PROCEEDINGS OF SCIENCE

Parameter Optimization of Domain-Wall Fermion using Machine Learning

Shunsuke Yasunaga,[a,b,*] **Kenta Yoshimura,**[a,b] **Akio Tomiya**[c,b,d] **and Yuki Nagai**[e,f]

[a] *Department of physics, Institute of Science Tokyo, Meguro, Tokyo 152-8550, Japan*

[b] *RIKEN Center for Computational Science, Kobe, Hyogo 650-0047, Japan*

[c] *Department of Information and Mathematical Sciences, Tokyo Woman's Christian University, Suginami, Tokyo 167-8585, Japan*

[d] *Department of Physics, Kyoto University, Kyoto, Kyoto 606-8502, Japan*

[e] *Information Technology Center, The University of Tokyo, Kashiwa, Chiba 277–0882, Japan*

[f] *Department of Advanced Materials Science, The University of Tokyo, Kashiwa, Chiba 277–8561, Japan*

We study a parameter optimization of domain-wall fermions to improve chiral symmetry based on machine learning. Domain-wall fermions involve coefficients along the fifth dimension, which can be treated as trainable parameters to reduce the chiral symmetry violation caused by the finite extent of the fifth dimension. As the loss function, we use the residual mass estimated stochastically on a single gauge configuration. Numerical tests on a $L^3 \times T \times L_5 = 4^3 \times 8 \times 8$ lattice demonstrate the feasibility of this framework.

The 42nd International Symposium on Lattice Field Theory (LATTICE2025)
3-8 November 2025
Mumbai, India

*Speaker

PoS(LATTICE2025)035

1. Introduction

Chiral symmetry and its spontaneous breaking are fundamental concepts in quantum chromodynamics (QCD). However, in the lattice formulation, which provide one of the most promising regularizations of nonperturbative field theory, cannot preserve chiral symmetry in a straightforward way because of the Nielsen-Ninomiya theorem [1]. This obstruction motivates some descretizations of the lattice fermions. One powerful approach is the domain-wall fermion formulation, which realizes good chiral symmetry on the lattice by introducing an additional fifth dimension. In the limit where the extent of the fifth dimension becomes infinite, domain-wall fermions satisfy the Ginsparg-Wilson relation [2] and possess an exact lattice chiral symmetry even at finite lattice spacing.

Domain-wall fermions were originally proposed as a formulation of chiral fermions on the lattice, motivated by chiral gauge theories as in the Standard Model [3]. Subsequently, their application for lattice QCD with chiral symmetry was developed [4, 5], and nowadays domain-wall fermions are widely employed as a chiral symmetric discretization in the lattice QCD. Tuning of domain-wall coefficients has been explored as a practical way to improve chiral symmetry in domain-wall fermions. Dirac operators satisfying the Ginsparg-Wilson relation can be expressed in terms of the sign function [6]. The Zolotarev approximation provides a rational approximation to the sign function and enables a highly accurate realization of chiral symmetry by optimization of the coefficients appearing in the numerator [7, 8]. In contrast, the Möbius domain-wall formulation introduces a transformation in the denominator as well [9]. Furthermore, an extension to complex coefficients, known as zMöbius, has also been developed [10].

Machine learning has recently been explored in lattice QCD as a general framework to improve the efficiency of numerical simulations [11, 12]. Because it provides a flexible way to handle complicated optimization problems in high-dimensional parameter spaces, it may also be useful for tuning internal parameters of lattice actions. Nevertheless, applications aimed at improving physical properties have so far remained relatively limited. For domain-wall fermions, the coefficients in the fifth-dimensional structure control the approximation to the sign function and hence the quality of chiral symmetry. This feature makes them a natural target for optimization. In this work, we investigate their optimization using a machine-learning-based framework.

2. Method

2.1 Parameters of domain-wall fermion

The five-dimensional domain-wall Dirac operator D_5 is written as

$$(D_5)_{nm,\,st} = (D_{\mathrm{W}})_{nm}(b_s\delta_{st} + c_s F_{st}) + \delta_{nm}(\delta_{st} - F_{st}), \tag{1}$$

where the subscripts n, m denote four-dimensional lattice sites, and $s, t = 1, \ldots, L_5$ label the slices in the fifth dimension. The operator D_{W} denotes the Wilson–Dirac operator, while the matrix F_{st} represents hopping in the fifth dimension and implements the fermion with chiral symmetry on the boundaries of the fifth dimension. The coefficients b_s and c_s control the mapping from the five-dimensional formulation to the corresponding effective four-dimensional operator. In practice, the effective four-dimensional operator can be expressed in terms of a product of transfer matrices along the fifth dimension. This product encodes the approximation to the sign function, and thus determines the quality of chiral symmetry. The transfer matrix is constructed from the Hermitian operator H_s defined by

$$H_s = \gamma_5 \frac{(b_s + c_s)D_{\mathrm{W}}}{2 + (b_s - c_s)D_{\mathrm{W}}}. \tag{2}$$

