

Lattice fermion formulation via physics-informed neural networks: Ginsparg-Wilson relation and overlap fermions

Tatsuhiro Misumi^{*}*Department of Physics, Kindai University, Higashi-Osaka, Osaka 577-8502, Japan*

(Received 13 May 2026; accepted 24 June 2026; published 16 July 2026)

We propose a novel, machine-learning-based framework for constructing lattice fermions using physics-informed neural networks (PINNs). Our approach treats the formulation of the Dirac operator as an optimization problem guided by physical requirements, such as symmetries, locality and doubler-decoupling conditions. We first demonstrate that, when trained to satisfy the Ginsparg-Wilson (GW) relation as a soft constraint, a neural network reproduces the overlap fermion operator to high numerical accuracy and learns an effective sign-function mapping without explicitly using a prescribed polynomial or rational approximation. Second, we extend the framework from operator construction to machine-assisted algebraic discovery. Within a generalized polynomial ansatz, the network autonomously drives higher-order terms to zero and recovers the standard Ginsparg-Wilson relation. Remarkably, by changing the initial search bias, the same framework also finds a distinct solution corresponding to a Fujikawa-type generalized GW relation.

DOI: [10.1103/756m-stz5](https://doi.org/10.1103/756m-stz5)

I. INTRODUCTION

Quantum chromodynamics (QCD) formulated on a discretized spacetime lattice represents the most robust nonperturbative framework for understanding the strong interaction. However, formulating fermions on a discrete lattice is a long-standing fundamental challenge, primarily due to the fermion doubling problem and its inextricable link to the chiral anomaly.

The fundamental obstacle is rigorously formalized by the Nielsen-Ninomiya no-go theorem [1–4], which dictates that no single-flavor lattice Dirac operator on the Euclidean spacetime can simultaneously preserve basic conditions: (1) locality, (2) translational invariance, (3) (γ_5) -hermiticity, and (4) on site continuous chiral symmetry $\{D, \gamma_5\} = 0$.

Historically, navigating these stringent constraints has relied exclusively on human analytical ingenuity, effectively creating a game of choosing which fundamental condition to sacrifice. Wilson fermions [5] successfully decouple the unphysical doublers by introducing an irrelevant dimension-five operator that lifts their masses to the cutoff scale, but at the heavy cost of explicitly breaking chiral symmetry. Staggered fermions [6,7]

distribute spinor components across lattice sites to maintain a remnant symmetry, sacrificing the strict single-flavor interpretation.

A profound paradigm shift occurred with the discovery of the Ginsparg-Wilson (GW) relation [8]. The GW relation evades the Nielsen-Ninomiya theorem by replacing the exact, ultralocal onsite chiral symmetry with a lattice-deformed version, which was later proved to leave the path integral invariant under a well-defined lattice chiral transformation [9]. This elegant algebraic constraint culminated in exact analytical solutions, most notably the overlap fermion [10] and the closely related domain-wall fermion [11–13].

Despite their theoretical beauty, the exact chiral lattice fermions entail immense practical difficulties. The overlap operator requires the evaluation of a matrix sign function. In standard numerical simulations, this sign function must be approximated via Chebyshev polynomials or Zolotarev rational approximations [14,15], which incurs massive computational overhead and introduces truncation errors. Furthermore, the standard overlap fermion often suffers from delicate real-space locality issues that depend heavily on the smoothness of the background gauge fields. To circumvent these bottlenecks, lattice theorists have spent decades proposing improved versions, such as optimized and approximate GW operators [16–18], Brillouin-improved fermions [19–23], and related improved kernel constructions [24,25], which utilize hypercubic isotropic derivatives to drastically optimize the dispersion relation and improve locality.

^{*}Contact author: misumi@phys.kindai.ac.jp

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

Concurrently, machine learning (ML) techniques have revolutionized numerous scientific domains. Particularly, physics-informed neural networks (PINNs) [26] have proven that deep learning models can be constrained by fundamental physical laws—such as partial differential equations or algebraic symmetries—embedded directly into their loss functions. While ML has seen rapid adoption in lattice QCD for configuration generation [27–29], solving inverse problems and the formulation of improved gauge actions [30–33], its capacity as an engine for the fundamental formulation of the Dirac action itself remains an uncharted frontier [34].

In this paper, we propose a fundamental paradigm shift: utilizing PINNs not only to construct lattice fermion operators but to discover their underlying algebraic structures. Specifically, in the free-field setting, we demonstrate that a generic neural network trained with the GW relation as a soft constraint can reproduce the overlap Dirac operator to high numerical accuracy. Furthermore, by implementing a differentiable architecture search (DARTS [35])-inspired optimization over a generalized polynomial ansatz, we demonstrate that the network can autonomously recover the standard Ginsparg-Wilson relation from an agnostic initial condition. This recovery should be understood within the prescribed polynomial search space: starting from unbiased coefficients, the optimization selects the $n = 1$ structure and suppresses higher-order terms under the combined requirements of locality and doubler decoupling. Strikingly, by changing the initial logits, the same framework also finds a distinct solution corresponding to a Fujikawa-type generalized GW relation [36,37].

The idea of optimizing lattice discretizations under physical constraints has a long history in the program of perfect and fixed-point actions [38,39]. Optimized Dirac operators, including overlap-hypercube constructions, further showed that improved locality and dispersion properties can be pursued together with chiral symmetry constraints [40]. Our PINN framework can be interpreted as providing a complementary machine-learning route to such an optimization problem.

This paper is organized as follows. In Sec. II, we outline the general methodology of our machine-learning framework and its potential extensions. In Sec. III, we review the foundational mechanics of the GW relation and its relation to overlap fermions. In Sec. IV, we demonstrate that a neural network can successfully learn the overlap mapping. We also discuss its straightforward application to 4D. Finally, in Sec. V, we present the main result of this study: the autonomous algebraic discovery of the GW relation and its higher-order generalizations. Conclusions and future directions are summarized in Sec. VI. To demonstrate the stability of our framework, we provide an analysis of its robustness against variations in the loss weights in Appendix A, and its sensitivity to the locality decay parameter α in Appendix B.

II. LATTICE FERMIONS FROM SOFT CONSTRAINTS

The core philosophy of our proposal is to replace the analytical derivation of lattice operators with constraint-driven machine learning. A generic, translationally invariant lattice Dirac operator $D(k)$ can be parametrized by a highly expressive function approximator, such as a multi-layer perceptron (MLP) defined in momentum space.

The desired physical properties of the fermion are translated into “soft constraints” within a total loss function during the training phase,

$$\mathcal{L}_{\text{total}} = \sum_i w_i \mathcal{L}_i(D). \quad (1)$$

The constraints, or loss functions, can naturally include:

- (i) *Symmetry constraints* ($\mathcal{L}_{\text{sym}}$): Enforcing continuous symmetries, discrete symmetries, or deformed algebraic structures.
- (ii) *Locality constraints* ($\mathcal{L}_{\text{loc}}$): Penalizing long-range real-space hoppings via Fourier transforms to encourage physically acceptable locality properties.
- (iii) *Spectrum constraints* ($\mathcal{L}_{\text{spec}}$): Explicitly suppressing doublers, or pinning doubler locations in momentum space.

By carefully balancing the relative weights w_i , called hyperparameters, the neural network systematically navigates the landscape of valid lattice actions. The model discovers optimal compromises between competing physical requirements (e.g., extreme locality versus exact symmetry), effectively automating the design of lattice field theories.

This modular framework provides immense flexibility and is strictly not limited to the standard single-flavor Dirac operator. For instance, the methodology can be naturally extended to one-component fermions such as staggered fermions. Furthermore, it opens new robust avenues for exploring minimally doubled fermions, where specific discrete symmetries ($\mathcal{C}, \mathcal{P}, \mathcal{T}$) or full hypercubic symmetries are intentionally broken [41–62]. The framework can also be adapted to the Hamiltonian formalism of lattice fermions, establishing a solid foundation for future studies aiming at quantum simulations and tensor network applications [63–71].

III. REVIEW OF THE GINSPARG-WILSON RELATION AND OVERLAP FERMIONS

To provide a concrete realization of our PINN framework, we target the construction of a GW-compliant fermion. We briefly review the underlying mechanics of lattice chiral symmetry to formalize our optimization objectives.

In continuum Euclidean space, the massless Dirac operator $D = \gamma_\mu \partial_\mu$ perfectly anticommutes with γ_5 . A naive

symmetric finite-difference discretization yields the momentum-space operator

$$D_{\text{naive}}(k) = \frac{i}{a} \sum_{\mu} \gamma_{\mu} \sin(ak_{\mu}). \quad (2)$$

While $\{\gamma_5, D_{\text{naive}}\} = 0$ holds, $D_{\text{naive}}(k)$ possesses zeros at 2^d points in the Brillouin zone, corresponding to 2^d doublers. From now, we consider the dimensionless operators and momenta as $aD \rightarrow D$ and $ak \rightarrow k$.

This is resolved by adding a momentum-dependent mass term, or the Wilson kernel M_W ,

$$D_W(k) = i \sum_{\mu} \gamma_{\mu} \sin k_{\mu} + M_W(k),$$

$$M_W(k) = m_0 + r \sum_{\mu} (1 - \cos k_{\mu}). \quad (3)$$

The physical mode at $k = 0$ retains the mass m_0 , whereas the doublers acquire masses of $O(1/a)$. The unavoidable cost is the explicit breaking of chiral symmetry $\{\gamma_5, D_W\} \neq 0$.

