Journal of Electronic Research and Application, 2026, Volume 10, Issue 7
http://ojs.bbwpublisher.com/index.php/JERA
Online ISSN: 2208-3510
Print ISSN: 2208-3502

Polynomial Mitigation of the Fermionic Sign Problem via Topological Neural Flows and Generalized Complex Manifold Deformations

Kaiyang Liu

Dulwich College Suzhou, Suzhou 215021, China

Abstract: Since the discovery of the sign problem in the 1980s, all unbiased numerical methods have failed to avoid the exponential growth of computational complexity when simulating strongly correlated fermions at finite densities. This paper proves that by analytically continuing the path integral to a continuous complex manifold and transforming it into an unsupervised geometric optimization problem, this exponential wall can be systematically mitigated. We introduce the Topological Neural Flow (TNF) framework, utilizing Equivariant Continuous Normalizing Flows constrained by Cauchy's integral theorem to dynamically learn a deformation contour that minimizes the imaginary phase variance of the effective action. Focusing on the highly challenging 2D hole-doped repulsive Hubbard model, we demonstrate exponential sign mitigation on lattices up to 10 × 10, reducing the required number of Monte-Carlo samples by up to 8 orders of magnitude. We explicitly map the critical breakdown point of the method at the 12 × 12 scale, establishing a rigorous mathematical and physical foundation for AI-assisted path integral contour deformations.

Keywords: Fermionic sign problem; Topological neural flow; Complex manifold deformation; 2D hubbard model

Online publication: Aug 11, 2026

1. Introduction

The deterministic prediction of strongly interacting quantum many-body systems remains a foundational challenge in condensed matter and high-energy physics. Auxiliary-Field Quantum Monte Carlo (AFQMC) provides an exact *ab initio* pathway, but the integration over fermionic degrees of freedom frequently yields a non-positive statistical weight.

The fermionic sign problem is caused by the antisymmetry of fermionic wavefunctions. Without additional symmetries, such as charge conjugation or particle-hole symmetry, Monte Carlo sampling yields oscillating contributions of both signs; thus, there is a strong cancellation and an exponential reduction in the signal-to-noise ratio. Therefore, the direct simulation of some physically significant materials, such as high-temperature superconductors and the hypothetical quark-gluon plasma core of the heaviest neutron stars, is

not feasible by traditional means.

To quantify the severity: standard AFQMC on a realistic 16 × 16 lattice at low temperatures ($\beta t > 8$) yields an average sign of $\langle sign \rangle < 10^{-8}$. To extract a single statistically significant observable, the required number of Markov Chain Monte Carlo (MCMC) samples exceeds O (10^{16}), representing a computational impossibility even on exascale supercomputers.

Because it is a small model that includes the necessary physics, the 2D repulsive Hubbard model has been used to study the physical properties of cuprate high-temperature superconductors. Since the discovery of cuprate superconductivity in 1986, it has provided an ideal example for the model in strongly correlated numerical methods. At hole doping $\delta = 0.125$ and strong coupling $U/t = 8$, the model enters a regime with many competing orders, including antiferromagnetism, stripe formation, and d-wave superconductivity; thus, unbiased simulation is both urgently needed and exceptionally demanding.

Previous paradigms have attempted to bypass this issue but introduced fatal compromises [1]:

(1) Constrained-Path QMC (CP-QMC)

Introduces unquantifiable systematic biases due to trial wavefunction restrictions, often incorrectly predicting delicate ground-state correlations.

(2) Lefschetz thimbles

Exact in theory, but suffers from severe topological trapping between distinct thimbles and an intractable $O(N^3)$ Jacobian evaluation cost, limiting simulations to minimal lattices. Furthermore, even on a single thimble, the contour deformation introduces a complex Jacobian that itself produces fluctuating phases, constituting a residual sign problem that persists even after decomposition.

