Metropolis Monte Carlo on the Lefschetz thimble: application to a one-plaquette model

Abhishek Mukherjee,1,2* Marco Cristofooretti,1,2† and Luigi Scorzato1,2‡

1ECT*, Villa Tambosi, I-38123 Villazzano (Trento), Italy
2LISC, Via Sommarive 18, I-38123 Povo (Trento), Italy

We propose a new algorithm based on the Metropolis sampling method to perform Monte Carlo integration for path integrals in the recently proposed formulation of quantum field theories on the Lefschetz thimble. The algorithm is based on a mapping between the curved manifold defined by the Lefschetz thimble of the full action and the flat manifold associated with the corresponding quadratic action. We discuss an explicit method to calculate the residual phase due to the curvature of the Lefschetz thimble. Finally, we apply this new algorithm to a simple one-plaquette model where our results are in perfect agreement with the analytic integration. We also show that for this system the residual phase does not represent a sign problem.

Introduction — In the path integral formulation of quantum field theory (QFT), the expectation value of observables is written as ratios of multidimensional functional integrals involving the exponential of an (effective) action, \( S \). When \( S \) is real, \( e^{-S} \) can be interpreted as a probability distribution and the functional integral can be evaluated very efficiently and accurately using stochastic methods, viz. Monte Carlo sampling (see, e.g., [1]). For large systems at low temperatures, quantum Monte Carlo is arguably the most accurate method for calculating observables, at present.

Unfortunately, systems with real actions are special cases. In general, \( S \) will be complex (although the full integral is still real), and \( e^{-S} \) cannot be interpreted as a probability distribution. In principle, one can use reweighting: the absolute value of \( e^{-S} \), i.e., \( e^{-\text{Re}S} \) as the probability weight and include \( e^{-\text{Im}S} \) in the redefinition of the value of the observable for a given field configuration. However, reweighting is effective only if the fraction of configurations with negative weight is limited, rendering the method of little use for large systems and/or at low temperatures. This is a manifestation of the infamous ‘sign problem’ which plagues the application of Monte Carlo methods to quantum field theories.

Numerous methods have been proposed to deal with sign problem [2–4], and they have had important but partial success in particular classes of models. However, a general solution is missing, and the sign problem is a major hindrance to accurate calculation in many interesting physical systems: lattice QCD at finite density [5] or with a \( \theta \)-vacuum [6], real-time field theories [7], electronic systems [8], the repulsive Hubbard model [9], the nuclear shell model [10], polymer field theory [11], to name a few. Any new method to evade or at least mollify the sign problem in the generic situation represents an important advance.

Recently we proposed that a way to alleviate the sign problem is to use the formulation of the QFT on a Lefschetz thimble [12, 13] for the Monte Carlo integration [14, 15]. Lefschetz thimbles are many dimensional generalizations of the paths of steepest descent. By construction the imaginary part of the action remains constant on each thimble. However, because the Lefschetz thimbles are in general curved complex manifolds, we may pick up an additional residual phase due this curvature. We argued that the sign problem due to this residual phase, if present at all, should be much milder than the sign problem in the original integration domain.

The Lefschetz thimble formulation of QFT is, in principle, independent from methods used to sample field configurations on the thimble. The latter, in itself, presents a non-trivial problem due to complexity of the measure on the thimble. In previous work, we proposed an algorithm based on discretized Langevin dynamics. While the testing of the algorithm proposed in [14] is in progress, it is also worth exploring alternative algorithms to achieve the challenging goal of performing Monte Carlo simulations on a Lefschetz thimble.

In this paper, we present a different method to sample field configuration on the Lefschetz thimble, which is based on the Metropolis algorithm and uses a mapping between the Lefschetz thimble and a flat manifold associated with the corresponding quadratic action. We also discuss an explicit procedure to calculate the residual phase within this method.

We apply this method to the \( U(1) \) one-plaquette model. The integrals involved in this model are one variable integrals and can be performed analytically. However, it provides an interesting benchmark which can be seen as a limiting case of more realistic QFTs on a lattice. It is non-trivial from the point of view of a Monte Carlo integration. In fact, the complex Langevin method fails for this particular system. It also provides a case where different aspects of our methodology can be visualized quite clearly.