It was proposed [8] that the anticommutation relation should be deformed by a lattice artifact as

$$D\gamma_5 + \gamma_5 D = D\gamma_5 D. \quad (4)$$

Assuming γ_5 -hermiticity $D^{\dagger} = \gamma_5 D \gamma_5$, multiplying Eq. (4) by γ_5 gives $D + D^{\dagger} = D^{\dagger} D$. Thus, the eigenvalues of any operator satisfying the GW relation must lie exactly on a circle in the complex plane, centered at (1,0) with radius 1. An exact analytical solution to this relation is the overlap Dirac operator [10]

$$D_{ov} = 1 + \gamma_5 \text{sgn}(\gamma_5 D_W). \quad (5)$$

A crucial aspect of exact chiral lattice fermions is their spatial locality. Unlike the ultralocal Wilson Dirac operator, operators satisfying the GW relation, such as the overlap fermion, are inherently nonultralocal (i.e., their couplings do not strictly vanish beyond a finite distance). However, it has been rigorously shown that as long as the underlying input kernel is sufficiently smooth and local, the resulting overlap operator exhibits strict exponential decay in real space $|D(x, y)| \leq C e^{-\alpha \|x-y\|}$ [72–74]. Consequently, explicitly incorporating and evaluating the exponential spatial locality of the derived operator—characterized by the decay rate α —is a paramount criterion when exploring the landscape of GW-compliant solutions via machine learning.

IV. DEEP LEARNING OF OVERLAP OPERATOR

We first demonstrate that a neural network can successfully learn the exact overlap mapping to high numerical

accuracy when the GW relation is explicitly provided as a target constraint.

We parametrize the Dirac operator using an MLP defined in momentum space. To ensure the network learns a universal mapping rather than overfitting to a specific configuration, we employ domain randomization by uniformly sampling the bare mass $m_0 \in [-1.75, -0.25]$ at each training epoch. The total loss function is defined as

$$\mathcal{L}_{\text{total}} = w_{\text{GW}} \mathcal{L}_{\text{GW}} + w_{\text{pin}} \mathcal{L}_{\text{pin}} + w_{\text{loc}} \mathcal{L}_{\text{loc}}. \quad (6)$$

Here, $\mathcal{L}_{\text{GW}}$ penalizes deviations from the exact GW relation, defined as

$$\mathcal{L}_{\text{GW}} = \frac{1}{N_k} \sum_k \|\{D_{\theta}(k), \gamma_5\} - D_{\theta}(k) \gamma_5 D_{\theta}(k)\|^2. \quad (7)$$

The detailed roles of $\mathcal{L}_{\text{pin}}$ and $\mathcal{L}_{\text{loc}}$ are elaborated in the following subsection.

A. Neural network implementation details

The Dirac operator is modeled by a MLP with three hidden layers (256 units each) using smooth tanh activation functions to ensure infinite differentiability in momentum space. For the 2D case, the input layer of the network takes a 3-dimensional vector consisting of the basic lattice momentum functions

$$\text{Input: } (\sin k_1, \sin k_2, M_W(k)) \in \mathbb{R}^3. \quad (8)$$

The network, parametrized by weights θ , then outputs three real-valued components $(\mathcal{V}_1, \mathcal{V}_2, \mathcal{M}) \in \mathbb{R}^3$, which construct the general momentum-space Dirac operator as

$$\text{Output: } D_{\theta}(k) = i \sum_{\mu=1}^2 \gamma_{\mu} \mathcal{V}_{\mu}(k) + \mathbb{I} \mathcal{M}(k). \quad (9)$$

For 4D, the inputs and outputs are straightforwardly expanded to 5 dimensions to accommodate the four Euclidean gamma matrices and mass or Wilson terms. In essence, our neural network functions as a highly nonlinear map that takes the conventional Wilson fermion kernel as its input and yields an optimized chiral lattice fermion operator as its output.

Parity symmetry is explicitly enforced by symmetrizing the network's forward pass: $\mathcal{V}_{\mu}(k) = \frac{1}{2}(\mathcal{V}_{\mu}^{fwd}(k) - \mathcal{V}_{\mu}^{bwd}(-k))$ and $\mathcal{M}(k) = \frac{1}{2}(\mathcal{M}^{fwd}(k) + \mathcal{M}^{bwd}(-k))$, ensuring that the spatial vector components are exactly odd functions, while the scalar mass component is strictly an even function of momentum.

In our loss function formulation, $\mathcal{L}_{\text{pin}}$ plays a critical role in enforcing the doubler-decoupling conditions. This spectrum constraint explicitly requires the physical zero mode to remain massless ($D(0) = 0$) while forcefully pushing the

unphysical doubler modes to the cutoff scale ($D(k_{\text{doub}}) = 2$), leading to the loss function

$$\mathcal{L}_{\text{pin}} = \|D_{\theta}(0, 0)\|^2 + \|D_{\theta}(0, \pi) - 2\mathbb{I}\|^2 + \|D_{\theta}(\pi, 0) - 2\mathbb{I}\|^2 + \|D_{\theta}(\pi, \pi) - 2\mathbb{I}\|^2. \quad (10)$$

Without this doubler pinning loss, the network might easily fall into trivial parameter spaces, such as a null operator. Alternatively, this strict penalty can be relaxed into an inequality constraint, such as $\|D_{\theta}(k_{\text{doub}}) - 1.5\mathbb{I}\| > 0$, by employing a ReLU-based loss function.

Finally, the locality penalty $\mathcal{L}_{\text{loc}}$ penalizes long-range real-space hoppings and thereby biases the learned operator toward an exponentially localized profile such as $|D(r)| \sim e^{-\alpha|r|}$ discussed in Sec. III. This is implemented via a fast Fourier transform (FFT) to real space as $D_{\theta}(r) = \text{FFT}[D_{\theta}(k)]$, applying a weighted exponential penalty

$$\mathcal{L}_{\text{loc}} = \sum_r \|D_{\theta}(r) \cdot (e^{\alpha|r|} - 1)\|^2. \quad (11)$$

Here, the distance $|r|$ is defined as the shortest Euclidean distance on the periodic lattice. For a d -dimensional lattice of linear size L , it is explicitly calculated as $|r| = \sqrt{\sum_{\mu} \min(|x_{\mu}|, L - |x_{\mu}|)^2}$, which properly accounts for the periodic boundary conditions.

For the investigation in this section, we utilize the lattice size $L = 32$ with a decay parameter $\alpha = 0.8$ for 2D, and $L = 10$ with $\alpha = 0.8$ for 4D. The network parameters are optimized using the Adam optimizer. The learning rate is initialized to $\eta = 0.002$ and is smoothly annealed during the training process to ensure stable convergence. The relative weights of the total loss function are chosen as

$w_{\text{GW}} = 1.0$, $w_{\text{pin}} = 0.1$, and $w_{\text{loc}} = 0.1$. This specific weighting prevents the harsh spatial locality penalty from overpowering the optimization in the early epochs, allowing the network to first satisfy the exact algebraic and spectrum constraints before fine-tuning the spatial profile.

During the training process, the network smoothly navigates the complex optimization landscape. While the initial random weights produce arbitrary spectra that violate all symmetries, the simultaneous minimization of these three distinct loss components steadily guides the network toward a GW-compliant operator with suppressed long-range hoppings.

To explicitly demonstrate the robustness of our framework against the choice of these specific hyperparameters, we provide a stability analysis against variations in the loss weights in Appendix A, and against variations in the locality decay parameter α in Appendix B. Furthermore, in addition to the hyperparameter stability discussed above, to verify the robustness and reproducibility of our framework against different network initializations, we performed multiple independent training runs with different random initializations. Across all trials (each with 10 000–20 000 epochs), the network consistently converged to nearly identical physical results, demonstrating the high stability of the optimization process.

B. Results: Learned overlap mapping in 2D

Post-training analysis demonstrates a good agreement with the GW formulation. As shown in Fig. 1, the eigenvalue spectrum of the learned operator collapses onto the GW circle. It is notable that, while m_0 is randomized during training, the spectrum is evaluated here at $m_0 = -1.0$.

