Extensively parallelizable chiral fermion

Yuki Nagai
CCSE, Japan Atomic Energy Agency, 178-4-4, Wakashiba, Kashiwa, 277-0871, Chiba, Japan and Mathematical Science Team, RIKEN Center for Advanced Intelligence Project (AIP), 1-4-1 Nihonbashi, Chuo-ku, 103-0027, Tokyo, Japan

Akio Tomiya
Faculty of Technology and Science, International Professional University of Technology, 3-3-1, Umeda, Kita-ku, Osaka, 530-0001, Osaka, Japan

(Dated: April 5, 2022)

Chiral symmetry is a key to investigating quantum physics, from condensed matter to particle physics. We propose a novel way of realizing a chiral fermion, known as the overlap-Dirac operator, without explicitly calculating the low modes of the Wilson-Dirac operator. We introduce a projection operator inspired by the Sakurai-Sugiura method and formulate the exact sign function and overlap-Dirac operator with a contour-integral form. Like the Sakurai-Sugiura method, the proposing method is multi-scale parallelizable, which fits the multi-core/multi-GPGPU paradigm. We confirm that the quality of chiral symmetry realized with the proposed method is sufficient for double precision. We evaluate the strong scaling of the proposing method.

INTRODUCTION

Chiral symmetry in physics plays a crucial role in a wide energy scale. For example, in the standard model of particle physics, chiral symmetry guarantees the lightness of light quarks and pions [1]. In the context beyond the standard model physics, chiral symmetry in supersymmetric Higgs fields protects it from quadratic divergence [2]. Moreover, it is essential to build the super string theory [3]. In fields of condensed matter physics, the chiral symmetry is important for realizing low-energy spectrum in Graphene or topologically nontrivial materials [4]. Nielsen and Ninomiya [5] have shown that chiral symmetry is tightly connected to the discretizations of spacetime and is non-trivial to realize on the lattice. Through the theorem, the structure of chiral symmetry on the lattice is related to the topological insulators [6].

Lattice gauge theory is a tool to investigate strongly coupled gauge systems like QCD [7]. Lattice QCD has been investigated for more than 40 years [8], and recently it is a mandatory technique not only for theoretical physics but also for experimental physics [9]. The formulation is based on discretized spacetime kept gauge symmetry and evaluated using supercomputers, which achieves great success [10–15]. Lattice QCD has demanded a formulation that keeps (precise) chiral symmetry, but that took a long time [16].

Neuberger introduced the overlap Dirac operator [17],

\[ D^{ov} (m_q) = \frac{1 + m_q}{2} 1 + \frac{1 - m_q}{2} \gamma_5 \text{sgn}(H_W). \]  

(1)

where \( H_W = \gamma_5 D_W (-m_0) \) is the hermitian Wilson-Dirac operator with negative mass \((-m_0 < 0)\) and \( m_q \) is a fermion mass. This Dirac-operator has exact chiral symmetry on the lattice [18] at \( m_q = 0 \). \( \text{sgn}(x) \) is the sign function takes 1 for \( x > 0 \), -1 for \( x < 0 \) and otherwise 0. However, the sign function causes difficulty with numerical computations. We must calculate lower-lying eigenvalues and eigenvectors for the Wilson-Dirac operator involved in the overlap-Dirac operator in practice [21–23]. In terms of the eigenmodes of \( H_W \), the sign function is represented as,

\[ \text{sgn}(H_W) = \sum_i \text{sgn}(\lambda_i) \vec{v}_i \vec{v}_i^\dagger, \]  

(2)

where \( \lambda_i \) is an eigenvalue of \( H_W \) and \( \vec{v}_i \) is the eigenvector for \( \lambda_i \). Using these arguments, we can evaluate the multiplication of the overlap-Dirac operator on a vector with the use of the Zorotarev approximation [21–23]. Even using many cores, it cannot be efficiently accelerated because calculations for each core are numerically expensive.