QFT on a Lefschetz thimble — Consider a QFT on a lattice (or any other system with a finite number of continuous degrees of freedom) defined by the action \( S(\phi) \), where \( \phi \) is a vector field whose number of components,
$n$, is equal to the number of degrees of freedom in the system. Suppose that the initial field theory is defined for real fields, i.e., the expectation value of any observable $O$ is given by,

$$\langle O \rangle = \frac{\int_D d\phi O(\phi) e^{-S(\phi)}}{\int_D d\phi e^{-S(\phi)}}$$  \quad (1)

where $D$ is the appropriate integration cycle for $S$ in the real domain $\mathbb{R}^n$. Now, consider $S$ in terms of the complexified fields, i.e., the field components $\phi_i$ are now allowed to be complex. Suppose, $S(\phi)$ is holomorphic in this complexified space and its critical points $\phi^*$ given by

$$\frac{\partial S}{\partial \phi^*} = 0$$  \quad (2)

are non-degenerate,

$$\det \left[ \frac{\partial^2 S}{\partial \phi^* \partial \phi^*} \right] \neq 0.$$  \quad (3)

Then, under suitable conditions on $S$ and $O$ (typically fulfilled in physical systems) and for a sufficiently generic choice of parameters, we have the following crucial result \cite{12, 13, 16}

$$\int_D d\phi O(\phi) e^{-S(\phi)} = \sum_{\sigma} m_{\sigma} \int_{J_{\sigma}} d\phi O(\phi) e^{-S(\phi)},$$  \quad (4)

where $m_{\sigma} \in \mathbb{Z}$ (see later). That is, an integral over the real domain $D$ is equivalent to sum of integrals over the Lefschetz thimbles $J_{\sigma}$. This result can be seen as a generalization of contour deformation in one dimension. The Lefschetz thimbles $J_{\sigma}$ associated with the critical points are many dimensional generalizations of the paths of steepest descent. The thimble $J_{\sigma}$ is defined as the union of all paths governed by,

$$\frac{d\phi}{dt} = -\frac{\partial S}{\partial \phi}$$  \quad (5)

and which end at the critical point $\phi^*$ for $\tau \to \infty$. They are hypersurfaces of real dimension $n$ embedded in the complex manifold $\mathbb{C}^n$. Here, and later, the overhead bar represents complex conjugation. In this paper we will assume that $S$ is a Morse function, i.e., it has only non-degenerate critical points.$^1$

Then, the expectation value of an observable can be written as

$$\langle O(\phi) \rangle = \frac{\sum_{\sigma} m_{\sigma} \int_{J_{\sigma}} d\phi O(\phi) e^{-S(\phi)}}{\sum_{\sigma} m_{\sigma} \int_{J_{\sigma}} d\phi e^{-S(\phi)}}.$$  \quad (6)

From the point of view of stochastic integration, the main benefit of the above formulation is that along a given thimble $J_{\tau}$, the imaginary part of the action $\Im S(\phi)$ remains constant. The only fluctuation in the complex phase comes from the residual phase due to the curvature of the thimble itself. We expect this to be a significantly milder sign problem than the original one.

The critical points of the action can be found by looking at all the solutions of Eq. (2). The integer coefficients $m_{\sigma}$ are the intersection numbers between $D$ and $K_{\sigma}$, where $K_{\sigma}$ is the unstable thimble, i.e., it is the union of all paths which are governed by Eq. (5), but go to $\phi^*$ at $\tau \to -\infty$. It is also a hypersurface of real dimension $n$. Then, $m_{\sigma}$ is simply the number of times the two hypersurfaces $D$ and $K_{\sigma}$ intersect.

We are not aware of a general method to calculate the $m_{\sigma}$ for an arbitrary QFT. But we argued in \cite{14} that only a limited set of thimbles are expected to dominate and, moreover, a single thimble is typically sufficient to regularize a QFT. However, in order to test the algorithm presented in this paper, it may be interesting to consider also the case in which we want to study more thimbles at the same time. Hence, in the rest of this paper we will keep a general $m_{\sigma}$, but we will assume that the intersection numbers $m_{\sigma}$ are known, and comment when relevant.