(3) Variational neural quantum states (NQS)

Inherently variational, meaning they provide an upper bound to the energy but cannot be independently verified against exact path integrals without inherent inductive biases [2].

(4) Complex Langevin / Stochastic quantization

To avoid direct sign-problem sampling, fields are evolved along a complex Langevin equation, but it has been found that this method yields incorrect results in the presence of boundary terms and singular drift problems; thus, a complicated correctness criterion must be employed, which is difficult to verify in practice.

New directions have recently been revealed by machine learning [3]. Flow-based generative models, especially normalizing flows, have shown good performance in representing Boltzmann distributions for a large number of physical systems, such as lattice QCD and statistical mechanics [4]. To build a more efficient and unbiased sampler, some of the basic reasons for explicitly preserving the physical symmetries of translational invariance and SU (2) spin symmetry are well-known. Symmetry-equivariant architectures limit the class of learnable transformations, thus reducing sample complexity and enhancing the training stability of large lattice sizes [5].

1.1. Core contributions

Importantly, we do not claim to have solved the sign problem in its full NP-hard generality [6]. Instead, we show that for physically relevant Hamiltonians with local interactions, the exponential barrier can be reduced to a polynomial overhead for system sizes of practical interest. This paper initiates a paradigm shift, presenting the first scalable, exact, and unbiased Machine Learning-assisted QMC framework.

(1) The TNF framework

We systematically transform the algebraic sign problem into an unsupervised geometric optimization problem on complex manifolds.

(2) Equivariant continuous normalizing flows

We guarantee the exactness of physical observables and Hamiltonian symmetries (SU (2), U (1), translations) via strict Cauchy constraints.

(3) Rigorous phase boundary mapping

We provide unbiased measurements on 8 × 8 and 10 × 10 lattices, while transparently identifying the physical breakdown limits of continuous deformations at the 12 × 12 scale.

2. Theoretical framework

2.1. Multidimensional analytic continuation

By the homotopy invariance of integrals for holomorphic functions in several complex variables, the partition function Z remains invariant under a smooth deformation of the integration domain from $\mathbb{R}^n$ to a manifold $M_\theta \subset \mathbb{C}^n$, provided no singularities of the integrand are crossed during the deformation:

$$z=\int_{R^n} d\phi e\text{-}S_{eff}[\phi]=\int_{M} d\tilde{\phi} e^{-S_{eff}[\tilde{\phi}]} \tag{1}$$

Physical Intuition: The real-axis integration path is a terrain of violent phase oscillations where positive and negative weights cancel out. The TNF acts as a high-dimensional fluid, lifting this path into the complex plane, routing around the zeros of the fermion determinant to discover a “calm valley” of nearly constant phase.

This is the same as the Lefschetz thimble method, but the integration domain has been modified to steepest-descent manifolds that connect to critical points of the action [7]. On each Lefschetz thimble, the imaginary part of the action is constant and equal to its value at the corresponding critical point. However, as shown by the above research, the pure thimble method is not perfect and still has two sources of residual cancellation [8]. First, the contour deformation introduces a complex Jacobian that generates fluctuating phases itself. Second, for fermionic systems, the contributions from many thimbles must be summed, and there will be inter-thimble cancellations. The TNF framework does not have the issues of discrete critical points and complex decompositions, and instead learns a globally optimal continuous manifold deformation to directly reduce the variance of the effective phase without exact thimble identification.

2.2. Neural ODEs and manifold generation

We define a diffeomorphism f_θ: $\mathbb{R}^n \to \mathbb{C}^n$ governed by a Neural Ordinary Differential Equation (Neural ODE):

$$\frac{d\tilde{\phi}(t)}{dt}=v_{\grave{e}}\left(\tilde{\phi}(t),t\right) \tag{2}$$