Although the overlap-Dirac operator has numerical difficulty, it has been used in a lot of studies with lattice QCD [24–29]. In addition to numerical calculations, mathematical physics e.g. proof of the index theorem [30]. Recent numerical studies with lattice QCD employ the domain-wall fermions [31–41], which can be regarded...
as an approximated overlap fermion \cite{12}; however, the approximation occasionally causes pathological behavior \cite{13}.

We propose a completely different and exact approach to evaluate the matrix-valued sign function. We point out that the sign function is expressed as

\[
\text{sgn}(H_W) = P_+(H_W) - P_-(H_W),
\]

where \(P_+(H_W) = \sum_{\lambda_i > 0} \tau_i \tilde{v}_i \tilde{v}_i^\dagger\) is a projection operator matrix which projects the matrix \(H_W\) to the sub-space spanned by the eigenvectors associated with positive (negative) eigenvalues. We emphasize that the zero eigenvalue \((\lambda_i = 0)\) do not contribute to the sign function because \(\text{sgn}(\lambda_i) = 0\) by definition of the sign function. The projection matrix can be expressed by the contour integral on complex plane \cite{44,45}:

\[
P_{\pm}(H_W) = \frac{1}{2\pi i} \int_{\Gamma_{\pm}} \frac{dz}{z - H_W},
\]

since the resolvent of \(H_W\) is \((zI - H_W)^{-1} = \sum_i \tilde{v}_i \tilde{v}_i^\dagger / (z - \lambda_i)\). Here, all positive (negative) eigenvalues are located inside the closed loop \(\Gamma_{+/}\) as shown in Fig. 2. Therefore, the multiplication of the sign function \(\text{sgn}(H_W)\) and a vector \(\vec{x}\) is evaluated by

\[
\text{sgn}(H_W)\vec{x} = \frac{1}{2\pi i} \int_{\Gamma_{+}} \frac{dz}{z - H_W} \vec{y}(z, H_W, \vec{x}) - \frac{1}{2\pi i} \int_{\Gamma_{-}} \frac{dz}{z - H_W} \vec{y}(z, H_W, \vec{x}),
\]

where \(\vec{y}(z, H_W, \vec{x})\) is a solution of the linear equations

\[
(zI - H_W)\vec{y} = \vec{x}.
\]

We remark that we do not have to solve the eigenvalue equation for low-laying modes in contrast to the conventional method \cite{22}.

We show a method to evaluate a contour integral with numerical quadrature. Using \(N_q\)-point quadrature rule, the vector \(P_{\Gamma}\vec{x}\) is approximately written as

\[
P_{\Gamma}\vec{x} \approx \frac{1}{N_q} \sum_{j=1}^{N_q} \rho w_j \vec{y}_j,
\]

with \(w_j = \alpha \cos \theta_j + i \sin \theta_j\), \(z_j = \gamma + \rho(\cos \theta_j + i \alpha \sin \theta_j)\), and \(\theta_j = 2\pi(j - 1/2)/N_q\). Here, \(\alpha\) is a vertical scaling factor. The vector \(\vec{y}_j = \vec{y}(z_j, H_W, \vec{x})\) is the solution of the linear equation (6). We note that, the approximately equal symbol in (7) can be exact in double precision if the \(N_q\) is enough large. This integration can be improve by employing more sophisticated integration scheme.

The linear equations with different \(z_j\) can be simultaneously evaluated by a shift solver \cite{46}. In addition, this expression can be extensively parallelized since the integration can be decomposed in small parts. We remark that the Sakurai-Sugiura method, which is the eigenvalue solver using the contour integrals, is applicable with multi-scale parallel computations. Our proposing method also has the same properties as the parallelization (Fig. 1). With the use of the shift solver, the calculation of the sign function \(\text{sgn}(H_W)\vec{x}\) consists of two parts: a calculation of a solution at a given seed point \(\vec{x}_\text{seed}\) by the biconjugate gradient method and constructions of solutions on contours \(\Gamma\). Therefore, the computational cost is a sum of the matrix-vector multiplication \(H_W\vec{v}\) and construction of solutions \(\vec{y}_j\). In this letter, we do not parallelize the former for simplicity.