*Mapping the Lefschetz thimble on a flat manifold* — In the neighborhood of a non-degenerate critical point $\phi^*$, the holomorphic action function $S(\phi)$ can be written as,

$$S(\phi) = S(\phi^*) + S_G(\eta) + O(|\eta|^3)$$  \quad (7)

where the Gaussian action $S_G$ is given by,

$$S_G = \frac{1}{2} \sum_k \lambda_k \eta_k^2,$$  \quad (8)

and $\eta$ is related to $\phi$ by a (complex) linear transformation,

$$\phi_i = \phi_i^* + \sum_k w_{ki} \eta_k$$  \quad (9)

The $w_{ki}$ are components of the vectors $w_k$. We call the flat thimble associated with the Gaussian action $S_G$, the Gaussian thimble $G_\sigma$.

The $\lambda_k$ and $w_k$ can be found from the solutions of the generalized eigenvalue equation,

$$H w_k = \lambda_k \bar{w}_k.$$  \quad (10)

The elements of the hessian matrix $H$ are given by,

$$H_{ij} = \frac{\partial^2 S}{\partial \phi_i \partial \phi_j}.$$  \quad (11)

\footnote{Degenerate minima, as they typically occur in the presence of symmetries, can be either lifted or treated as discussed in \cite{14}.}

\footnote{See also \cite{17} for a different point of view, that is complementary and consistent with the one of \cite{14}.}
In practice, we find the $\lambda_k$ and the $w_k$ from the positive eigenvalues and the corresponding eigenvectors of the real symmetric $2n \times 2n$ matrix

$$
\hat{H} = \begin{pmatrix} H^R & H^I \\ H^I & -H^R \end{pmatrix}
$$

where

$$
H^R_{ij} = \frac{\partial RS}{\partial \phi_i \partial \phi_j}, \quad H^I_{ij} = -\frac{\partial RS}{\partial \phi_i \partial \phi_j}.
$$

The eigenvalues of $\hat{H}$ come in pairs $\{\pm \lambda_k\}$ with $k = 1, \ldots, n$, and the $\lambda_k$ being real and positive. Let $(u_k, v_k)$ be normalized $n$-dimensional vectors such that $(u_k^T, v_k^T)^T$ is an eigenvector of $\hat{H}$ with a positive eigenvalue $\lambda_k$. Then, the pair $\lambda_k$ and $w_k = \frac{1}{2} (u_k + iv_k)$ satisfies Eq. (10).

With this parametrization, the directions of steepest descent/ascent of $RS$ (and constant $3S$) correspond to directions where the $\eta_k$ are real. Consider, the equations of steepest descent of the variables $\eta_k$ (assumed real) for the Gaussian action $S_G$ in terms of the new parameter $r = e^{-\tau}$,

$$
\frac{d\eta_k}{dr} = \frac{1}{r} \frac{\partial S_G}{\partial \eta_k} = \frac{1}{r} \lambda_k \eta_k
$$

which yields the solution,

$$
\eta_k \propto r^{\lambda_k}.
$$

Now, we can define a mapping between the Gaussian thimble, parametrized by the vectors $\eta$, and the Lefschetz thimble, parametrized by the field $\phi$. First, we find the corresponding configuration $\xi$ at $r = \epsilon$,

$$
\xi_k = e^{\lambda_k \eta_k}.
$$

For a sufficiently small $\epsilon$, the Lefschetz thimble and the Gaussian thimble will coincide at $r = \epsilon$. Thus, the field configuration on the Lefschetz thimble at $r = \epsilon$ is given by,

$$
\phi_i(r = \epsilon) = \phi_i^0 + \sum_k w_{ki} \xi_k
$$

Using this as the boundary condition, we can now integrate the equation of steepest descent of the full action $S$ for the fields $\phi_i(r)$,

$$
\frac{d\phi_i}{dr} = \frac{1}{r} \frac{\partial S}{\partial \phi_i}
$$

from $r = \epsilon$ to 1. The field configuration at $r = 1$ is the one we seek. For brevity, we will simply denote it by $\phi$.