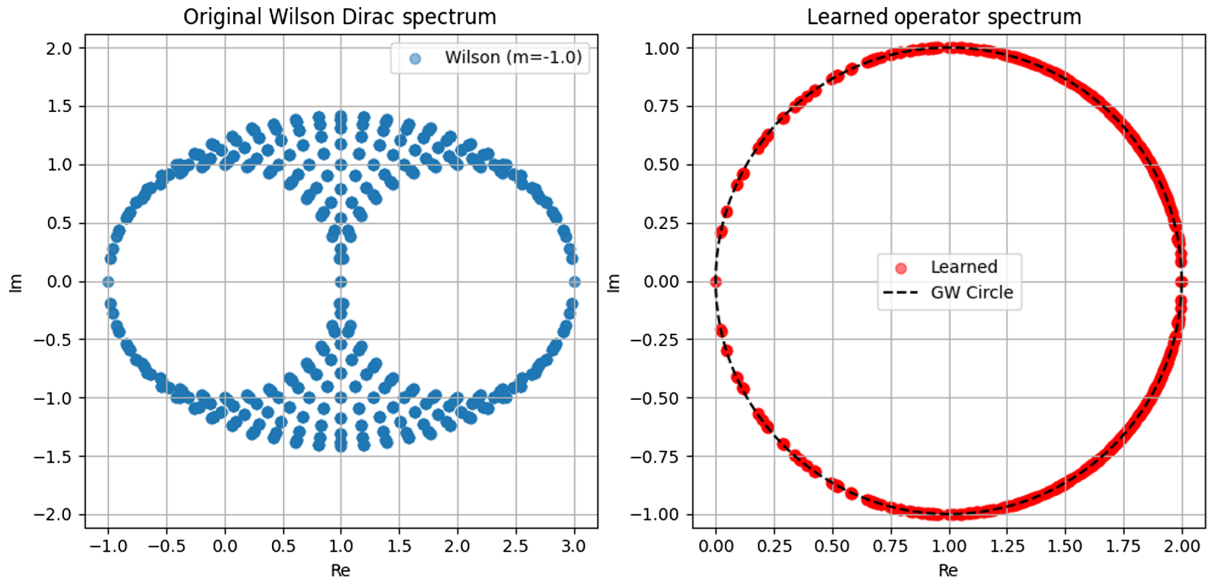


FIG. 1. Comparison of the eigenvalue spectrum in the complex plane. Left: the original 2D Wilson Dirac spectrum (input). Right: the spectrum of the learned 2D operator (output), which collapses onto the Ginsparg-Wilson circle. While m_0 is randomized during training, the spectrum is evaluated here at $m_0 = -1.0$. The parameters are $\alpha = 0.8$, $L = 32$ and $w_{\text{GW}} = 1.0$, $w_{\text{pin}} = 0.1$, $w_{\text{loc}} = 0.1$.

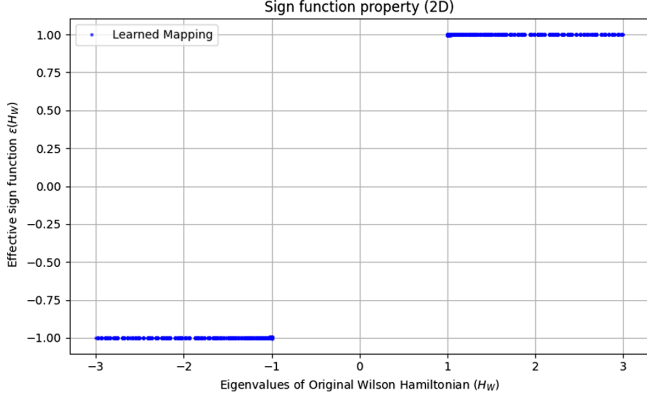


FIG. 2. The effective sign function $\epsilon_{\text{eff}}(H_W)$ extracted from the learned operator, plotted against the eigenvalues of the original Wilson Hamiltonian H_W . The network realizes a sharp step function. The parameters are $\alpha = 0.8$, $L = 32$ and $w_{\text{GW}} = 1.0$, $w_{\text{pin}} = 0.1$, $w_{\text{loc}} = 0.1$.

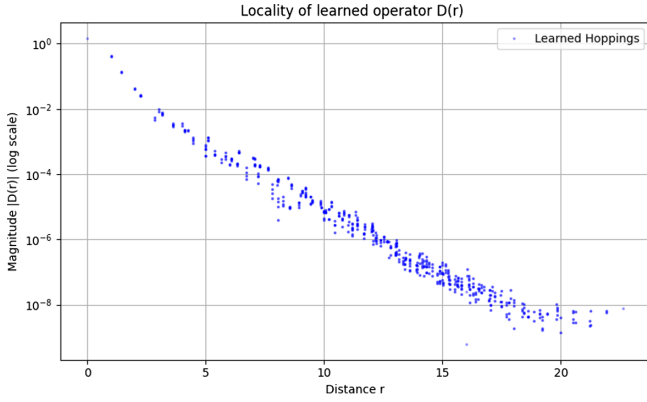


FIG. 3. Magnitude of the real-space hopping parameters $|D(r)|$ of the learned 2D operator as a function of the distance r . The observed exponential-like suppression provides a numerical check of the locality properties of the learned operator. The parameters are $\alpha = 0.8$, $L = 32$ and $w_{\text{GW}} = 1.0$, $w_{\text{pin}} = 0.1$, $w_{\text{loc}} = 0.1$.

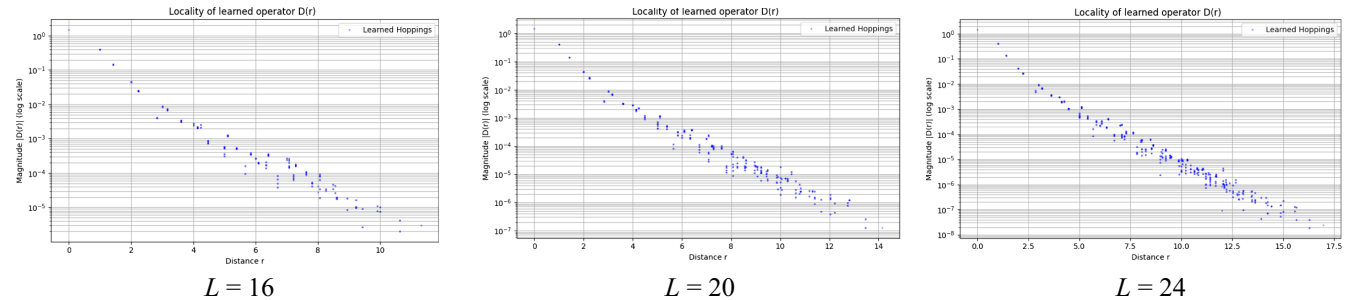


FIG. 4. Magnitude of the real-space hopping parameters $|D(r)|$ of the learned 2D operator as a function of the distance r for smaller lattices $L = 16, 20, 24$ with $\alpha = 0.8$ and $L = 32$. The plateau shifts according to the volume.

The near zero mode deviates from the origin by approximately 3.3870×10^{-4} , showing the high precision of the doubler-pinning loss. To test the accuracy of the learned mapping, we extract the effective sign function ϵ_{eff} from the learned operator and project it against the eigenvalues of the original Wilson Hamiltonian $H_W = \gamma_5 D_W$.

As illustrated in Fig. 2, the extracted effective sign function forms a sharp steplike profile and is numerically saturated near ± 1 on the sampled eigenvalues of H_W . Quantitatively, the maximum absolute deviation of the learned points from the exact theoretical sign function is $\sim 5.2439 \times 10^{-3}$, with a root-mean-square error (RMSE) $\sim 1.0845 \times 10^{-3}$. The result shows that the nonlinear neural-network ansatz can represent the overlap sign-function mapping with high numerical accuracy, without explicitly specifying a polynomial or rational approximation scheme.

Real-space checks in Fig. 3 show that the hoppings of the learned operator are strongly suppressed at large distances, consistently with an exponentially localized profile. An exponential fit $Ce^{-\mu|r|}$ to the decay behavior in the intermediate range yields a fitted locality exponent of $\mu \sim 1.02$. Furthermore, at very large distances, the locality plot exhibits a plateau around $O(10^{-8})$. To clarify whether this feature is a finite-volume artifact or an immanent artifact of the numerical procedure, we performed identical training on smaller lattice volumes $L = 16, 20, 24$ in Fig. 4. The comparison reveals that for these smaller volumes, the plateau does not appear, as the maximum distance is reached before the magnitude hits the $O(10^{-8})$ noise floor. This clearly indicates that the plateau primarily originates from the numerical precision limit (single-precision floating-point format) of the current network training, rather than physical finite volume effects.

We also plot the kinetic term magnitude $|\mathcal{V}_x(k)|^2 + |\mathcal{V}_y(k)|^2$ in Fig. 5 and the momentum-dependent mass term $\mathcal{M}(k)$ in Fig. 6 for Wilson-overlap, Brillouin-overlap [19,20] and our learned operator. By examining the dispersion near the Brillouin zone boundaries, we observe

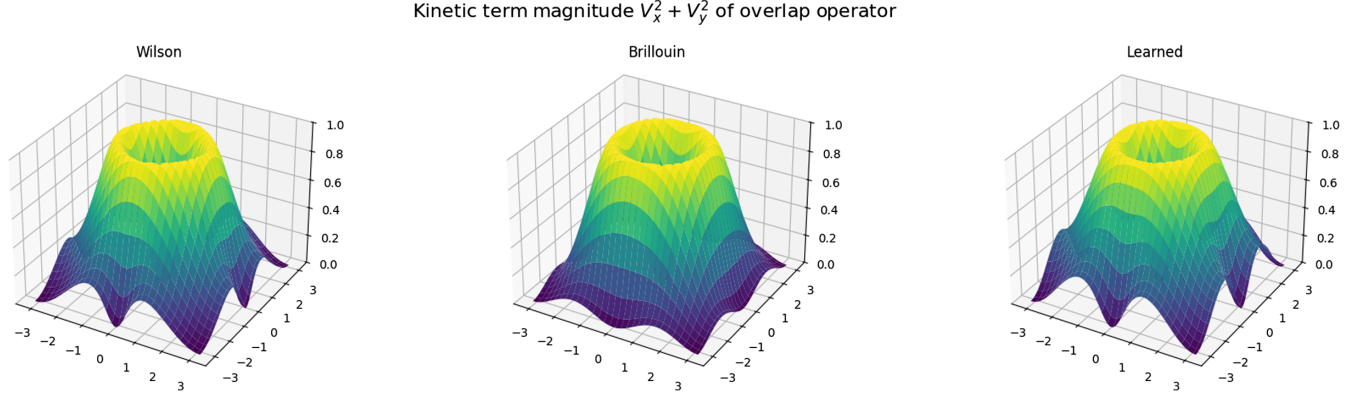


FIG. 5. Plots of the kinetic term magnitude $|\mathcal{V}_x|^2 + |\mathcal{V}_y|^2$. Left: standard Wilson-overlap. Center: Brillouin-overlap with characteristic flat doubler basins. Right: the learned operator with $\alpha = 0.8$, $L = 32$ and $w_{\text{GW}} = 1.0$, $w_{\text{pin}} = 0.1$, $w_{\text{loc}} = 0.1$.