Typically, the coefficients b_s and c_s are taken to be independent of s; in this case they are referred to as the Möbius parameters [9]. One may also allow only the combination $b_s + c_s$ in the numerator of Eq. (2) to depend on s. This makes it possible to improve the chiral symmetry by optimizing the rational approximation to the sign function, as in the Zolotarev approximation [7, 8]. In this work, we generalize the coefficients by allowing both b_s and c_s to depend on s, and treat them as independent parameters on each slice in the fifth dimension.

2.2 Residual mass

The violation of chiral symmetry due to lattice artifacts is quantified by the residual mass m_{res}. From the axial Ward–Takahashi identity, m_{res} can be defined as the plateau value of the ratio of appropriate pseudoscalar correlation functions [13–15]. It is also useful to introduce the "global" residual mass, which can be evaluated without identifying a plateau in Euclidean time. Following Ref. [8], the global residual mass can be expressed in terms of the effective four-dimensional operator D_4 as

$$m_{\mathrm{res}} = \frac{\mathrm{Re}\left\langle \mathrm{Tr}\, D_4^{-1} \right\rangle}{\left\langle \mathrm{Tr}\left(D_4^\dagger D_4\right)^{-1} \right\rangle} - m_q, \tag{3}$$

where m_q is the input bare quark mass. The traces are taken over the four-dimensional lattice sites as well as spin and color indices. Note that these two definitions can differ in the presence of near-zero modes [16]. The effective four-dimensional operator D_4, with the contact term subtracted, can be obtained from the boundary of the fifth dimension as [17]

$$D_4^{-1} = \left[P^\dagger D_5^{-1} R P\right]_{(s,t)=(1,1)}, \tag{4}$$

where P ($P_{st} = P_L \delta_{st} + P_R \delta_{s+1,t}$ for $s \neq L_5$ and $P_{L_5,t} = P_L \delta_{t,L_5} + P_R \delta_{t,1}$) and R ($R_{st} = \delta_{s,L_5+1-t}$) denote the chiral projection and reflection operators, respectively.

2.3 Loss function and its deriavtive

Machine learning provides a systematic optimization framework in which one specifies a loss (objective) function and iteratively updates the model parameters to minimize it. In gradient-based methods, the parameters are updated using the gradient of the loss with respect to those parameters.

To apply this framework to domain-wall fermions, we treat the domain-wall coefficients b_s and c_s as parameters to be optimized and choose a loss function that quantifies the violation of chiral symmetry. As a per-configuration loss, we use the global residual mass in Eq. (3) evaluated on a single gauge configuration:

$$\mathcal{L} = \mathcal{M}^2 = \left(\frac{\mathrm{Re}\,\mathrm{Tr}\, D_4^{-1}}{\mathrm{Tr}\left(D_4^\dagger D_4\right)^{-1}} - m_q \right)^2. \tag{5}$$

Although one can take a configuration average using a mini-batch of gauge configurations, in this work we update the parameters on a per-configuration basis.

Next, we calculate the gradient of the loss function with respect to the fifth-dimensional vectors b_s and c_s. Using a set of noise vectors η_k, the numerator and denominator of the loss function are rewritten as stochastic estimators

$$\mathcal{M} = \frac{\mathcal{N}}{\mathcal{D}} = \frac{\mathrm{Re}\left(\sum_k \eta_k^\dagger P^\dagger D_5^{-1} R P \eta_k\right)}{\sum_k \eta_k^\dagger \left(D_5^{-1} R P\right)^\dagger Q \left(D_5^{-1} R P\right) \eta_k}, \tag{6}$$

where $Q = \mathrm{diag}(P_L, 0, \dots, 0, P_R)$ is a projection operator which acts only on the boundaries of fifth dimension. The derivatives with respect to the domain-wall parameters follow from Eq. (1):

$$\frac{\partial (D_5)_{tu}}{\partial b_s} = \delta_{st}\delta_{tu} D_{\mathrm{W}}, \tag{7}$$

$$\frac{\partial (D_5)_{tu}}{\partial c_s} = \delta_{st} F_{tu} D_{\mathrm{W}}. \tag{8}$$