For a constant $\epsilon$, we have the following relation between the measures of integration

$$
\int_J d\phi = \int_{\mathbb{R}^n} \det [J^{\phi}_\eta] d\eta = \int_{\mathbb{R}^n} \left( \prod_k e^{\lambda_k} \right) \det [J^{\phi}_\xi] d\eta.
$$

The matrix $J^{\phi}_\eta$ ($J^{\phi}_\xi$) is the Jacobian of the transformation between the $\eta$ ($\xi$) and $\phi$ fields.

The matrix $J^{\phi}_\xi$ can be calculated along the path of steepest descent from the equation

$$
\frac{d}{dr} J^{\phi}_{\xi_{jk}}(r = \epsilon) = \frac{1}{r} \frac{\partial^2 S}{\partial \phi_i \partial \phi_j} J^{\phi}_{\xi}.
$$

along with the boundary condition,

$$
J^{\phi}_{\xi_{jk}}(r = \epsilon) = w_{ki}.
$$

In the limit $\epsilon \to 0$, the above procedure produces an explicit mapping between the flat Gaussian thimble and the Lefschetz thimble. In practice, it is necessary to perform calculations at a few sufficiently small values of $\epsilon$ in order to perform the extrapolation to the limiting case. For later reference, we note that setting $\epsilon = 1$, corresponds to a mapping from the Gaussian thimble to itself.

Note that Eq. (21) involves the evolution of a $N \times N$ matrix whose determinant must also be computed. The latter is expected to cost $O(N^3)$. This may be still too expensive for some models, but it is already a huge cost reduction compared to the $O(e^N)$ scaling expected in general and it should be sufficient to enable the Monte Carlo simulation of some important models, which are currently not feasible. Techniques of noise estimation of the trace (see, e.g., [18, 19]) may further reduce the cost of the computation of the determinant, but we do not consider them in this paper.

**Metropolis sampling on the Lefschetz thimble** — Given the mapping above, it is straightforward to formulate a Metropolis algorithm on the Lefschetz thimble. Below we give the simplest version.

Suppose we start from a set $\{\sigma^{old}, \eta^{old}, \phi^{old}\}$. First, we propose a thimble $\sigma^{new}$ from the distribution $m_{\sigma^{new}} / \sum m_{\sigma}$. Note that, in view of the arguments presented earlier, this step is typically not needed in simulations of QFT. It is done here to compare with the exact analytical result, which is available.

Next, we choose $n$ independent standard normal deviates $\{\tilde{\eta}_k\}$. The $\eta^{new}$ is then obtained as,

$$
\eta^{new}_k = \frac{1}{\sqrt{\lambda_k}} \tilde{\eta}_k.
$$

Subsequently, $\phi^{new}$ is obtained from $\eta^{new}$ using the procedure outlined above.

The new field configuration is accepted according to the probability,

$$
P_{\text{accept}} = \min\{1, e^{-RS(\phi^{new}) + RS(\phi^{old}) + S_G(\eta^{new}) - S_G(\eta^{old})}\}
$$
Note that each new configuration proposed in this way is completely independent from the previous ones. The acceptance of such proposals may be good as long as the quadratic approximation of the action (that constitutes the basis for the proposal) approximates well the full action. This may not be hopeless, thanks to the basic property of the Lefschetz thimble. In fact, along the thimble, the dominant part of the integral is optimally concentrated close to the stationary point. Indeed, this fact was exploited also in [2]. In any case, the present approach does not rely essentially on the proposal in Eq. (23): it is conceivable to devise a proposal based on a Markov chain, by introducing small random variations to a previous configuration. The key idea of the present algorithm is rather the mapping between the Lefschetz thimble and the Gaussian thimble \( \mathcal{G}_\alpha \).

In either case, given a set of \( N \) uncorrelated field configurations labeled by \( \alpha = 1, \ldots, N \), the expectation values of observables are given by

\[
\langle O \rangle = \frac{\sum_{\alpha} O_\alpha J_\alpha e^{-iS_\alpha}}{\sum_{\alpha} e^{-iS_\alpha}} \tag{25}
\]

where \( S_\alpha \), \( O_\alpha \) and \( J_\alpha \) are, respectively, the values of the action, the observable, the determinant of the Jacobian defined in Eqs. (21 - 22) for the \( \alpha \)-th field configuration. Note that, although the \( iS \) remain constant over each thimble, it can vary from thimble to thimble.