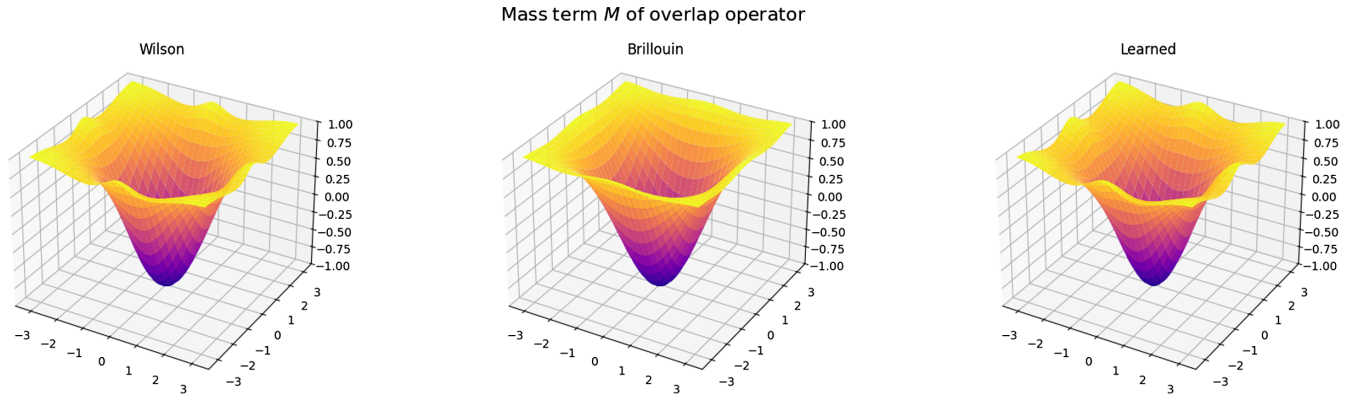


FIG. 6. Plots of the mass term $\mathcal{M}$. Left: standard Wilson-overlap. Center: Brillouin-overlap with characteristic flat doubler basins. Right: the learned operator with $\alpha = 0.8$, $L = 32$ and $w_{\text{GW}} = 1.0$, $w_{\text{pin}} = 0.1$, $w_{\text{loc}} = 0.1$.

that the operator learned by the neural network closely resembles the standard Wilson-overlap operator rather than the Brillouin-overlap, under the present choice of w_{loc} and α . Notably, the neural network framework has the potential to discover improved formulations, such as the Brillouin-overlap operator, through appropriate tuning of the penalty parameters w_{loc} and the decay parameter α . We investigate this issue in Appendix A.

C. Extension to 4D

We successfully extend our methodology to four dimensions to confirm its broad applicability. The network architecture is expanded to output five components, and the exponential locality penalty is evaluated via a 4D FFT over a hypercubic grid. The optimization successfully converges, and the eigenvalue spectrum of the resulting 4D operator lies close to the GW circle as shown in Fig. 7, where we set $w_{\text{GW}} = 1.0$, $w_{\text{pin}} = 0.1$, $w_{\text{loc}} = 0.1$, $\alpha = 0.8$ and $L = 10$. The near zero mode deviates from the exact origin by only $\sim 4.8639 \times 10^{-3}$, demonstrating that the doubler-pinning condition is accurately satisfied in 4D.

The effective sign function similarly realizes a steplike function as shown in Fig. 8. The maximum absolute deviation from the exact 4D sign function is $\sim 3.1569 \times 10^{-2}$ with $\text{RMSE} \sim 1.4993 \times 10^{-3}$. This demonstrates that the machine-learning framework can be similarly applied to higher dimensions. However, in this 4D study, we only have the small lattice as $L = 10$ and need larger lattices to investigate the locality and the property of bare kernels as done for 2D. Future work will be devoted to this investigation.

V. AUTONOMOUS DISCOVERY OF GINSPARG-WILSON RELATION

The previous section demonstrated that a neural network can solve the problem of finding a local lattice operator when explicitly given the GW relation $(D + D^\dagger = D^\dagger D)$ as a target constraint. However, a more ambitious question remains: can the network recover the GW algebraic structure within a prescribed polynomial search space, guided by locality and doubler-decoupling constraints?

To answer this, we extend our framework from inverse construction to machine-assisted algebraic-structure search.

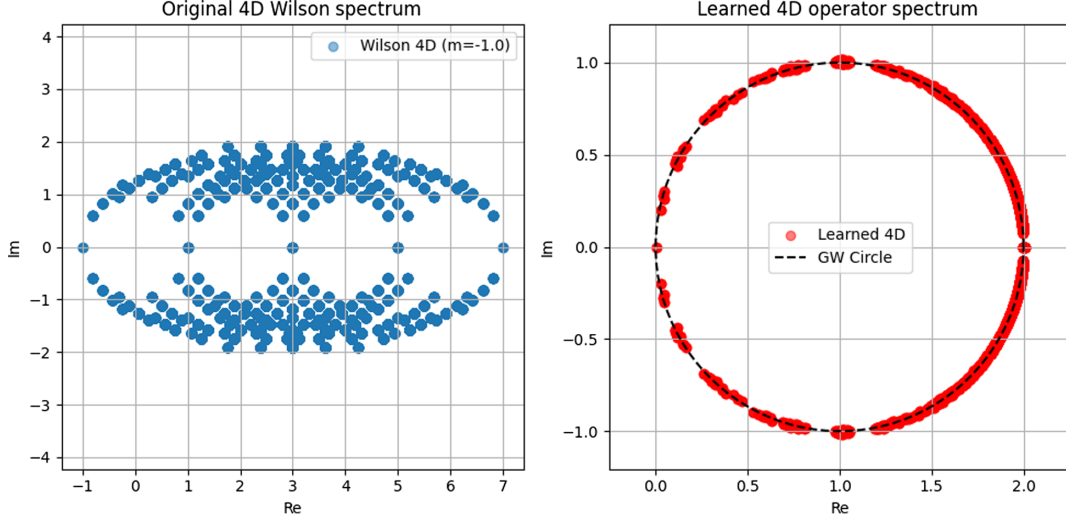


FIG. 7. The 4D spectrum of the learned operator. Left: the original 4D Wilson Dirac spectrum (input). Right: the spectrum of the learned 4D operator (output), which lies close to the Ginsparg-Wilson circle. While m_0 is randomized during training, the spectrum is evaluated here at $m_0 = -1.0$. The parameters are $\alpha = 0.8$, $L = 10$ and $w_{\text{GW}} = 1.0$, $w_{\text{pin}} = 0.1$, $w_{\text{loc}} = 0.1$.

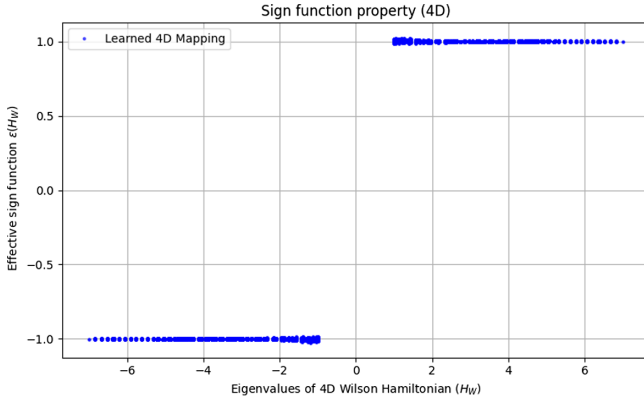


FIG. 8. The effective sign function $\epsilon(H_W)$ extracted from the 4D learned operator, where the network realizes a step function. The parameters are $\alpha = 0.8$, $L = 10$ and $w_{\text{GW}} = 1.0$, $w_{\text{pin}} = 0.1$, $w_{\text{loc}} = 0.1$.

We abandon the explicit GW constraint $\mathcal{L}_{\text{GW}}$ and instead propose a generalized polynomial ansatz for the algebraic structure,

$$D + D^\dagger = \sum_{n=1}^N c_n (D^\dagger D)^n, \quad (12)$$

or equivalently,

$$\mathcal{L}_{\text{pol}} = \frac{1}{N_k} \sum_k \left\| D_\theta(k) + D_\theta^\dagger(k) - \sum_{n=1}^N c_n (D_\theta^\dagger(k) D_\theta(k))^n \right\|^2. \quad (13)$$

The coefficients c_n define the specific algebraic structure of the fermion. In our numerical demonstration of 2D Dirac

operators, we truncate this expansion at $N = 5$. We note that, in this investigation, m_0 is again randomized during training, while the spectrum is evaluated at $m_0 = -1.0$.

The standard GW relation corresponds precisely to the subset $\{c_1 = 1.0, c_{n>1} = 0.0\}$. Our objective is to determine whether the network can independently derive this exact subset of coefficients from a completely agnostic starting state.