In this letter, we examine the expression with the contour integrals (4).

**NUMERICAL EXPERIMENTS**

In this proof-of-principle study, we perform experiments for 30 quenched configurations generated with SU(3) Wilson gauge action with \(\beta = 5.7\) on \(L^4 \times 8^4\) lattice using LatticeQCD.jl \cite{17}, a package for simulation of the lattice QCD written in Julia programming language. The configurations are generated with the heatbath algorithm by skipping 10 trajectories. There are three overrelaxations per a heatbath update. We employ the standard overlap operator (Wilson kernel) and we choose \(m_0 = 1\). We confirm that our ensemble contains enough fluctuations of the topological charge as described in Supplemental material. Detailed extensive study will be provided in another publication.

It is well known that the massless overlap operator has exact chiral symmetry, namely, it satisfies the Ginsparg-Wilson relation \cite{18,20}.

\[
\{D^{\sigma\nu}(0), \gamma_5\} = D^{\sigma\nu}(0)\gamma_5 D^{\sigma\nu}(0),
\]

and it is equivalent to \((\text{sgn}(H_W))^2 = I\) \cite{48,49}. We evaluate quality of the sign function \cite{48},

\[
\sigma = \frac{\|\vec{s}^\dagger \vec{s} - \vec{y}^\dagger \vec{y}\|}{\|\vec{y}\|^2}
\]
where $\vec{s} = \text{sgn}(H_W)\vec{y}$ and $\vec{y}$ is a random vector. This $\sigma$ should be 0 if the Ginsparg-Wilson relation is exactly satisfied. Numerically, it is enough that this satisfies in the double precision.

We numerically perform the contour integral in (4). The number of divisions of the numerical integration for the contour integral is taken as $N_{qtot} = 100, 250, 500$ and 1000 to establish this formulation. We parallelize the code with MPI to evaluate the integral and compare the number of processes as 1 – 120. In contrast to the conventional way of parallelization for the lattice QCD, we can freely divide the system along with the contour integral independent to the lattice geometry. We develop a code based on LatticeDiracOperators.jl [50] written in Julia programming language and perform calculations in the supercomputer HPE SGI8600 in the Japan Atomic Energy Agency (Two Intel Xeon Gold 6242R CPUs (3.1GHz, 20core) per a computational node).

**Results**

Here we show results from numerical experiments. The statistical error is estimated using the Jackknife method.

First, we show the violation of the sign function for each configuration (Fig. 3). The violation $\sigma$ does not strongly depend on configurations and instead depends on the number of divisions. As discussed in the Supplemental material, the result is not correlated to the topology of configurations, which is clear from the definition of the matrix sign function.

Second, we show the result of the violation of the sign function along with the number of divisions (Fig. 4). The quality indicator $\sigma$ is averaged over configurations. The error is the same size as the symbol. One can see that the quality can be improved with the power of the divisions, and this divisions can be parallelly evaluated as in Fig. 1.

Fig. 3 shows the elapsed time of multiplication $\text{sgn}(H_W)$ to a vector as a functional of the number of the CPU cores. The vertical axis is the time that is averaged over configurations. Each CPU core solves linear equations with different $z_j$, simultaneously, with the use of the shifted Biconjugate Gradient (BICG) method. The total number of the integration points $N_q$ is divided by the number of the cores. Since we do not parallelize the matrix-vector multiplication, the elapsed time is governed by the single matrix-vector multiplication when the computational cost to evaluate shifted solutions $\vec{y}$ is reduced by many CPU cores (the Amdahl’s law). The residual elapsed time due to matrix-vector operation can be reduced by the low-level parallelization as shown in Fig. 1. The most expensive case in our parameter $N_{qtot} = 1000$ is shown in Tab. 1. It should be noted that the actual value of the elapsed time is not important in this table since we do not use fully optimized code and algorithm. For example, there might be better contour shape, position of the seed, and convergence criteria to reduce the computational cost.