After summing over the index t, these derivatives can be regarded as the (s, u) components of a matrix acting in the fifth-dimensional space. Since the t index is contracted, the matrix products surrounding $\boldsymbol{\nabla}_{b,c} D_5$ can be separated into left and right factors. Then, the gradients of N and D can be written as

$$\boldsymbol{\nabla}_{b,c}\mathcal{N} = \mathrm{Re}\sum_k \left[\left(D_5^\dagger\right)^{-1} P\eta_k\right]^* \odot \left[\left(\boldsymbol{\nabla}_{b,c} D_5\right)\chi_k\right], \tag{9}$$

$$\boldsymbol{\nabla}_{b,c}\mathcal{D} = 2\mathrm{Re}\sum_k \left[\left(D_5^\dagger\right)^{-1} Q\chi_k\right]^* \odot \left[\left(\boldsymbol{\nabla}_{b,c} D_5\right)\chi_k\right], \tag{10}$$

where $\odot$ denotes the Hadamard product of vectors in the fifth dimension, and $\chi_k = D_5^{-1} R P \eta_k$.

3. Numerical setup

We compute parameter updates using the adaptive moment estimation (Adam) optimizer implemented in `Optimisers.jl`, with the learning rate set to 10^{-2}. Both the loss function and its derivatives are evaluated using ten independent noise vectors for each learning epoch. The four-dimensional lattice size is $L^3 \times T = 4^3 \times 8$, and we use the Wilson gauge action at $\beta = 6.0$. For the domain-wall fermion, we use a bare quark mass $m_q a = 0.02$ and Pauli–Villars mass $m_{\mathrm{PV}} = 1.0$, domain-wall height $M_5 = 1.9$ ($\kappa = 0.2381$), and fifth-dimensional extent $L_5 = 8$. Simulations and measurements are carried out using `LatticeQCD.jl`, which we extended to implement residual-mass measurement, gradient computation, and parameter updates. In the present setup, parameter updates and measurements are performed alternately for each gauge configuration; namely, after generating a configuration, we evaluate the loss and its gradient on that configuration and then update the parameters before proceeding to the next configuration.

We consider two parameterizations of the domain-wall kernel along the fifth direction, namely the *Möbius* setting and the *general* setting. In the Möbius setting, the coefficients are taken to be uniform in the fifth dimension. On the other hand, in the general setting, independent parameters are assigned to each fifth-dimensional slice, so that the parameter space is a $2L_5$-dimensional space spanned by $\{b_s, c_s\}$. Unless otherwise stated, the initial parameters are set to $b_s = 1.5$, $c_s = 0.5$, which corresponds to the so-called scaled-Shamir kernel.

4. Preliminary results

4.1 Decreasing of the loss function

Figure 1 shows the evolution of the loss function as a function of the training step for the Möbius and general settings. In both cases, the loss decreases stably as the optimization proceeds, indicating that the training behaves as intended and improves the target quantity associated with chiral symmetry. A comparison between the general and Möbius settings reveals that allowing independent parameters for each fifth-dimensional slice leads to a systematically smaller residual mass M_{res} than imposing a uniform value across all slices. In other words, within the $2L_5$-dimensional parameter space spanned by $\{b_s, c_s\}$, there exist configurations that achieve better chiral symmetry than those lying on the Möbius constraint surface.

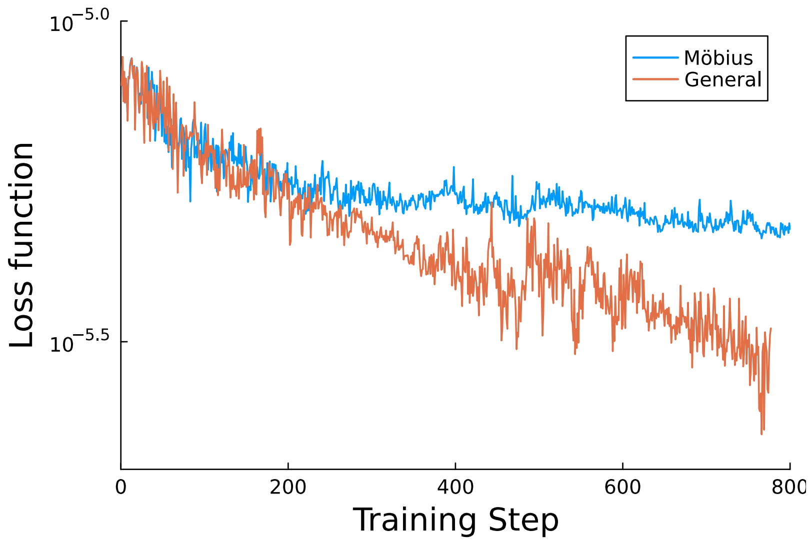

Figure 1: Training history of the loss function for the Möbius and general settings.