A. Differentiable architecture search and global scale

We implement a differentiable architecture search (DARTS) inspired mechanism [35]. We decouple the selection of the algebraic structure from its overall coefficient. The network is parametrized by internal logits α_n , which dictate the relative selection probability of each polynomial term via a Softmax function with temperature annealing. Concurrently, a completely independent parameter, s_{global} , governs the overall coefficient of the relation.

The coefficients are dynamically determined during training as

$$c_n = s_{\text{global}} \times \frac{\exp(\alpha_n/\tau)}{\sum_j \exp(\alpha_j/\tau)} \times 4^{-(n-1)}. \quad (14)$$

The factor $4^{-(n-1)}$ represents a standard basis normalization preventing gradient explosion near the Brillouin zone boundaries, not a structural constraint.

Crucially, the algebraic coefficients are initialized without bias toward the standard GW relation. For instance, all structural logits are set to zero ($\alpha_n = 0$, ensuring a completely flat probability distribution), and the global scale is initialized to $s_{\text{global}} = 0.1$. The network is then completely unaware that the physical target scale is 1.0, nor does it know which combination of c_n is optimal. The total

loss function incorporates the polynomial algebraic condition $\mathcal{L}_{\text{pol}}$, the locality penalty $\mathcal{L}_{\text{loc}}$ in Eq. (11) with the decay parameter $\alpha = 0.8$ and the lattice size $L = 32$, and the doubler-decoupling conditions $\mathcal{L}_{\text{pin}}$ in Eq. (10) as

$$\mathcal{L}_{\text{total}} = w_{\text{pol}}\mathcal{L}_{\text{pol}} + w_{\text{pin}}\mathcal{L}_{\text{pin}} + w_{\text{loc}}\mathcal{L}_{\text{loc}}. \quad (15)$$

For this autonomous discovery phase, the loss weights are set to $w_{\text{pol}} = 1.0$, $w_{\text{pin}} = 0.1$, and $w_{\text{loc}} = 0.01$, and the network is trained using the Adam optimizer with a learning rate of $\eta = 0.002$. The Softmax temperature τ is exponentially annealed from $\tau = 0.5$ down to $\tau = 0.01$ over the 10 000 epochs. The network must optimize its own algebraic rules to satisfy these basic requirements.

B. Results: Emergence of the GW relation

The dynamics of the optimization process illustrate how the proposed framework performs machine-assisted algebraic-structure search. Starting from random initial network weights, the initial outputs yield arbitrary, nonlocal spectra that entirely violate chiral symmetry. However, constrained by the generalized polynomial framework and guided purely by the necessity to fulfill the doubler-pinning condition ($D(k_{\text{doub}}) = 2$) and the spatial locality penalty, the network systematically traverses the loss landscape.

As training progresses over 10 000 epochs, the network smoothly transitions from a broad exploration of the generalized polynomial space to a sharp decision via temperature annealing in the Softmax function ($\tau = 0.5 \rightarrow 0.01$). During the early phases, the optimization suppresses higher-order terms, suggesting that they are disfavored by the combined polynomial, locality, and pinning losses in the present setup. Consequently, the Softmax gating dynamically suppresses the higher-order logits, completely isolating the $n = 1$ structural term ($c_{n>1} \rightarrow 0.000$). Simultaneously, the network incrementally drives the global scale from its naive starting point of 0.1 up to ~ 1.00 . As shown in Fig. 9, by the end of the 10 000 epochs, the resultant eigenvalue spectrum eventually converges to form a GW circle in the complex plane, confirming the emergence of the GW formulation. The near zero mode in this discovered operator deviates from the origin by only $\sim 4.9084 \times 10^{-3}$. Let us show the explicit results of coefficients

$$\begin{aligned} \text{Initial: } s_{\text{global}} &= 0.1, \quad c_n = 0.1 \times 0.2 \times 4^{-n+1} \quad (\alpha_n = 0), \\ \text{Final: } s_{\text{global}} &= 1.001, \quad c_1 = 1.001, \quad c_{n>1} \simeq 0.000. \end{aligned} \quad (16)$$

The other initial coefficients also result in the same final coefficients

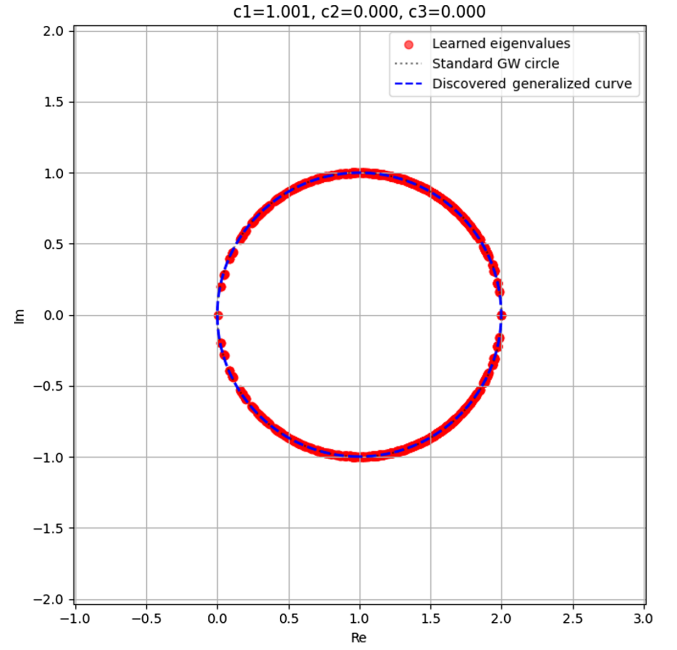


FIG. 9. The learned eigenvalue spectrum in 2D. Guided by the generalized polynomial constraint, spatial locality, and doubler-decoupling conditions, the network selects $s_{\text{global}} \rightarrow 1.001$ ($c_1 = 1.001$) and zeroes out higher-order coefficients ($c_{n>1} = 0.000$), indicating that the standard GW relation is selected as the preferred solution within the present polynomial optimization setup. The parameters are $\alpha = 0.8$, $L = 32$ and $w_{\text{pol}} = 1.0$, $w_{\text{pin}} = 0.1$, $w_{\text{loc}} = 0.01$.

$$\text{Initial: } s_{\text{global}} = 0.1, \quad c_1 = 0.01, \quad c_2 = 0.016,$$

$$c_3 = 0.001, \quad c_{4,5} \simeq 0.000 \quad (\alpha_2 = 1.0, \alpha_{n \neq 2} = 0),$$

$$\text{Final: } s_{\text{global}} = 1.001, \quad c_1 = 1.001, \quad c_{n>1} \simeq 0.000.$$

(17)

We emphasize that the final 2D Dirac operator learned in this autonomous phase reproduces the overlap operator obtained in Sec. IV, with a similar momentum-space profile and real-space exponential locality.

The network was not explicitly tasked with satisfying the Ginsparg-Wilson relation. Instead, starting from a generalized polynomial framework and naive initializations, the optimization selects the $c_1 \simeq 1$, $c_{n>1} \simeq 0$ structure as the preferred solution within the prescribed polynomial ansatz and the present loss function.

C. Discovery of the generalized GW relation

The successful autonomous derivation of the standard GW relation raises an intriguing question regarding the structure of the optimization landscape: does the generalized polynomial space contain other physically valid local minima, and can the network autonomously construct them? To investigate this, we maintain the same global scale ($s_{\text{global}} = 0.1$) but introduce a slight initial bias in the structural logits, starting the network from

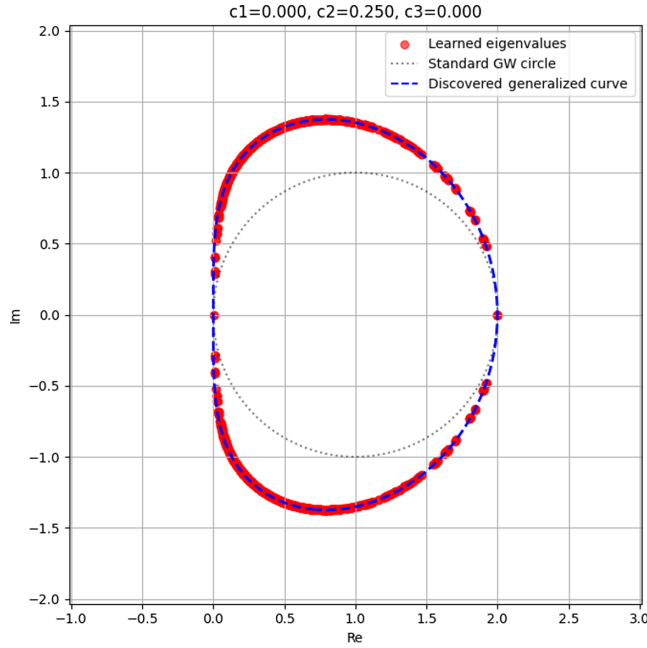


FIG. 10. The learned eigenvalue spectrum corresponding to the autonomous discovery of the generalized Fujikawa relation. By initializing the search with a slight bias toward the $n = 2$ term, the network zeroes out all other coefficients and precisely converges to $c_2 = 0.250$, $c_{n \neq 2} = 0$, showing that the same framework can converge to an alternative higher-order GW-type algebraic structure. The parameters are $\alpha = 0.8$, $L = 32$ and $w_{\text{pol}} = 1.0$, $w_{\text{pin}} = 0.1$, $w_{\text{loc}} = 0.01$.