In summary, $N_{qtot} \gtrsim 500$ achieves enough quality, $\sigma \lesssim 10^{-14}$ and the elapsed time is roughly 70 seconds with more than 50 processors.

**TABLE I.** Elapsed time for multiplying $\text{sgn}(H_W)$ with $N_{qtot} = 1000$

| # of procs | time (sec) | $\sigma$ |
|------------|------------|----------|
| 1          | 959(7)     | (2.97 ± 0.60) × 10^{-15} |
| 2          | 493(3)     | (3.0 ± 0.6) × 10^{-15}    |
| 16         | 142.8(3)   | (2.9 ± 0.6) × 10^{-15}    |
| 24         | 125.0(2)   | (3.0 ± 0.7) × 10^{-15}    |
| 40         | 94.8(3)    | (3.0 ± 0.7) × 10^{-15}    |
| 80         | 70.3(3)    | (2.95 ± 0.63) × 10^{-15}  |
| 120        | 62.3(3)    | (2.9 ± 0.6) × 10^{-15}    |
SUMMARY

In this work, we introduce a novel way to realize the overlap-Dirac operator without calculating eigenvalues and eigenvectors explicitly. We introduce projection operators represented by contour integrals involving the Wilson-Dirac operator $H_W$. Similar to the Sakurai-Sugiura method [44, 45], our method can be parallelized very efficiently.

The proposed method would pave the way for extensive studies of chiral physics from condensed matter to high-energy physics. In this work, we introduce a novel way to realize the Wilson-Dirac operator $H_W$. This work is partially supported by JSPS KAKENHI Grant Numbers 22K03539 (YN and AT).

The authors thank Akinori Tanaka for fruitful discussions. The part of this research was conducted with the supercomputer HPE SGI8600 in the Japan Atomic Energy Agency. Numerical computation in this work was supercomputer HPE SGI8600 in the Japan Atomic Energy Agency. Numerical computation in this work was

FIG. 5. Elapsed time of multiplication $\text{sgn}(H_W)$ to a vector. The time is averaged over configurations. The statistical error is evaluated but the same size with the symbol.
Ichi Noaki. Finite temperature study of the axial U(1) symmetry on the lattice with overlap fermion formulation. *Phys. Rev. D*, 87(11):114514, 2013. [Erratum: Phys.Rev.D 88, 019901 (2013)].

[29] M. Gong et al. Strangeness and charmness content of the nucleon from overlap fermions on 2+1-flavor domain-wall fermion configurations. *Phys. Rev. D*, 88:014503, 2013.

[30] Kazuo Fujikawa. Relation $\text{Tr} \gamma(5) = 0$ and the index theorem in lattice gauge theory. *Phys. Rev. D*, 60:074505, 1999.

[31] Peter Boyle. 2+1 flavour Domain Wall Fermion simulations by the RBC and UKQCD collaborations. *PoS*, LATTICE2007:005, 2007.

[32] Ting Wai Chiu, Tung Han Hsieh, and Yao Yuan Mao. Topological Susceptibility in Two Flavors Lattice QCD with the Optimal Domain-Wall Fermion. *Phys. Rev. D*, 88:014503, 2013.

[33] Kazuo Fujikawa. Relation $\text{Tr} \gamma(5) = 0$ and the index theorem in lattice gauge theory. *Phys. Rev. D*, 60:074505, 1999.

[34] Ting-Wai Chiu, Wen-Ping Chen, Yu-Chih Chen, Han-Yi Chou, Tian-Shin Guu, and Tung-Han Hsieh. Chiral symmetry and axial U(1) symmetry in finite temperature QCD with domain-wall fermion. *PoS*, LATTICE2013:165, 2014.