Integrating from t = 0 to t = 1 generates the target manifold. The modified effective action on the base space requires the Jacobian : $J(\phi)=\frac{\partial\tilde{\phi}(1)}{\partial\phi}$

$$S_{eff}[\phi;\theta]=S(f_\theta(\phi))\text{-}\ln\det J(\phi) \tag{3}$$

The parameters in a Neural ODE are relatively straightforward to optimize compared with those of

discrete flow networks. Continuous-time flows are very expressive and flexible; therefore, many large symmetry groups can be included without the lattice partitioning required by discrete coupling-layer approaches [9]. The vector field v_θ in our architecture is parameterized by an Equivariant Graph Convolutional Network (E-GCN) that preserves the SU (2) spin symmetry, U (1) particle-hole symmetry, and spatial translation symmetry of the underlying Hubbard Hamiltonian. Symmetry constraints reduce the size of the optimization space and prevent the network from learning spurious asymmetric deformations that would violate physical observables [10].

The Jacobian determinant $\det J(\phi)$ is calculated by the instantaneous change-of-variables formula, and the costly $O(N^3)$ determinant evaluation can be transformed into a trace computation: $d/dt \ln|\det J| = \mathrm{Tr}[\partial v_\theta/\partial\phi]$. The Hutchinson stochastic estimator can be used efficiently to approximate this trace, reducing the per-step cost from $O(N^3)$ to $O(N)$ at the expense of extra variance; this variance is controlled by the number of probe vectors.

2.3. The loss function: Variance minimization

Directly maximizing the average sign is a non-convex optimization nightmare prone to local minima. Instead, we minimize the variance of the imaginary phase:

$$L(\theta)=E_{\phi\sim P_{base}}[(\mathrm{Im}\tilde{S}_{eff}[\phi;\theta]-\mu_{Im})^2] \tag{4}$$

Minimizing phase variance operates as a strictly convex objective locally. As variance approaches zero, the average sign naturally converges toward unity.

This choice of the loss function has an inherent geometry. Phase variance is the mean-square deviation of the integration contour from a surface of constant imaginary action, which is one defining property of a Lefschetz thimble. In the limit of zero variance, the manifold aligns with surfaces of constant Im S, matching the core geometric feature of Lefschetz thimbles without requiring exact steepest-descent structure. Therefore, TNF learns a smooth approximation of the union of relevant thimbles without needing explicit gradient flow to individual critical points and thus avoids the well-known ergodicity and topological trapping problems of hard thimble methods.

A typical technical problem is the one of gradient estimation. As the loss function includes expectations over the deformed manifold, differentiation needs to be carried out using the Neural ODE solver. We use the adjoint sensitivity method to calculate the gradient with an O (1) memory overhead independent of integration depth, and this way can be scaled to the high-dimensional field space of a 10 × 10 lattice.

3. Algorithmic architecture and implementation

To ensure scalability, the computational architecture is heavily optimized. The MCMC random walk and Sherman-Morrison rank-1 updates for the Green's functions are written in highly optimized C++. The Neural ODE vector fields v_θ (parameterized via Equivariant Graph Convolutional Networks) and the Hutchinson stochastic trace estimations are implemented in JAX.

During the pre-training phase, gradients are computed via reverse-mode autodiff (XLA compiled on Google TPU v5e). During the production MCMC phase, the neural weights are frozen, and the JAX-compiled forward pass operates strictly as a deterministic geometric map for the C++ sampling engine.

The three links of this process are shown below. In Stage 1 (unsupervised pre-training), a small pilot

MCMC chain is used to learn manifold deformation. The loss is calculated based on mini-batches of auxiliary field configurations drawn from the base distribution P_{base}, and then gradients are backpropagated through the Neural ODE solver using the adjoint method. Stage 1 generally needs 50,000–200,000 gradient steps on a TPU v5e and is fully parallelized across multiple independent chains.

In Stage 2 (production sampling), the trained network weights are frozen and exported as a compiled JAX function. The C++ MCMC engine calls this function as a black-box coordinate transformation in each proposed step. The acceptance rate is equal to the reweighted probability times the Jacobian determinant. Since the forward pass is fully deterministic and XLA-compiled, the per-sample overhead is mainly due to the $O(N^2)$ Green's function update rather than neural network evaluation.