$[\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5] = [0.0, 2.0, 0.0, 0.0, 0.0]$. This gently nudges the initial search space toward the $n = 2$ term.

The resulting training dynamics exhibits a second stable solution in the polynomial search space. As the epochs progress, the network systematically zeroes out the c_1 , c_3 , c_4 , and c_5 coefficients. Remarkably, by epoch 6 000, the network rigorously drives the global scale to 1.001 and precisely converges to the coefficient set $\{c_2 = 0.250, c_{n \neq 2} = 0.000\}$ as

$$\begin{aligned} \text{Initial: } s_{\text{global}} &= 0.1, \quad c_1 = 0.002, \quad c_2 = 0.023, \\ c_{3,4,5} &\simeq 0 \quad (\alpha_2 = 2.0, \quad \alpha_{n \neq 2} = 0), \\ \text{Final: } s_{\text{global}} &= 1.001, \quad c_2 = 0.250, \quad c_{n \neq 2} \simeq 0.000. \end{aligned} \quad (18)$$

Mathematically, this specific coefficient selection translates the generalized polynomial ansatz into

$$D + D^\dagger = \frac{1}{4}(D^\dagger D)^2. \quad (19)$$

This is not an arbitrary result; it is precisely the generalized Ginsparg-Wilson relation originally proposed by Fujikawa [36], which plays a pivotal role in the context of higher-order Dirac operators and index theorems. The corresponding eigenvalue spectrum, depicted in Fig. 10, confirms a symmetric, topologically valid structure distinct

from the standard GW circle. Here again, the near zero mode is pinned, with an offset of only $\sim 5.3882 \times 10^{-3}$.

This result suggests that the optimization landscape contains more than the standard GW solution within the chosen polynomial ansatz. Instead, it can access different algebraic structures depending on the initial search bias. Thus, by altering the initial search bias, the same optimization framework can converge to distinct GW-type algebraic structures within the prescribed polynomial ansatz.

VI. CONCLUSION

We have proposed a novel framework that utilizes physics-informed neural networks to systematically construct, optimize, and theoretically explore lattice fermions. By integrating doubler-decoupling conditions, spatial locality, and symmetry constraints into the loss function, we treated the formulation of the Dirac operator as a constraint-driven optimization problem.

We first demonstrated that a neural network can learn the overlap mapping from the Wilson kernel to high numerical accuracy when trained with the GW relation as a soft constraint. As a proof of concept, we confirmed that the framework extends to 4D, where the learned spectrum satisfies the GW-circle constraint. Furthermore, we demonstrated the stability of our framework by analyzing its robustness against variations in the loss weights and its sensitivity to the locality decay parameter.

Most importantly, we elevated this framework from operator construction to machine-assisted algebraic discovery. Within a generalized polynomial ansatz and with an agnostic initial scale, the neural network autonomously suppresses higher-order structural terms and recovers the standard Ginsparg-Wilson relation, including its overall normalization. Furthermore, by exploring different initial search biases, the same framework finds a distinct algebraic structure corresponding to a Fujikawa-type generalized GW relation. These results indicate that the standard and generalized GW relations appear as preferred solutions in the present optimization landscape, selected by the combined requirements of locality, doubler decoupling, and polynomial algebraic consistency.

Looking ahead, there are several promising directions for future research. First, extending this framework to the Hamiltonian formalism will be crucial for tensor network simulations and quantum computing applications. Second, applying this methodology to minimally doubled fermions could systematically identify optimal kernels where specific discrete symmetries such as $\mathcal{C}$, $\mathcal{P}$, $\mathcal{T}$ or hypercubic symmetries are partially relaxed. Third, exploring one-component fermions like staggered fermions under this machine-learning-driven framework may reveal new classes of highly localized operators. Finally, incorporating background gauge fields directly into the training process remains an important next step. These efforts will further establish deep learning as an indispensable theoretical tool

for advancing the fundamental mathematical structures of quantum field theory.

ACKNOWLEDGMENTS

The author would like to thank Norihiro Tanahashi and Akio Tomiya for encouraging the author to publish the present work. The author also thanks the seminar “Deep Learning and Physics” and the workshop “Topology in Lattice Systems 2026” for providing opportunities of useful discussions. This work was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grants No. 23K03425 and No. 22H05118.

DATA AVAILABILITY

The author welcomes discussions on the computational framework and is happy to provide the Python source code used in this study to interested readers upon request.

APPENDIX A: STABILITY AGAINST VARIATIONS IN LOSS WEIGHTS

To verify that the learned operator is not a delicate artifact of a specific choice of hyperparameters, we investigated the stability of the optimization with respect to variations in the loss weights (w_{GW} , w_{pin} , w_{loc}).

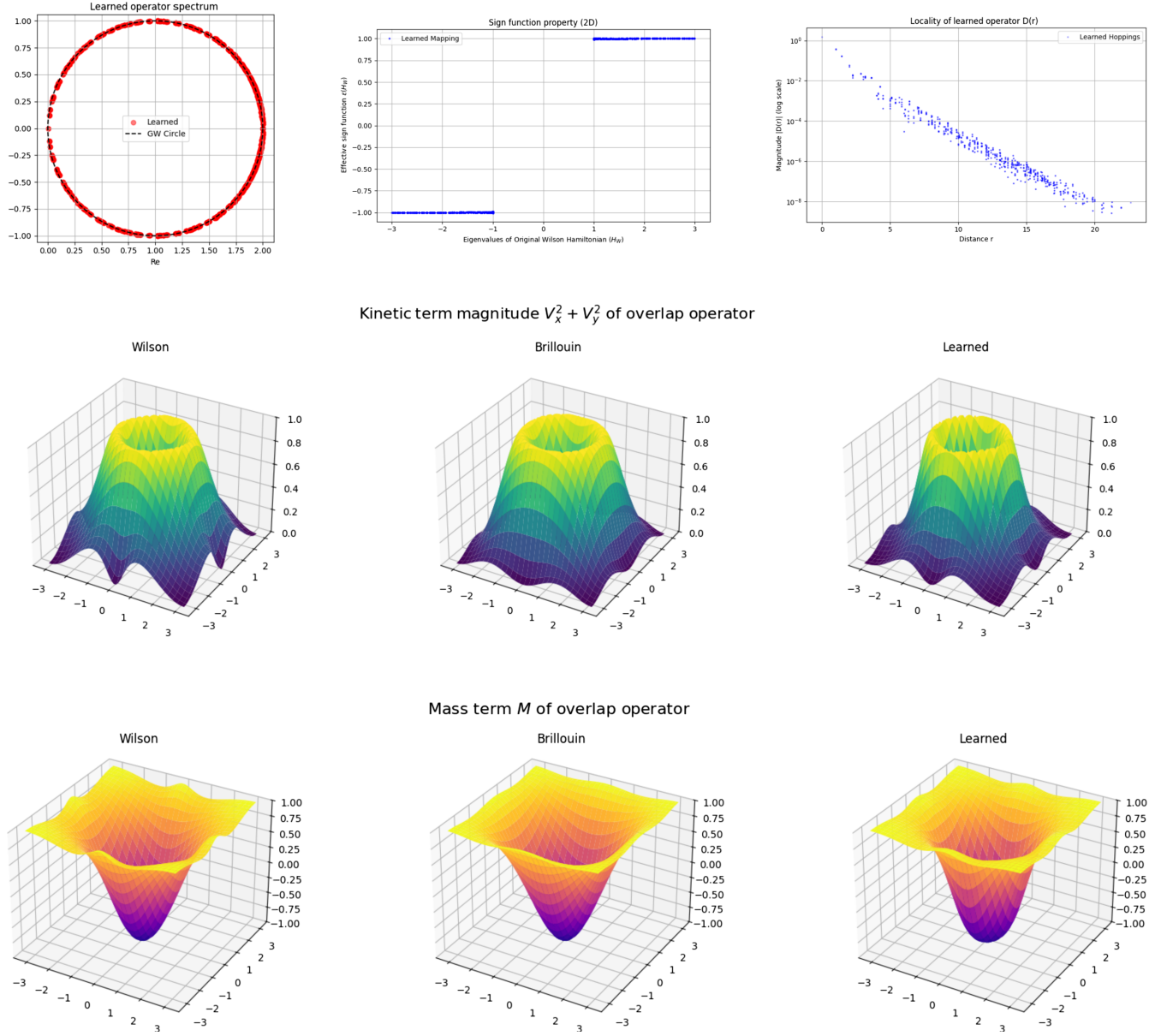


FIG. 11. Properties of the learned operator trained with a reduced locality penalty ($w_{\text{GW}} = 1.0$, $w_{\text{pin}} = 0.1$, $w_{\text{loc}} = 0.01$) with $\alpha = 0.8$ and $L = 32$ in 2D. It seems to converge to an overlap operator, as evidenced by the eigenvalue spectrum and the effective sign function. The dispersion relations exhibit flattened basins near the Brillouin zone boundaries. The optimization of the momentum-space dispersion comes at the cost of real-space locality.

Specifically, we set $w_{\text{GW}} = 1.0$, $w_{\text{pin}} = 0.1$, and $w_{\text{loc}} = 0.01$, and performed the same training procedure for $\alpha = 0.8$ and $L = 32$ in 2D.