4.2 Evolution of the optimized parameters

Figures 2 and 3 summarize the evolution of the optimized parameters in the Möbius and general settings. In Fig. 2, the two Möbius parameters b and c are displayed in a single panel. In contrast, Fig. 3 shows, in the left panel, each fifth-dimensional component of b_s, while the right panel is dedicated to c_s in the general setting. From Fig. 3, two characteristic features are observed. Firstly, the dominant variations occur near the boundaries of the fifth dimension, i.e., at $s = 1$ and $s = L_5$, whereas the bulk parameters remain comparatively stable throughout the optimization. This behavior is consistent with patterns known from Zolotarev-type optimizations, where endpoint coefficients play a central role, and it suggests that the boundary slices control most of the effective freedom relevant for improving chiral symmetry. Secondly, b_s and c_s exhibit qualitatively different

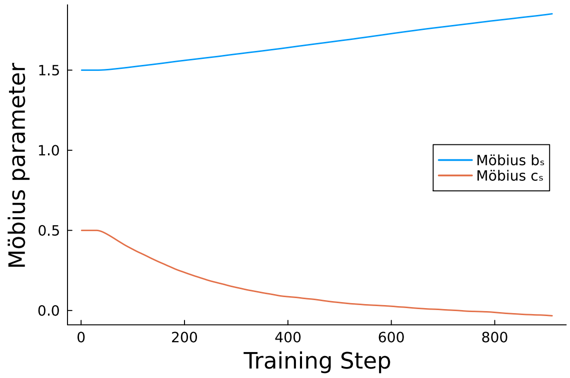

Figure 2: Training history of the Möbius parameters b and c.

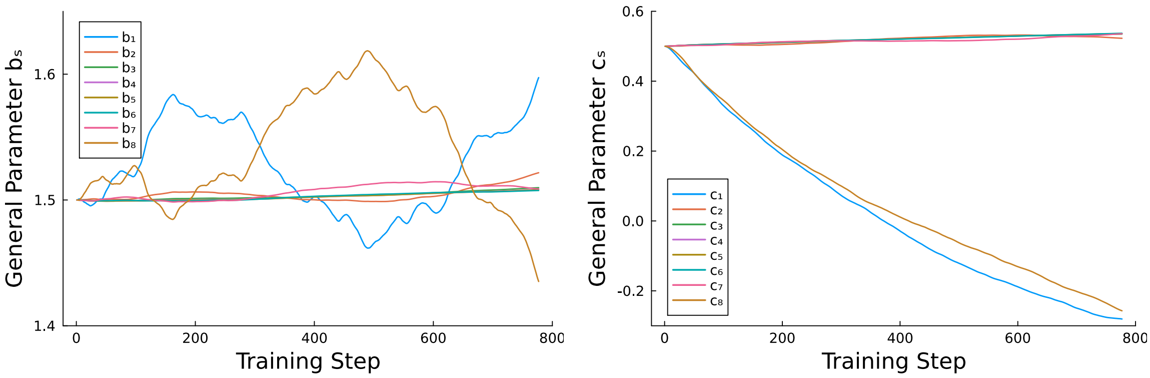

Figure 3: Training histories of the slice-dependent domain-wall parameters b_s (left) and c_s (right) in the general setting.

convergence behavior. In contrast to endpoint components of b_s fluctuating during training, c_s shows an approximately monotonic drift over the explored training range and does not exhibit clear saturation, even when the training is extended. This contrast suggests that the loss function is more sensitive to variations in b_s than in c_s. In particular, the parameter subspace associated with c_s may contain relatively flat directions, along which the loss changes only weakly, leading to a slow drift rather than rapid convergence. Additionally, we find that highly negative c_s values lead to convergence instabilities of the conjugate-gradient (CG) solver for the Dirac operator. This indicates that, for a stable learning scheme, it is necessary to amend the loss function by adding constraint terms that moderate the variation of the c_s parameters.

5. Summary and outlook

We studied parameter optimization of domain-wall fermions to improve chiral symmetry using machine learning. By adopting a stochastic estimator of the global residual mass as the loss function, we demonstrated that the loss can be reduced through parameter updates.

In future work, we will validate whether the obtained parameters effectively reduce the residual mass by measuring the conventional plateau definition on larger lattice volumes. A systematic study of the lattice-volume dependence, as well as an investigation of the behavior toward the continuum limit by varying the lattice spacing, would also be important.

