relevant regions of the integration domain.

Field theoretic models of interest often have a discrete group of translational symmetries on the lattice. These symmetries are respected by the flowed manifold, and therefore should be imposed on the learnifold which approximates it. Translation symmetry can be implemented in our setup in a simple way. Let  $T_i$  be the lattice translation that places lattice site  $i$  at the origin. We want that  $\tilde{\phi}(T_i\phi) = T_i\tilde{\phi}(\phi)$  which requires that  $\tilde{f}(T_i\phi) = T_i\tilde{f}(\phi)$ . A kernel function  $f : \mathbb{R}^N \rightarrow \mathbb{R}$  can be used to define a translational invariant function  $\tilde{f} : \mathbb{R}^N \rightarrow \mathbb{R}^N$  by:

$$\tilde{f}_i(\phi) = f(T_i\phi) \quad (19)$$

When multiple degrees of freedom are associated with each lattice site (for the model of interest to us, there are 2),  $f$  will have that many components. For our case we train the kernel function  $f$  to match the values of the imaginary component of the flowed configuration at origin, that is  $f(\text{Re } \tilde{\phi}) \approx (\text{Im } \tilde{\phi})_0$ , where  $\tilde{\phi}$  are the configurations from the training set generated by flowing from  $\mathbb{R}^N$ . The procedure to get  $\tilde{\phi}$  on the learnifold is to start with a configuration  $\phi$  in the real plane, and evaluate  $f(\phi)$  to determine the imaginary part of the degrees of freedom associated to the lattice site at the origin. Then, translate the lattice so that site  $i$  is moved to the origin, and evaluate  $f(T_i\phi)$  again to get the imaginary part of the degrees of freedom associated to lattice site  $i$ . This procedure is then repeated for all possible translations of the lattice.

The inputs to the network represent the real degrees of freedom  $\phi$  at each lattice site; however, in our model, these degrees of freedom are periodic. We impose this periodicity by passing to the network not  $\phi$ , but  $\sin \phi$  and  $\cos \phi$  separately — for a model with  $N$  degrees of freedom, the network will take  $2N$  inputs. This is not simply an optimization: if the learnifold lacks this periodicity, it will generically belong to a different homology class from  $\mathbb{R}^N$ , that is, it will describe a manifold of integration that is not equivalent to the original domain  $\mathbb{R}^N$ .

Implementing translation invariance as described above, a point  $\tilde{\phi}$  on the learnifold, parameterized by its real coordinates  $\phi = \text{Re } \tilde{\phi}$ , is given by

$$\tilde{\phi}_i(\phi) = \phi_i + if(T_i\phi), \quad (20)$$

where  $f(\phi)$ , computed by a feed-forward network, outputs the imaginary parts of the degrees of freedom associated to a single lattice point.

## V. THIRRING MODEL

The model we use to illustrate our method, the  $1+1$  dimensional Thirring model at finite chemical potential, has been studied before by the generalized thimble method [20]. It is defined in the continuum by the

Euclidean action

$$S = \int d^{1+1}x [\bar{\psi}^\alpha (\not{D} + \mu\gamma_0 + m)\psi^\alpha + \frac{g^2}{2N_F} \bar{\psi}^\alpha \gamma_\mu \psi^\alpha \bar{\psi}^\beta \gamma_\mu \psi^\beta], \quad (21)$$

where the flavor indices take values  $\alpha, \beta = 1, \dots, N_F$ ,  $\mu$  is the chemical potential and the Dirac spinors  $\bar{\psi}, \psi$  have two components. It is convenient to treat the four-fermion interaction by introducing an auxiliary vector field  $A_\mu$ . We use the Wilson discretization given by

$$S = \sum_{x,\nu} \frac{N_F}{g^2} (1 - \cos A_\nu(x)) + \sum_{x,y} \bar{\psi}^\alpha(x) D_{xy}^W(A) \psi^\alpha(y), \quad (22)$$

with

$$D_{xy}^W = \delta_{xy} - \kappa \sum_{\nu=0,1} \left[ (1 - \gamma_\nu) e^{iA_\nu(x) + \mu\delta_{\nu 0}} \delta_{x+\nu,y} + (1 + \gamma_\nu) e^{-iA_\nu(x) - \mu\delta_{\nu 0}} \delta_{x,y+\nu} \right], \quad (23)$$

and  $\kappa = 1/(2m + 4)$ .