This algorithm is inherently stable. As \( \epsilon \to 0 \), the field configurations will be sampled with the correct measure on the Lefschetz thimble. At finite \( \epsilon \), the distance of sampled field configurations from the Lefschetz thimble is not accumulated over simulation time and there is no chance of divergences. This is because successive \( \phi \)-s are calculated by first generating the \( \eta \)-s.

**One-plaquette model with \( U(1) \) symmetry** — We now discuss the application of the above algorithm for a system with one degree of freedom, viz. the one-plaquette model with \( U(1) \) symmetry. The action is given in terms of the gauge link \( U = e^{i\phi} \) as

\[
S = -i \frac{\beta}{2} (U + U^{-1}) = -i\beta \cos \phi. \tag{26}
\]

where \( \phi \) in this case is a one component field. For real \( \beta \) the action is complex, similar to real time gauge theories.

For this simple model, all the integrals can be evaluated analytically, which offers the chance to compare every detail of our numerical results to exact results. In particular the plaquette average of the phase \( e^{i\phi} \) is given by,

\[
\langle e^{i\phi} \rangle = i \frac{J_1(\beta)}{J_0(\beta)} \tag{27}
\]

with \( J_\alpha(\beta) \) being Bessel functions of the first kind. This analytic result offers the chance of a clear test of our algorithm.

Obtaining this result using stochastic methods is quite non-trivial. For example the complex Langevin method without ad-hoc optimizations gives the wrong result for this model [20].

In order to apply our method, we treat the field \( \phi \) as complex. The action \( S \) has two critical points at \( \phi = 0 \) and \( \pi \). By explicitly constructing the Hessian, it is easy to show that both the critical points are non-degenerate. In this simple model we can also compute the intersection numbers \( (m_\alpha) \), which turn out to be equal to 1 for both thimbles. The field configurations on the two thimbles are related by the discrete symmetry transformation \( \phi \to \pi - \phi \), and expectation values of observables can be written in terms of integrals over one thimble only. However, in order to illustrate the above algorithm, we perform stochastic integration using the full Eq. (6).

For this model, one can explicitly derive the expression for the thimbles attached to the two saddle points. This can be obtained by requiring that the imaginary part of the action be constant along the flow, which gives

\[
\cos \Re \phi \cosh \Im \phi = \pm 1 \tag{28}
\]

as the equations for the Lefschetz thimbles attached to the two saddle points. Such a simple characterization of the thimble is not available for systems with more than one degree of freedom. Of course, our algorithm does not make use of Eq. (28), but in Fig. 1 we show that the fields obtained using the method described above reproduce well the exact thimble defined by Eq. (28).

We see systematic improvement in our results on increasing \( N_\tau = e\tau^{-1} \); with increasing \( N_\tau \) the sampled field configurations uniformly converge on the true thimble. In contrast, the flat Gaussian thimble \( (N_\tau = 1) \) approximates the thimble quite well near the saddle point, but it noticeably different further away from the saddle point.

In Fig. 2 we show the results for the expectation value of the observable \( e^{i\phi} \) for different \( \beta \). Again, the results from our method systematically approach the exact analytical result with increasing \( N_\tau \). For \( N_\tau = 200 \), the results from our method are identical (within statisti-
residual phase is an important one. We expect it to produce a milder sign problem (if at all), than the original sign problem. Nevertheless, it should be included in any quantitative estimate. In our formulation the full (complex) measure of integration is given by $\det \begin{bmatrix} J_\phi \end{bmatrix} e^{-S}$.

The full integrals on the Lefschetz thimble are always real. This means that $\sin \left( \arg \{ \det \begin{bmatrix} J_\phi \end{bmatrix} e^{-S} \} \right)$ does not contribute to the integral. The statement that the sign problem in our method is mild (or absent) means that $\cos \left( \arg \{ \det \begin{bmatrix} J_\phi \end{bmatrix} e^{-S} \} \right)$ (residual phase) will vary very little (or not at all), in the region where $|\det \begin{bmatrix} J_\phi \end{bmatrix} e^{-S}|$ (probability measure) is significant.