[35] Wen-Ping Chen, Yu-Chih Chen, Ting-Wai Chiu, Han-Yi Chou, Tian-Shin Guu, and Tung-Han Hsieh. Decay Constants of Pseudoscalar D-mesons in Lattice QCD with Domain-Wall Fermion. *Phys. Rev. D*, 93(9):094503, 2016.

[36] Katsumasa Nakayama, Brendan Fahy, and Shoji Hashimoto. Short-distance charmonium correlator on the lattice with Möbius domain-wall fermion and a determination of charm quark mass. *Phys. Rev. D*, 94(5):054507, 2016.

[37] A. Tomiya, G. Cossu, S. Aoki, H. Fukaya, S. Hashimoto, T. Kaneko, and J. Noaki. Evidence of effective axial U(1) symmetry restoration at high temperature QCD. *Phys. Rev. D*, 96(3):034509, 2017. [Addendum: Phys.Rev.D 96, 079902 (2017)].

[38] Florian Burger, Ernst-Michael Ilgenfritz, Maria Paola Lombardo, and Anton Trunin. Chiral observables and topology in hot QCD with two families of quarks. *Phys. Rev. D*, 98(9):094501, 2018.

[39] G. Colangelo et al. Prospects for precise predictions of $a_\mu$ in the Standard Model. In 2022 Snowmass Summer Study, 3 2022.

[40] Brian Colquhoun, Shoji Hashimoto, Takashi Kaneko, and Jonna Koponen. Form factors of $B \to \pi \ell \nu$ and a determination of $|V_{ub}|$ with Möbius domain-wall-fermions. 3 2022.

[41] S. Yamamoto, T. Sogabe, T. Hoshi, S.-L. Zhang, and T. Fujiwara. Shifted conjugate-orthogonal–conjugate-gradient method and its application to double orbital extended hubbard model. *Journal of the Physical Society of Japan*, 77(11):114713, 2008.
Supplemental Materials: Extensively parallelizable chiral fermion

NUMERICAL SETUP FOR THE INTEGRATION

In the contour integral with numerical quadrature, we set a vertical scaling factor $\alpha = 0.1$, a radius $\rho = 1.1999$ and the origin of the ellipse $\gamma = \pm 1.2$. We note that we can tune parameters if all eigenvalues are located inside the contours. We set the seed point, $z_{\text{seed}} = 1 + 0.05i$, where the linear equation should be solved. For simplicity, we do not use a seed switching technique [S51].

TOPOLOGICAL SECTOR

The overlap Dirac operator is tightly related to the topological charge through the index theorem. In particular, a configuration $Q \neq 0$ accompanies zero-modes of $D^{ov}(0)$.

First, we examine the topological fluctuation of our ensemble. We employ $O(a^2)$ improved definition of the topological charge and measure it after the Wilson flow $t = 4.0$ to make them approximately integers. Fig. S1 is the history of the topological charge and it shows that our ensemble contains $Q \neq 0$ configurations.

![Fig. S1. History of the topological charge.](image)

Next, we see the correlation between the topological charge and the violation of sign function $\text{sgn}(H_W)$. As we have mentioned, $\text{sgn}(x) = 0$ for $x = 0$ in the introduction; however, some readers might suspect topological zero-modes could affect the quality, so we check the correlation between them. Fig. S2 - S5 show the correlation for $N_{\text{qtot}}$. One can see that, no clear correlation between the topological sector and the violation $\sigma$. 
FIG. S2. Correlation between the violation of sign function $\sigma$ and topological charge for $N_{\text{tot}} = 100$.

FIG. S3. Same figure to Fig. S2 but for $N_{\text{tot}} = 250$.

FIG. S4. Same figure to Fig. S2 but for $N_{\text{tot}} = 500$. One data point reaches $\sigma = 0$ numerically and it is dropped from the figure.
FIG. S5. Same figure to Fig. S2 but for $N_{\text{tot}} = 1000$. 