The integration over the fermion fields leads to

$$S = N_F \left( \frac{1}{g^2} \sum_{x,\nu} (1 - \cos A_\nu(x)) - \log \det D(A) \right). \quad (24)$$

For  $\mu \neq 0$  the determinant  $\det D(A)$  is not real so this model has a sign problem. In this work we use  $N_F = 2$ . Notice that the variables  $A_\nu(x)$  are periodic so the original (real) domain of integration of the path integral is  $(S^1)^N$  with  $N = nL_t L_s^n$ ,  $L_t, L_s$  being the temporal and spatial sizes of the lattice, respectively.

## VI. SIMULATION DETAILS

Our procedure begins by using the flow-based algorithm of [17] to generate a training set, that is, a set of points on  $\mathcal{M}_T$ . A quality training set need not sample the probability distribution  $e^{-S_R}$  but should provide information about all of  $\mathcal{M}_T$ , or at least the region likely to be sampled by a long Monte-Carlo run. A small set of training configurations sampled with  $e^{-S_R}$  turns out to be insufficient. The learnifold generated this way develops “pockets”, that is the distribution on it is multimodal and the Metropolis process becomes trapped. It is then crucial to provide additional information about  $\mathcal{M}_T$ .

Much freedom exists in generating this additional set: configurations normally thrown away during thermalization can be kept, and they can be generated in parallel. To cure the multimodality problem we need to include configurations on  $\mathcal{M}_T$  that have large  $S_R$ . For that we include in the training set configurations from an ensemble sampled from the distribution  $e^{-S_R/\tau_i}$ , with  $\tau_i \geq 1$ . Sets with  $\tau_i > 1$  sample higher- $S_R$  regions of  $\mathcal{M}_T$  than  $\tau_i = 1$ . We use  $\tau_i = 1, 2$ . It should be emphasized that the

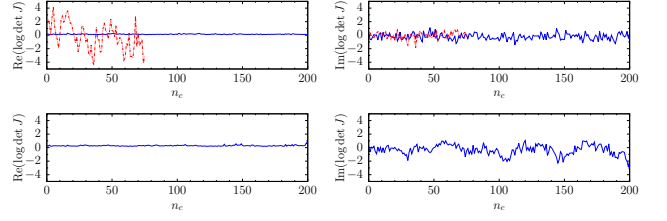


FIG. 3. The Jacobian for the learnifold for the  $10 \times 10$  ensemble (above) and for the  $40 \times 10$  lattice at  $\mu/m_f = 2.33$  (below). The left panels indicate the real part of the Jacobian and the right one its phase. In the row above we compare the fluctuations of the Jacobian of the learnifold (blue) with the Jacobian induced by the flow (red). Note the dramatic reduction in the size of the fluctuations.

ensemble generated for training is not sufficient for proper evaluation of any observable. They are not thermalized, they are highly correlated and the ones obtained with  $\tau_i \neq 1$  are not distributed correctly.

Since, as discussed above, the action is translationally invariant, a single flowed configuration can be used as a total of  $V$  training points, where  $V$  is the space-time volume of the lattice. Thus each flowed configuration sampled is translated to  $V$  other configurations. Since, in practice, the most time-consuming step of the algorithm is the generation of configurations on  $\mathcal{M}_T$ , this multiplication of the training set is critical in making this algorithm practical.

Once the training configurations are obtained, the feed-forward network is trained by minimizing the cost function. We use a network with 3 hidden layers, each consisting of 10 nodes. This choice is somewhat arbitrary and we have not yet fully investigated the behavior of the algorithm as the number of nodes and layers is changed. The training is accomplished by performing stochastic gradient descent to minimize the weights  $w_{ij}$  and biases  $b_i$  with respect to the cost function. Specializing Eq. 11 to our specific case:

$$C(w, b) = \sum_k \left[ \sum_\nu (\text{Im } \tilde{A}_\nu(0, 0) - \text{Im } A_\nu(0, 0))^2 \right]^{1/2} \quad (25)$$

where the sum over  $k$  is taken over all training points, and the sum over  $\nu$  is over the Lorentz indices. By minimizing the  $C(w, b)$ , we minimize the distance between the  $\mathcal{L}_T$  and  $\mathcal{M}_T$ .

Once the gradient descent is complete,  $f$  is used to define  $\mathcal{L}_T$  through Eq. (20). This manifold is parameterized by the real plane, and so we can perform an importance sampling on this manifold in the same manner as for a flowed manifold.