For the $U(1)$ one-plaquette model, the Jacobian of the transformation on each thimble is a single number and is simply given by,

$$J_\phi = -\frac{i\beta \sin\phi}{\eta}.$$

In Fig. 3 we show the residual phase vs the positive probability measure for this model. We see that the residual phase changes by very little for variations of the probability measure spanning many orders of magnitude. Moreover, the fluctuations of the residual phase grow milder as the true thimble is approached starting from the Gaussian thimble. Most importantly, the residual phase keeps the same sign throughout the full domain of integration, i.e., there is no sign problem for our method for this particular model. This is reassuring, although it is impossible to extrapolate from this simple model any claim about the residual phase on systems with many degrees of freedom.

**Conclusions** — In this paper we have described a new stable algorithm to sample field configurations on the Lefschetz thimble. We applied this method to the one plaquette model with $U(1)$ symmetry. Our results are in perfect agreement with the exact results from analytical integration. Also, the residual phase remains quasi-constant over configurations with large weight, indicating that our method does not suffer from a sign problem for this system. Further optimization of the algorithm in order to apply it to more challenging problems with a large number of degrees of freedom is underway.

**Acknowledgments** — We would like to thank Francesco Di Renzo, Giovanni Eruzzi, Christian Torrero and Christian Schmidt for useful discussions.

---

[1] C. Morningstar, arXiv preprint hep-lat/0702020 (2007).
[2] R. Baer, M. Head-Gordon, and D. Neuhauser, *The Journal of Chemical Physics* **109**, 6219 (1998).
[3] G. Aarts, PoS **LATTICE2012**, 017 (2012), arXiv:1302.3028 [hep-lat].
[4] S. Chandrasekharan, *Phys. Rev. D* **82**, 025007 (2010).
[5] E. Vicari and H. Panagopoulos, *Phys.Rept.* **470**, 93 (2009) arXiv:0803.1593 [hep-th].
[6] J. Berges, S. Borsanyi, D. Sexty, and I.-O. Stamatescu, *Phys.Rev.* **D75**, 045007 (2007), arXiv:hep-lat/0609058 [hep-lat].
[7] P. L. Silvestrelli, S. Baroni, and R. Car, *Phys. Rev. Lett.* **71**, 1148 (1993).
[8] D. M. Charutz and D. Neuhauser, *The Journal of Chemical Physics* **102**, 4495 (1995).
[9] E. Loh, J. Gubernatis, R. Scalettar, S. White, D. Scalapino, et al., *Phys.Rev.* **B41**, 9301 (1990).
[10] S. Koonin, D. Dean, and K. Lan-
[11] T. Vilgis, *Physics Reports* **336**, 167 (2000).

[12] E. Witten, in *Chern-Simons Gauge theory: 20 years after*, edited by J. E. Andersen and H. A. Boden and A. Hahn and B. Himpel (2010) pp. 347–446, arXiv:1001.2933 [hep-th].

[13] E. Witten, arXiv:1009.6032 [hep-th] (2010).

[14] M. Cristoforetti, F. Di Renzo, and L. Scorzato (Aurora-Science Collaboration), *Phys. Rev. D* **86**, 074506 (2012).

[15] M. Cristoforetti, F. Di Renzo, A. Mukherjee, and L. Scorzato, arXiv:1303.7204 [hep-lat] (2013).

[16] F. Pham, in *Proc. Symp. Pure Math*, Vol. 40 (1983) pp. 319–333.

[17] G. Guralnik and Z. Guralnik, *Annals of Physics* **325**, 2486 (2010).

[18] S.-J. Dong and K.-F. Liu, *Physics Letters B* **328**, 130 (1994).

[19] M. Cristoforetti, L. Scorzato, and F. Di Renzo, in *Journal of Physics: Conference Series*, Vol. 432 (IOP Publishing, 2013) p. 012025.

[20] J. Berges and D. Sexty, *Nucl. Phys. B* **799**, 306 (2008).