As shown in Fig. 11, we confirmed that even with this hyperparameter set, the learned operator consistently seems to converge to an overlap operator. Furthermore, the dispersion relation for this set exhibits a relatively flat behavior near the Brillouin zone boundary, yielding a more Brillouin-fermion-like overlap operator. However, due to the reduced locality weight $\mu \sim 0.95$, the spatial locality degrades compared to the baseline case ($w_{\text{GW}} = 1.0$, $w_{\text{pin}} = 0.1$, $w_{\text{loc}} = 0.1$).

APPENDIX B: DEPENDENCE ON THE LOCALITY DECAY PARAMETER α

In contrast to the robustness against the loss weights, we found that the optimization landscape is sensitive to the choice of the locality decay parameter α . To investigate this dependence, we systematically varied the exponent from $\alpha = 0.5$ to 0.9 with $w_{\text{GW}} = 1.0$, $w_{\text{pin}} = 0.1$, $w_{\text{loc}} = 0.1$ and $L = 32$ fixed in 2D.

As illustrated in Fig. 12, when the locality penalty is relatively weak such as $\alpha = 0.5, 0.6, 0.7$, the network seems to successfully converge to the overlap operator,

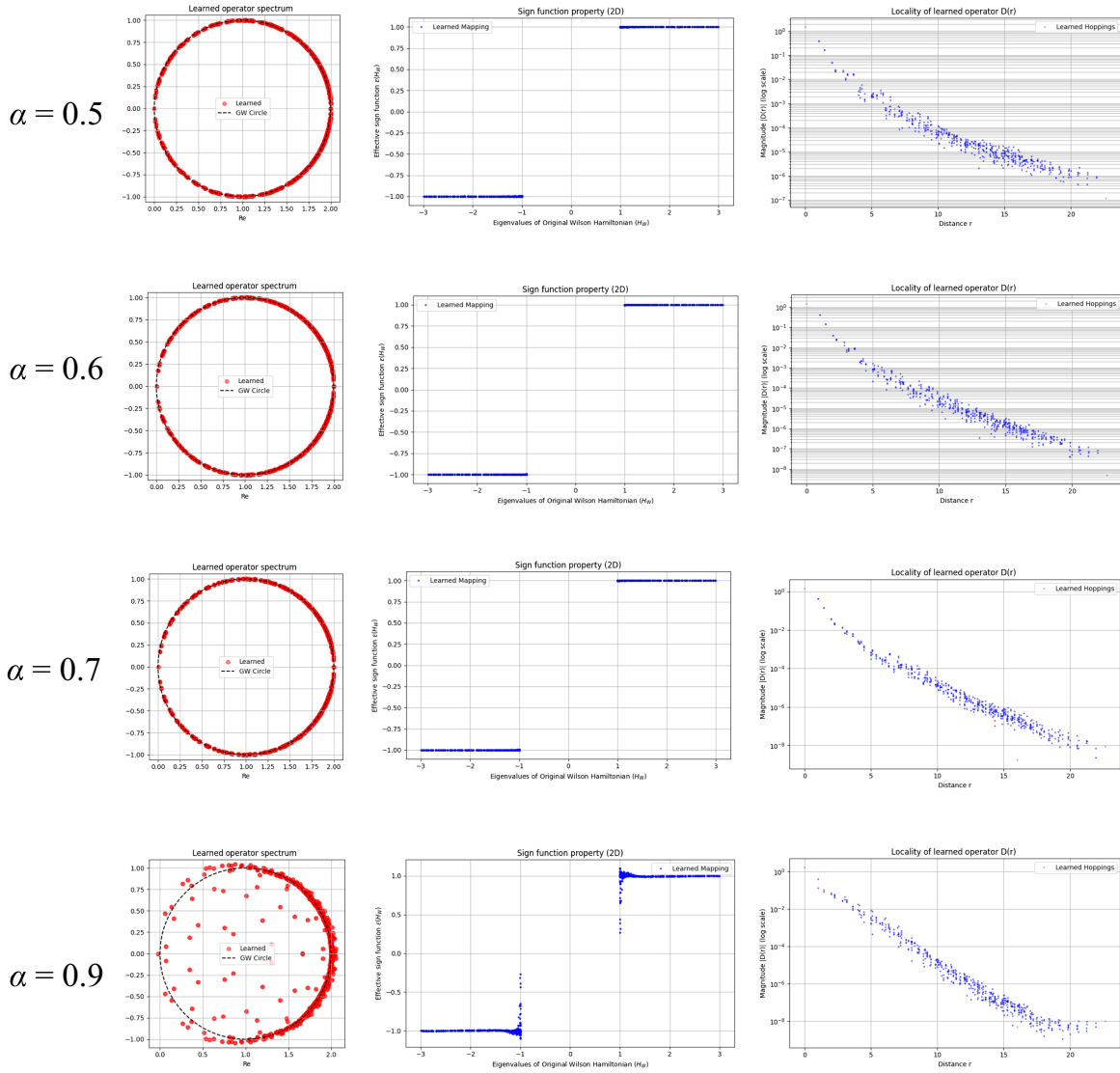


FIG. 12. The eigenvalue spectra, the effective sign functions and the magnitudes of the real-space hopping parameters $|D(r)|$ for $\alpha = 0.5, 0.6, 0.7, 0.9$ with $w_{\text{GW}} = 1.0$, $w_{\text{pin}} = 0.1$, $w_{\text{loc}} = 0.1$ and $L = 32$ fixed. For $\alpha = 0.9$, the network fails to reproduce the Ginsparg-Wilson circle of spectrum while it does for $\alpha = 0.5, 0.6, 0.7$.

reproducing both the Ginsparg-Wilson circle and the topological sign function. However, the spatial locality naturally degrades as α decreases. Conversely, when the penalty is excessively strong such as $\alpha = 0.9$, the network either fails to reproduce the Ginsparg-Wilson relation entirely, or the training process itself

becomes highly unstable and fails to converge. This strong sensitivity clearly highlights that a careful adjustment of the real-space locality constraint (around $\alpha \sim 0.8$) is a crucial ingredient for guiding the machine-learning framework to valid chiral lattice fermion formulations.

-
- [1] L. H. Karsten and J. Smit, Lattice fermions: Species doubling, chiral invariance, and the triangle anomaly, *Nucl. Phys.* **B183**, 103 (1981).
 - [2] H. B. Nielsen and M. Ninomiya, Absence of neutrinos on a lattice. 1. Proof by homotopy theory, *Nucl. Phys.* **B185**, 20 (1981).
 - [3] H. B. Nielsen and M. Ninomiya, Absence of neutrinos on a lattice. 2. Intuitive topological proof, *Nucl. Phys.* **B193**, 173 (1981).
 - [4] H. B. Nielsen and M. Ninomiya, No go theorem for regularizing chiral fermions, *Phys. Lett.* **105B**, 219 (1981).
 - [5] K. G. Wilson, Quarks and strings on a lattice, in *New Phenomena in Subnuclear Physics: Proceedings, International School of Subnuclear Physics, Erice, Sicily, 1975. Part A* (Plenum Press, New York, NY, 1975), pp. 69–142, [10.1007/978-1-4613-4208-3-6](#).
 - [6] J. B. Kogut and L. Susskind, Hamiltonian formulation of Wilson's lattice gauge theories, *Phys. Rev. D* **11**, 395 (1975).
 - [7] L. Susskind, Lattice fermions, *Phys. Rev. D* **16**, 3031 (1977).
 - [8] P. H. Ginsparg and K. G. Wilson, A remnant of chiral symmetry on the lattice, *Phys. Rev. D* **25**, 2649 (1982).
 - [9] M. Luscher, Exact chiral symmetry on the lattice and the Ginsparg-Wilson relation, *Phys. Lett. B* **428**, 342 (1998).
 - [10] H. Neuberger, More about exactly massless quarks on the lattice, *Phys. Lett. B* **427**, 353 (1998).
 - [11] D. B. Kaplan, A method for simulating chiral fermions on the lattice, *Phys. Lett. B* **288**, 342 (1992).
 - [12] Y. Shamir, Chiral fermions from lattice boundaries, *Nucl. Phys.* **B406**, 90 (1993).
 - [13] V. Furman and Y. Shamir, Axial symmetries in lattice QCD with Kaplan fermions, *Nucl. Phys.* **B439**, 54 (1995).
 - [14] T.-W. Chiu, T.-H. Hsieh, C.-H. Huang, and T.-R. Huang, A note on the Zolotarev optimal rational approximation for the overlap Dirac operator, *Phys. Rev. D* **66**, 114502 (2002).
 - [15] T.-W. Chiu, Optimal domain wall fermions, *Phys. Rev. Lett.* **90**, 071601 (2003).
 - [16] C. Gattringer, A new approach to Ginsparg-Wilson fermions, *Phys. Rev. D* **63**, 114501 (2001).
 - [17] W. Bietenholz, Solutions of the Ginsparg-Wilson relation and improved domain wall fermions, *Eur. Phys. J. C* **6**, 537 (1999).
 - [18] W. Bietenholz and I. Hip, The scaling of exact and approximate Ginsparg-Wilson fermions, *Nucl. Phys.* **B570**, 423 (2000).
 - [19] S. Dürr and G. Koutsou, Brillouin improvement for Wilson fermions, *Phys. Rev. D* **83**, 114512 (2011).
 - [20] S. Dürr, G. Koutsou, and T. Lippert, Meson and baryon dispersion relations with brillouin fermions, *Phys. Rev. D* **86**, 114514 (2012).
 - [21] Y.-G. Cho, S. Hashimoto, J.-I. Noaki, A. Jüttner, and M. Marinkovic, $O(a^2)$ -improved actions for heavy quarks and scaling studies on quenched lattices, *Proc. Sci. LATTICE2013* (2014) 255 [[arXiv:1312.4630](#)].
 - [22] Y.-G. Cho, S. Hashimoto, A. Jüttner, T. Kaneko, M. Marinkovic, J.-I. Noaki, and J. T. Tsang, Improved lattice fermion action for heavy quarks, *J. High Energy Phys.* **05** (2015) 072.
 - [23] S. Dürr and G. Koutsou, On the suitability of the brillouin action as a kernel to the overlap procedure, [arXiv:1701.00726](#).
 - [24] M. Creutz, T. Kimura, and T. Misumi, Index theorem and overlap formalism with naive and minimally doubled fermions, *J. High Energy Phys.* **12** (2010) 041.
 - [25] T. Misumi, New fermion discretizations and their applications, *Proc. Sci. LATTICE2012* (2012) 005 [[arXiv:1211.6999](#)].
 - [26] M. Raissi, P. Perdikaris, and G. E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *J. Comput. Phys.* **378**, 686 (2019).
 - [27] D. Boyda *et al.*, Applications of machine learning to lattice quantum field theory, in *Snowmass 2021* (2022), 2; [arXiv:2202.05838](#).
 - [28] A. Tomiya, Machine learning for lattice QCD, *J. Phys. Soc. Jpn.* **94**, 031006 (2025).
 - [29] U. Wenger, Machine learning for four-dimensional SU(3) lattice gauge theories, *Proc. Sci. LATTICE2025* (2026) 0020.
 - [30] K. Holland, A. Ipp, D. I. Müller, and U. Wenger, Fixed point actions from convolutional neural networks, *Proc. Sci. LATTICE2023* (2024) 038 [[arXiv:2311.17816](#)].
 - [31] K. Holland, A. Ipp, D. I. Müller, and U. Wenger, Machine learning a fixed point action for SU(3) gauge theory with a gauge equivariant convolutional neural network, *Phys. Rev. D* **110**, 074502 (2024).
 - [32] U. Wenger, K. Holland, A. Ipp, and D. I. Müller, HMC and gradient flow with machine-learned classically perfect fixed point actions, *Proc. Sci. LATTICE2024* (2025) 466 [[arXiv:2502.03315](#)].