The previous parameterization, based on deforming the domain of integration via flow, required computing the Jacobian when performing importance sampling. The  $\mathcal{L}_T$  parameterization is found to result in Jacobians that have small fluctuations as can be seen from Fig. 3. Computing the Jacobian is expensive, but since its fluctuations

are small it is preferable to ignore it when performing importance sampling, and include it in observables via reweighting. The Jacobian may be computed by direct application of the chain rule. In practice, it is sufficient to compute the Jacobian via finite differencing, that is, computing  $\partial f(\phi_i)/\partial \phi_j \approx (f(\phi_i + \Delta \delta_{ij}) - f(\phi_i))/\Delta$  for small  $\Delta$  by feeding the values  $\phi_i + \Delta \delta_{ij}$  and  $\phi_i$  through the network. We take this approach here.

After the network is trained and the manifold  $\mathcal{L}_T$  defined by it is specified, we use the Metropolis algorithm applied to the (real) parameterizing variables  $\phi_i$  and the effective action  $S_{\text{eff}}(\phi) = \text{Re} S[\phi(\phi)]$ , the real part of the Euclidean action. The Jacobian and the phase of the Euclidean action are included through reweighting:

$$\langle \mathcal{O} \rangle = \frac{\langle \mathcal{O} e^{-iS_I + \log \det J} \rangle_{S_{\text{eff}}}}{\langle e^{-iS_I + \log \det J} \rangle_{S_{\text{eff}}}}, \quad (26)$$

where in contrast to Eq. (8), the real part of the Jacobian is also included in the reweighting. The minimum number of configurations from  $\mathcal{M}_T$  required for training must be empirically determined and we find it to be roughly set by the number of degrees of freedom in the fit to be performed: if the network has more degrees of freedom than the number of training points available, a long training process will overfit the data, and the final product will be unusable.

## VII. RESULTS

Although our calculations are not done particularly close to the continuum limit, we choose the bare parameters of the action so that the (renormalized) particle masses are somewhat below the lattice scale. Two particle masses have been measured: a fermion and a boson. The dimensionless masses of these particles,  $am_f$  and  $am_b$  ( $a$  is the lattice spacing) respectively, are determined via fitting the large time behavior of the correlators  $\langle \mathcal{O}_\alpha(t) \mathcal{O}_\alpha(0)^\dagger \rangle$  where  $\mathcal{O}_f = \psi_1$  and  $\mathcal{O}_b = \bar{\psi}_i \gamma_5 (\tau_3)_{ij} \psi_j$ , where the subscripts indicate flavor. In a free theory,  $m_b/m_f = 2$  and we use this ratio to gauge the strength of interactions, where  $m_b/m_f \ll 2$  implies a strongly interacting theory. The parameters used for the simulations in this paper are  $g = 1.0$  and  $m = -0.25$ , which lead to  $am_f = 0.30(1)$  and  $am_b = 0.44(1)$  [20]. We have then  $m_b/m_f = 1.5(2)$ . Therefore we are studying a strongly coupled theory.

The flow time used to generate training points and a range for the size of each partition of the training set is given for each lattice in Table I. During the generation of the training set we use an estimator of the Jacobian that has been shown to track accurately the full Jacobian [27]. Even using the estimator the generation of the training set is computationally expensive. For example the ensemble used for training the  $20 \times 10$  learnifold at  $\mu/m_f = 3.83$  we use 380 configurations with  $\tau = 1$  and 225 with  $\tau = 2$ . To generate these configurations we use no thermalization and save a configuration after 200 Metropolis steps. This

takes about 90 CPU-hours. The training of the network takes 24 CPU-hours. To generate the ensemble of 7220 measurements with 800 Metropolis steps between on the learnifold takes about 140 CPU-hours. Taking the training set time as an proxy (which neglects computing the Jacobian), this set would have taken about 4100 CPU-hours using flow. More details about the measurements can be found in Table I. It should be noted that for larger lattices at larger  $\mu/m_f$ , to achieve the same statistics more configurations are needed due to the smallness of average sign, which is reflected by a range being reported for the number of measurements.