We have also observed that changing the parameters modifies the convergence behavior of the solver, suggesting that these parameters may be tuned as a precondition to improve the solver conditioning. In such a scenario, multiple evaluations such as chiral-symmetry violation and solver performance are required, and we expect the machine-learning framework to work effectively for this multiobjective optimization.

Acknowledgments

We would like to thank Y. Aoki, I. Kanamori, K. Nakayama, and K. Nitadori for their continued discussions. This work was supported by MEXT as Program for Promoting Researches on the Supercomputer Fugaku (Simulation for basic science: approaching the new quantum era; Grant Number JPMXP1020230411) and used computational resources of the Fugaku supercomputer provided by the RIKEN Center for Computational Science Project Numbers hp240295 and hp250224. We also used the Yukawa-21 supercomputer at the Yukawa Institute for Theoretical Physics, Kyoto University. S.Y. was supported by JST SPRING Grant Number JPMJSP2180. K.Y. was supported by JSPS KAKENHI Grant Number JP24KJ1110. A. T. was supported by JSPS KAKENHI Grant Numbers JP22H05111, JP22H05112, JP22K03539, JP25K07287, and by JST BOOST Grant Number JPMJBY24F1. Y. N. was supported by JSPS KAKENHI Grant Numbers JP22K03539, JP22K12052, JP22H05111, JP22H05114, and JP25K07287.

References

[1] H. B. Nielsen, M. Ninomiya, *A no-go theorem for regularizing chiral fermions*, Phys. Lett. B 105, 219 (1981).

[2] P. H. Ginsparg, K. G. Wilson, *A remnant of chiral symmetry on the lattice*, Phys. Rev. D 25, 2649 (1982).

[3] D. B. Kaplan, *A method for simulating chiral fermions on the lattice*, Phys. Lett. B 288, 342 (1992).

[4] Y. Shamir, *Chiral fermions from lattice boundaries*, Nucl. Phys. B 406, 90 (1993).

[5] V. Furman, Y. Shamir, *Axial symmetries in lattice QCD with Kaplan fermions*, Nucl. Phys. B 439, 54 (1995) .

[6] H. Neuberger, *Exactly massless quarks on the lattice*, Phys. Lett. B 417, 141 (1998).

[7] T.-W. Chiu, *Optimal Lattice Domain-Wall Fermions*, Phys. Rev. Lett. 90, 071601 (2003)

[8] Y.-C. Chen, T.-W. Chiu, *Chiral symmetry and the residual mass in lattice QCD with the optimal domain-wall fermion*, Phys. Rev. D 86, 094508 (2012).

[9] R. C. Brower, H. Neff, K. Orginos, *The Möbius domain wall fermion algorithm*, Comput. Phys. Commun. 220, 1 (2017).

PoS(LATTICE2025)035

[10] M. Abramczyk, S. Aoki, T. Blum, T. Izubuchi, H. Ohki, S. Syritsyn, *Lattice calculation of electric dipole moments and form factors of the nucleon*, Phys. Rev. D 96, 014501 (2017).

[11] A. Tomiya, *Machine Learning for Lattice QCD*, J. Phys. Soc. Jpn. 94, 031006 (2025).

[12] S. Lawrence, *Machine-learning approaches to accelerating lattice simulations*, arXiv:2502.02670 [hep-lat].

[13] S. Aoki, T. Izubuchi, Y. Kuramashi, Y. Taniguchi, *Domain wall fermions in quenched lattice QCD*, Phys. Rev. D 62, 094502 (2000).

[14] A. Ali Khan *et al.*, *Chiral properties of domain-wall quarks in quenched QCD*, Phys. Rev. D 63, 114504 (2001).

[15] T. Blum *et al*, *Quenched lattice QCD with domain wall fermions and the chiral limit*, Phys. Rev. D 69, 074502 (2004).

[16] G. Cossu, H. Fukaya, A.Tomiya, S. Hashimoto, *Violation of chirality of the Möbius domain-wall Dirac operator from the eigenmodes*, Phys. Rev. D 93, 034507 (2016).

[17] R. G. Edwards, U. M. Heller, *Domain wall fermions with exact chiral symmetry*, Phys. Rev. D 63, 094505 (2001).

[18] Y. Nagai, A. Tomiya, *JuliaQCD: Portable lattice QCD package in Julia language*, arXiv:2409.03030 [hep-lat], Code: `https://github.com/akio-tomiya/LatticeQCD.jl`.

[19] Y. Nagai, A. Tomiya, *Gauge covariant neural network for quarks and gluons*, Phys. Rev. D 111, 074501 (2025).