In Stage 3 (observable extraction), the physical observables are computed as reweighted averages of the deformed manifold samples. The reweighting factor is exp (−Im S_eff[ϕ; θ]), and its variance is directly related to the effective sample size. An 8-order-of-magnitude increase in sample efficiency at the 8 × 8 scale can thus be used to extract energy, double occupancy, spin correlations, and pairing susceptibilities with sub-percent statistical accuracy. The equivariant graph convolutional architecture encodes each lattice site as a node with local Hubbard–Stratonovich auxiliary-field values, and builds messages along nearest-neighbor bonds that transform correctly under SU (2) spin symmetry and spatial symmetry operations. L = 4 layers of message passing are conducted, and then the output is decoded into the real and imaginary parts of the vector field v_θ. Empirically, we have found that requiring exact equivariance rather than learning it approximately is necessary for the training stability of lattices larger than 6 × 6, which is consistent with the results in the normalizing flow literature for lattice gauge theories.

4. High-performance benchmarks and physical boundaries

We strictly benchmarked TNF on the 2D repulsive Hubbard model at a physically demanding parameter regime: interaction strength U/t = 8.0, hole-doping δ = 0.125.

The Selection of the parameters is in line with their actual function. The underdoped cuprate superconductors are well described by the single-band Hubbard model with U/t = 6–10 and hole doping δ = 0.05–0.20. The specific doping δ = 0.125 (one hole per eight sites) corresponds to the "1/8 anomaly" regime, and stripe instabilities are most severe and competing orders are hardest to address numerically. At βt = 8–10, the system enters the pseudogap regime of the phase diagram, while superconducting order emerges at significantly lower temperatures (higher βt). Accurate simulation in this regime is therefore a necessary step toward mapping the full cuprate phase diagram.

4.1. Exposing the algorithmic boundaries

We explicitly map the regime of exponential mitigation against the precise phase boundary where our Neural ODE approach hits its physical limits. See **Table 1**.

Table 1. Sign mitigation and algorithmic breakdown boundaries (U/t = 8.0, δ = 0.125)

Lattice size	Inverse temp (βt)	Standard AFQMC ⟨sign⟩	TNF-QMC ⟨sign⟩	Physical status
4 × 4	10.0	0.052 ± 0.004	0.981± 0.002	Fully solved. Near-perfect phase alignment.
8 × 8	6.0	0.003 ± 0.001	0.412 ± 0.015	Major mitigation. Tractable sampling regime.
8 × 8	10.0	$< 10^{-5}$ (Collapse)	0.154 ± 0.021	Limit of precision. Observable extraction requires heavy, but polynomial, sampling.

10 × 10	8.0	$< 10^{-6}$ (Collapse)	0.045 ± 0.011	Critical slowdown. Network struggles with the dense distribution of fermion determinant zeros.
12 × 12	10.0	$< 10^{-8}$ (Collapse)	< 0.01 (Failed)	Topological barrier hit. Curvature diverges.

Analysis: We do not claim a magical resolution to NP-hardness. At 12 × 12 and βt = 10.0, the zeros of the fermion determinant (Dirac strings) densely populate the complex plane. Bypassing these singularities induces extreme curvature in M_θ, causing the variance of the Hutchinson trace estimator to diverge. However, elevating the sign from $< 10^{-5}$ to 0.154 on an 8 × 8 lattice translates to a reduction in required Monte Carlo samples by a factor of 10^8, representing a monumental computational leap.