In Fig. 4, we show for lattices of size  $N_t \times N_x = 10 \times 10$ ,  $20 \times 10$ , and  $40 \times 10$  the average sign  $\langle e^{-iS_I + i \text{Im} \log \det J} \rangle$  in the left column, and in the right column the average fermion density (per flavor)  $\langle n \rangle$ . The results obtained by standard reweighting techniques on  $\mathbb{R}^N$  are shown with black circles. As may be anticipated, the average sign drops to 0 as  $\mu \approx m_f$  and reweighting becomes unfeasible. We further plot with red triangles the results obtained by choosing as a manifold of integration the tangent plane to the main thimble (which is computationally as cheap as integration over  $\mathbb{R}^N$ ). For this manifold, the average sign drops more slowly to zero, but for sufficiently large lattices is still inadequate. The final set of results, the blue squares, are obtained from the learnifold. We find that the average sign decrease even more slowly, extending the reach in  $\mu/m_f$ . As a check, we include in these figures the analytic result for the free fermion gas (with the renormalized fermion mass) with a dashed line. We find that for values of  $\mu/m_f > 2.5$ , the free gas approximation becomes a poor description of the Thirring model.

As explained in [20] the sign problem is greatly improved by simply shifting the domain of integration:  $A_0(x) \rightarrow A_0(x) + i\mathbf{A}$ ,  $A_i(x) \rightarrow A_i(x)$  for a certain real value of  $\mathbf{A}$ . This simple shift is enough in  $1 + 1$  dimensions to allow for calculations in lattices of size up to  $L_t \times L_s = 40 \times 10$  if a staggered fermion action is used. However, the Wilson fermion action has a worse sign problem and systems of these sizes require longer flow times which make the calculation more expensive. More impor-

Lattice	$T$	$n_\tau$		$n_{\text{therm}}$	$n_{\text{cor}}$	$n_{\text{meas}}$
		1	2			
$10 \times 10$	0.4	1000	1000	15000	800	2000
$20 \times 10$	0.2, 0.4	300-600	100-700	15000	800	2000-8000
$40 \times 10$	0.2	50-170	30-150	20000	800	800-1000

TABLE I. Training set generation and Metropolis sampling parameters.  $T$  is the flow time. Different values of  $\mu/m_f$  generate flowed configurations at different rates, so for brevity we quote the size of the training set as a ranges.  $n_\tau$  is the size of the training set for each  $\tau_i$ , the Metropolis “temperature”.  $n_{\text{therm}}$ ,  $n_{\text{cor}}$ , and  $n_{\text{meas}}$  are the thermalization, decorrelation lengths and number of measurements respectively in the Metropolis sampling.



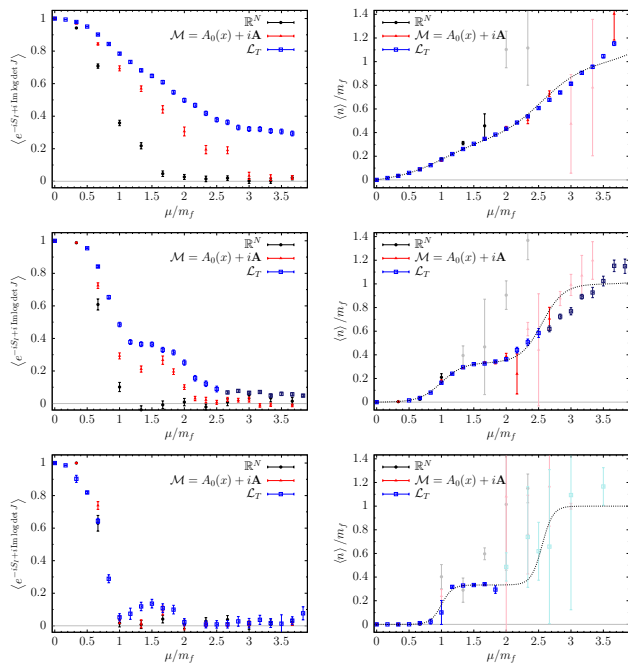


FIG. 4.  $\langle e^{-iS_I + i \text{Im log det } J} \rangle$  and  $\langle n \rangle / m_f$  as a function of  $\mu / m_f$  for Wilson fermions on lattices of size (top)  $10 \times 10$ , (center)  $20 \times 10$ , (bottom)  $40 \times 10$  with  $am_f = 0.30(1)$ . The dashed curve represents the free fermion gas with the same mass. The darker points in the  $20 \times 10$  graphs (middle row) correspond to a learnifold trained on  $\mathcal{M}_T$  with  $T_{\text{flow}} = 0.4$  whereas the lighter use  $T_{\text{flow}} = 0.2$ .

tantly, the large flow times lead to the trapping of the Monte Carlo chain as discussed above. For these reasons, Ref. [20] only contains simulations of the Wilson action in  $10 \times 10$  lattices. The use of the Wilson fermion model here demonstrates the utility of machine learning in applying the generalized thimble method to larger lattices.