- [33] K. Holland, A. Ipp, D. I. Müller, and U. Wenger, Machine-learned renormalization-group-Improved gauge actions and classically perfect gradient flows, *Phys. Rev. Lett.* **136**, 031901 (2026).
- [34] S. Yasunaga, K. Yoshimura, A. Tomiya, and Y. Nagai, Parameter optimization of domain-wall fermion using machine learning, in *42th International Symposium on Lattice Field Theory* (2026), 3;arXiv:2603.16329.
- [35] H. Liu, K. Simonyan, and Y. Yang, DARTS: Differentiable architecture search, arXiv:1806.09055.
- [36] K. Fujikawa, Algebraic generalization of the Ginsparg-Wilson relation, *Nucl. Phys.* **B589**, 487 (2000).
- [37] M. Clancy, D. B. Kaplan, and H. Singh, Generalized Ginsparg-Wilson relations, *Phys. Rev. D* **109**, 014502 (2024).
- [38] P. Hasenfratz and F. Niedermayer, Perfect lattice action for asymptotically free theories, *Nucl. Phys.* **B414**, 785 (1994).
- [39] W. Bietenholz and U. J. Wiese, Perfect lattice actions for quarks and gluons, *Nucl. Phys.* **B464**, 319 (1996).
- [40] W. Bietenholz, Optimised Dirac operators on the lattice: Construction, properties and applications, *Fortschr. Phys.* **56**, 107 (2008).
- [41] L. H. Karsten, Lattice fermions in Euclidean space-time, *Phys. Lett.* **104B**, 315 (1981).
- [42] F. Wilczek, On lattice fermions, *Phys. Rev. Lett.* **59**, 2397 (1987).
- [43] M. Creutz, Four-dimensional graphene and chiral fermions, *J. High Energy Phys.* **04** (2008) 017.
- [44] A. Borici, Creutz fermions on an orthogonal lattice, *Phys. Rev. D* **78**, 074504 (2008).
- [45] P. F. Bedaque, M. I. Buchoff, B. C. Tiburzi, and A. Walker-Loud, Broken symmetries from minimally doubled fermions, *Phys. Lett. B* **662**, 449 (2008).
- [46] P. F. Bedaque, M. I. Buchoff, B. C. Tiburzi, and A. Walker-Loud, Search for fermion actions on hyperdiamond lattices, *Phys. Rev. D* **78**, 017502 (2008).
- [47] S. Capitani, J. Weber, and H. Wittig, Minimally doubled fermions at one loop, *Phys. Lett. B* **681**, 105 (2009).
- [48] T. Kimura and T. Misumi, Characters of lattice fermions based on the hyperdiamond lattice, *Prog. Theor. Phys.* **124**, 415 (2010).
- [49] T. Kimura and T. Misumi, Lattice fermions based on higher-dimensional hyperdiamond lattices, *Prog. Theor. Phys.* **123**, 63 (2010).
- [50] M. Creutz and T. Misumi, Classification of minimally doubled fermions, *Phys. Rev. D* **82**, 074502 (2010).
- [51] M. Creutz, Minimal doubling and point splitting, *Proc. Sci. LATTICE2010* (2010) 078 [arXiv:1009.3154].
- [52] S. Capitani, M. Creutz, J. Weber, and H. Wittig, Renormalization of minimally doubled fermions, *J. High Energy Phys.* **09** (2010) 027.
- [53] B. C. Tiburzi, Chiral lattice fermions, minimal doubling, and the axial anomaly, *Phys. Rev. D* **82**, 034511 (2010).
- [54] S. Kamata and H. Tanaka, Minimal doubling fermion and hermiticity, *Prog. Theor. Exp. Phys.* **2013**, 023B05 (2013).
- [55] T. Misumi, Phase structure for lattice fermions with flavored chemical potential terms, *J. High Energy Phys.* **08** (2012) 068.
- [56] T. Misumi, T. Kimura, and A. Ohnishi, QCD phase diagram with 2-flavor lattice fermion formulations, *Phys. Rev. D* **86**, 094505 (2012).
- [57] S. Capitani, Reducing the number of counterterms with new minimally doubled actions, *Phys. Rev. D* **89**, 014501 (2014).
- [58] S. Capitani, New chiral lattice actions of the Borici-Creutz type, *Phys. Rev. D* **89**, 074508 (2014).
- [59] T. Misumi, Fermion actions extracted from lattice super Yang-Mills theories, *J. High Energy Phys.* **12** (2013) 063.
- [60] J. H. Weber, S. Capitani, and H. Wittig, Numerical studies of minimally doubled fermions, *Proc. Sci. LATTICE2013* (2014) 122 [arXiv:1312.0488].
- [61] J. H. Weber, Properties of minimally doubled fermions., Ph.D. thesis, Mainz U., 2015, arXiv:1706.07104.
- [62] S. Dürr and J. H. Weber, Dispersion relation and spectral range of Karsten-Wilczek and Borici-Creutz fermions, *Phys. Rev. D* **102**, 014516 (2020).
- [63] Y.-Y. Li, J. Wang, and Y.-Z. You, Quantum many-body lattice C-R-T symmetry: Fractionalization, anomaly, and symmetric mass generation, arXiv:2412.19691.
- [64] A. Chatterjee, S. D. Pace, and S.-H. Shao, Quantized axial charge of staggered fermions and the chiral anomaly, *Phys. Rev. Lett.* **134**, 021601 (2025).
- [65] S. Catterall and A. Pradhan, Symmetries and anomalies of Hamiltonian staggered fermions, *Phys. Rev. D* **113**, 014504 (2026).
- [66] T. Yamaoka, Quantized axial charge in the Hamiltonian approach to Wilson fermions, *J. High Energy Phys.* **10** (2025) 102.
- [67] T. Onogi and T. Yamaoka, Non-singlet conserved charges and anomalies in $3 + 1D$ staggered fermions, arXiv:2509.04906.
- [68] S. Aoki, Y. Kikukawa, and T. Takemoto, Chiral anomaly of Kogut-Susskind fermion in $(3 + 1)$ -dimensional Hamiltonian formalism, *Phys. Rev. D* **113**, 034514 (2026).
- [69] N. Seiberg and W. Zhang, Tori, klein bottles, and modulo 8 parity/time-reversal anomalies of $2 + 1d$ staggered fermions, *J. High Energy Phys.* **05** (2026) 264.
- [70] T. Misumi, Minimal-doubling and single-Weyl fermion Hamiltonians, *Phys. Rev. D* **113**, 074521 (2026).
- [71] T. Misumi, T. Onogi, and T. Yamaoka, Taste-splitting mass and edge modes in $3 + 1D$ staggered fermions, arXiv:2604.02078.
- [72] P. Hernandez, K. Jansen, and M. Luscher, Locality properties of neuberger's lattice Dirac operator, *Nucl. Phys.* **B552**, 363 (1999).
- [73] I. Horvath, Ginsparg-Wilson relation and ultralocality, *Phys. Rev. Lett.* **81**, 4063 (1998).
- [74] W. Bietenholz, On the absence of ultralocal Ginsparg-Wilson fermions, arXiv:hep-lat/9901005.