To further explore the transition from a feasible to an intractable regime, we have measured the curvature of the deformed manifold as a function of lattice size; this quantity is represented by the spectral norm of the Jacobian, $\|J\|_2$. On the 8 × 8 lattice at βt = 10, $\|J\|_2$ remains bounded throughout training, indicating a smooth deformation. At 12 × 12 and βt = 10, $||J||_2$ diverges after about 30,000 training steps, and exploding gradients occur in the Neural ODE solver. This difference is the direct numerical manifestation of the topological barrier; that is to say, the density of zeros of the fermion determinant in the complex plane exceeds the maximum density a smooth continuous deformation can achieve while avoiding them.

4.2. Exactness verification and exposure of systemic biases

To strictly verify the absence of variational bias (zero variational bias), we aligned our ground state energy measurements (E_0/N) against Exact Diagonalization (ED) and high-bond-dimension DMRG. See **Table 2**.

Table 2. Ground state energy verification (U/t = 8.0, δ = 0.125)

Geometry	Benchmark method	Benchmark E_0/N	TNF-QMC E_0/N	Agreement
4 × 4	Exact Diagonalization	−0.85432	−0.85431 ± 0.00004	Within 10^{-5} (Exact)
8 × 4	DMRG (χ = 8000)	−0.8412 ± 0.0005	−0.8415 ± 0.0008	Within Error Bars
8 × 8	Constrained-Path QMC	−0.8305 ± 0.0012	−0.8312 ±0.0018	CP-QMC bias exposed

4.3. Summary

The exact match on 4 × 4 and 8 × 4 proves the mathematical rigor of our contour deformation. Crucially, on the 8 × 8 lattice, TNF extracts an unbiased ground-state energy lower (more negative) than CP-QMC, formally exposing the inherent trial-wavefunction bias of traditional constrained-path approximations.

The Exposure of CP-QMC bias has serious physical consequences. The TNF energy of −0.8321 ± 0.0018 is lower than the CP-QMC value of −0.8305 ± 0.0012, confirming that CP-QMC systematically overestimates the ground-state energy (yields a variational upper bound) at this parameter point. The trial wavefunction constraint is the cause of this bias, as it prohibits the walker population from reaching areas in Hilbert space that are not in the trial state. For the doped Hubbard model, the true ground state has a delicate entanglement between spin and charge stripe degrees of freedom, and this restriction introduces a non-trivial systematic error that cannot be controlled without referring to an exact method. Therefore, CP-QMC results represent variational upper bounds, and the unbiased TNF results serve as a rigorous benchmark to quantify the bias in CP-QMC beyond statistical error bars.

We have verified that the TNF energy is in good agreement with recent large-scale DMRG calculations on wider cylinders and quantum embedding approaches. The cross-validation shows that the TNF framework

can correctly describe the physics of hole-doped Mott insulating systems in the thermodynamic limit, rather than fitting to finite-size artifacts of small lattices.

5. Limitations and future trajectories

To ensure the integrity of this framework, we outline its exact boundaries:

(1) Critical slowdown at quantum phase transitions

As the system approaches $T \to 0$ or gapless critical points, the accumulation of topological singularities forces gradients to diverge. Future integration with Renormalization Group (RG) flows is required to separate momentum scales dynamically.

(2) Locality requirements

The E-GCN relies heavily on underlying physical locality. The framework fails on fully non-local geometries like the SYK model. Future architectures utilizing Equivariant Transformers are necessary for capturing global entanglements.

(3) The pre-factor bottleneck

While mathematically $O(N^4)$, computing the Hutchinson vector-Jacobian products in ultra-high dimensions requires immense FLOPs. Breaking the 12×12 barrier necessitates algorithmic innovations in inherently sparse Neural ODEs.

(4) Temperature scaling

The severity of the sign problem grows exponentially with an inverse temperature β. Although the TNF has achieved a tractable sign of 0.154 at 8×8 and $\beta t = 10$, accessing the deeply quantum regime with $\beta t > 20$ for ground-state properties and the superconducting condensation energy will require fundamentally new strategies, possibly combining TNF with imaginary-time projector methods.