On the  $20 \times 10$  lattice, the sign problem for  $\mu / m_f \leq 2.66$  is sufficiently improved for a flow time of  $T = 0.2$  that reliable results can be obtained. For  $\mu / m_f > 2.66$ , by simply flowing longer ( $T = 0.4$ ) we were able to again raise the average sign to viable levels again.

At the colder temperatures, we demonstrate that the learnifold method is capable of reproducing the independence of observables below the threshold  $\mu \approx m_f$ , the so-called “Silver Blaze” phenomenon [28]. This is important, because other treatments of the sign problem can fail and wash out the plateau. In particular, we note that Lefschetz-thimble based approaches in which only the main thimble is sampled, or generalized thimble methods where trapping occurs, are likely to fail to produce the cor-

rect features and instead produce straight lines [18]. We believe this is evidence that while our  $\tau = 1$  training sets are trapped for large flows, the learnifold parameterization reduces trapping to manageable levels in the Metropolis sampling such and the higher  $\tau$  configurations give some information about the phase on additional thimbles to keep the sign problem manageable.

## VIII. DISCUSSION AND PROSPECTS

We have presented a method, based on machine learning techniques, to bypass the computationally expensive steps in the generalized thimble attack on the sign problem. The idea is that a reduced number of field configurations obtained from solving the flow equations can be used to train a feed-forward neural network that can roughly approximate the full manifold defined by the flow. The trained network defines a new manifold, the learnifold, which is equivalent to the real space for the computation of the path integral, but where the sign problem is ameliorated. The manifold defined by the network can be sampled very fast and a large number of measurements can be easily made, enough to bypass the sign problems in models where that was not previously possible.

This paper represents a first, exploratory study of the possibility of coupling the generalized thimble method and machine learning to solve the sign problem. Results were shown for the Wilson fermion version of the  $1 + 1$  dimensional Thirring model. As a bonus feature of the method, the parameterization we used avoids the “trapping” of Monte Carlo chains which plagues some calculations with large flow times. This method is general and should be applicable to other theories of interest. The large freedom in flow time, size and bias of the training set, and number of layers and nodes in the network together suggest that this method can be further optimized for efficiency which would extend the practical range of applicability.

## ACKNOWLEDGMENTS

A.A. is supported in part by the National Science Foundation CAREER grant PHY-1151648 and by U.S. Department of Energy grant 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 U.S. Department of Energy under Contract No. DE-FG02-93ER-40762.

- 
- [1] G. Aarts and I.-O. Stamatescu, *JHEP* **09**, 018 (2008), [arXiv:0807.1597 \[hep-lat\]](#).
  - [2] K. Langfeld and B. Lucini, *Proceedings, International*

*Meeting Excited QCD 2016: Costa da Caparica, Portugal, March 6-12, 2016*, *Acta Phys. Polon. Supp.* **9**, 503 (2016), [arXiv:1606.03879 \[hep-lat\]](#).

- [3] A. Alexandru, M. Faber, I. Horvath, and K.-F. Liu, *Phys. Rev.* **D72**, 114513 (2005), [arXiv:hep-lat/0507020 \[hep-lat\]](#).
- [4] P. de Forcrand and S. Kratochvila, *Hadron physics, proceedings of the Workshop on Computational Hadron Physics, University of Cyprus, Nicosia, Cyprus, 14-17 September 2005*, *Nucl. Phys. Proc. Suppl.* **153**, 62 (2006), [62(2006)], [arXiv:hep-lat/0602024 \[hep-lat\]](#).
- [5] Z. Fodor and S. D. Katz, *Phys. Lett.* **B534**, 87 (2002), [arXiv:hep-lat/0104001 \[hep-lat\]](#).
- [6] C. R. Allton, S. Ejiri, S. J. Hands, O. Kaczmarek, F. Karsch, E. Laermann, C. Schmidt, and L. Scorzato, *Phys. Rev.* **D66**, 074507 (2002), [arXiv:hep-lat/0204010 \[hep-lat\]](#).
- [7] S. Chandrasekharan, *Eur. Phys. J.* **A49**, 90 (2013), [arXiv:1304.4900 \[hep-lat\]](#).
- [8] P. de Forcrand and O. Philipsen, *JHEP* **01**, 077 (2007), [arXiv:hep-lat/0607017 \[hep-lat\]](#).
- [9] M. Cristoforetti, F. Di Renzo, and L. Scorzato (Aurora-Science), *Phys. Rev.* **D86**, 074506 (2012), [arXiv:1205.3996 \[hep-lat\]](#).
- [10] M. Cristoforetti, F. Di Renzo, A. Mukherjee, and L. Scorzato, *Proceedings, 31st International Symposium on Lattice Field Theory (Lattice 2013): Mainz, Germany, July 29-August 3, 2013*, *PoS LATTICE2013*, 197 (2014), [arXiv:1312.1052 \[hep-lat\]](#).
- [11] M. Cristoforetti, F. Di Renzo, G. Erucci, A. Mukherjee, C. Schmidt, L. Scorzato, and C. Torrero, *Phys. Rev.* **D89**, 114505 (2014), [arXiv:1403.5637 \[hep-lat\]](#).
- [12] F. Di Renzo and G. Erucci, *Phys. Rev.* **D92**, 085030 (2015), [arXiv:1507.03858 \[hep-lat\]](#).
- [13] A. Mukherjee, M. Cristoforetti, and L. Scorzato, *Phys. Rev.* **D88**, 051502 (2013), [arXiv:1308.0233 \[physics.comp-ph\]](#).
- [14] H. Fujii, S. Kamata, and Y. Kikukawa, *JHEP* **12**, 125 (2015), [Erratum: *JHEP*09,172(2016)], [arXiv:1509.09141 \[hep-lat\]](#).
- [15] K. Fukushima and Y. Tanizaki, *PTEP* **2015**, 111A01 (2015), [arXiv:1507.07351 \[hep-th\]](#).
- [16] A. Alexandru, G. Basar, and P. Bedaque, *Phys. Rev.* **D93**, 014504 (2016), [arXiv:1510.03258 \[hep-lat\]](#).
- [17] A. Alexandru, G. Basar, P. F. Bedaque, G. W. Ridgway, and N. C. Warrington, *JHEP* **05**, 053 (2016), [arXiv:1512.08764 \[hep-lat\]](#).
- [18] Y. Tanizaki, Y. Hidaka, and T. Hayata, *New J. Phys.* **18**, 033002 (2016), [arXiv:1509.07146 \[hep-th\]](#).
- [19] J. Nishimura and S. Shimasaki, *JHEP* **06**, 023 (2017), [arXiv:1703.09409 \[hep-lat\]](#).
- [20] A. Alexandru, G. Basar, P. F. Bedaque, G. W. Ridgway, and N. C. Warrington, *Phys. Rev.* **D95**, 014502 (2017), [arXiv:1609.01730 \[hep-lat\]](#).
- [21] M. Fukuma and N. Umeda, (2017), [arXiv:1703.00861 \[hep-lat\]](#).
- [22] A. Alexandru, G. Basar, P. F. Bedaque, and N. C. Warrington, (2017), [arXiv:1703.02414 \[hep-lat\]](#).
- [23] S. Ruder, *ArXiv e-prints* (2016), [arXiv:1609.04747 \[cs.LG\]](#).
- [24] D. P. Kingma and J. Ba, *ArXiv e-prints* (2014), [arXiv:1412.6980 \[cs.LG\]](#).
- [25] A. Alexandru, G. Basar, P. F. Bedaque, S. Vartak, and N. C. Warrington, *Phys. Rev. Lett.* **117**, 081602 (2016), [arXiv:1605.08040 \[hep-lat\]](#).
- [26] A. Alexandru, G. Basar, P. F. Bedaque, and G. W. Ridgway, *Phys. Rev.* **D95**, 114501 (2017), [arXiv:1704.06404 \[hep-lat\]](#).
- [27] A. Alexandru, G. Basar, P. F. Bedaque, G. W. Ridgway, and N. C. Warrington, *Phys. Rev.* **D93**, 094514 (2016), [arXiv:1604.00956 \[hep-lat\]](#).
- [28] T. D. Cohen, *Phys. Rev. Lett.* **91**, 222001 (2003), [arXiv:hep-ph/0307089 \[hep-ph\]](#).