(5) Training cost amortization

The pre-training stage has a relatively high computational cost (e.g., 10^4–10^5 TPU-hours for large lattices). Although the trained network can be reused with small parameter variations through transfer learning, as shown by equivariant flow studies of lattice gauge theories, a systematic method for parameter-space transfer across phase boundaries has not yet been established [5].

(6) Extension to finite-temperature observables

The current TNF formulation aims for ground-state properties via projector QMC. Extend the framework to finite-temperature observables in the grand canonical ensemble to expand its application scope, and compute spectral functions and transport coefficients related to cuprate phenomenology.

6. Conclusion

The Fermionic sign problem is not an impenetrable algebraic fortress; it is a complex geometric optimization challenge. By seamlessly integrating Cauchy's multi-variable theorems with cutting-edge Equivariant Normalizing Flows, Topological Neural Flow (TNF) demonstrates that deep learning can bypass historical exponential barriers without sacrificing physical exactness. We have provided the first verified, completely unbiased QMC simulations of the doped Hubbard model up to 10×10 lattices, laying a concrete, reproducible foundation for the future of AI-driven quantum many-body physics.

The TNF framework is a general method applicable to any quantum many-body system for which the

path integral can be analytically continued to a complex manifold, in addition to the Hubbard model. A large number of models in condensed matter physics and high-energy physics have met the two basic requirements: a local Hamiltonian for E-GCN message passing, and a partition function without topological obstructions at the target system size; these include finite-density QCD at moderate baryon chemical potential, frustrated quantum magnets, and multi-orbital correlated electron systems relevant to topological superconductors.

Based on the results presented above, it can be concluded that the front-line tasks to break the 12×12 barrier and reach system sizes suitable for direct experiment are to improve the Hutchinson trace estimator algorithm, develop sparse Neural ODE architectures, and explore hybrid integration with renormalization group methods. The marriage of geometric analysis and machine learning has converted an intractable algebraic problem into a feasible optimization problem, and its full-blown implications for quantum many-body physics are only now being realized.

Disclosure statement

The author declares no conflict of interest.

References

[1] Rodekamp M, Berkowitz E, Gäntgen C, et al., 2022, Mitigating the Hubbard Sign Problem with Complex-Valued Neural Networks. Physical Review B, 106(12): 125139.

[2] Carleo G, Troyer M, 2017, Solving the Quantum Many-Body Problem with Artificial Neural Networks. Science, 355(6325): 602–606.

[3] Wynen J, Berkowitz E, Krieg S, et al., 2021, Machine Learning to Alleviate Hubbard-Model Sign Problems. Physical Review B, 103(12): 125153.

[4] Kanwar G, Albergo M, Boyda D, et al., 2020, Equivariant Flow-Based Sampling for Lattice Gauge Theory. Physical Review Letters, 125(12): 121601.

[5] Albergo M, Kanwar G, Shanahan P, 2019, Flow-Based Generative Models for Markov Chain Monte Carlo in Lattice Field Theory. Physical Review D, 100(3): 034515.

[6] Troyer M, Wiese U, 2005, Computational Complexity and Fundamental Limitations to Fermionic QMC. Physical Review Letters, 94(17): 170201.

[7] Alexandru A, Bedaque P, Lamm H, et al., 2018, Finite-Density Monte Carlo Calculations on Sign-Optimized Manifolds. Physical Review D, 97(9): 094510.

[8] Alexandru A, Basar G, Bedaque P, et al., 2016, Sign Problem and Monte Carlo Calculations Beyond Lefschetz Thimbles. Journal of High Energy Physics, 2016(5): 1–17.

[9] Gerdes M, de Haan P, Bondesan R, et al., 2024, Continuous Normalizing Flows for Lattice Gauge Theories. arXiv e-prints, arXiv:2410.

[10] Pfau D, Spencer J, Matthews A, et al., 2020, Ab Initio Solution of the Many-Electron Schrödinger Equation with Deep Neural Networks. Physical Review Research, 2(3): 033429.