

# Chiral symmetry and taste symmetry from the eigenvalue spectrum of staggered Dirac operators

Hwancheol Jeong<sup>1</sup>, Chulwoo Jung<sup>2</sup>, Seungyeob Jwa<sup>1</sup>, Jangho Kim<sup>3</sup>, Jeehun Kim<sup>1</sup>, Nam Soo Kim<sup>4</sup>,  
Sunghee Kim<sup>1</sup>, Sunkyu Lee<sup>1</sup>, Weonjong Lee<sup>1,\*</sup>, Youngjo Lee<sup>5</sup>, and Jeonghwan Pak<sup>1</sup>

(SWME Collaboration)

<sup>1</sup>*Lattice Gauge Theory Research Center, FPRD, and CTP, Department of Physics and Astronomy, Seoul National University, Seoul 08826, South Korea*

<sup>2</sup>*Physics Department, Brookhaven National Laboratory, Upton, New York 11973, USA*

<sup>3</sup>*Institut für Theoretische Physik, Goethe University Frankfurt am Main, Max-von-Laue-Str. 1, 60438 Frankfurt am Main, Germany*

<sup>4</sup>*Department of Electrical and Computer Engineering and the Institute of New Media and Communications Seoul National University, Seoul 08826, South Korea*

<sup>5</sup>*Department of Statistics, Seoul National University, Seoul 08826, South Korea*



(Received 1 September 2020; accepted 29 June 2021; published 26 July 2021)

We investigate general properties of the eigenvalue spectrum for improved staggered quarks. We introduce a new chirality operator  $[\gamma_5 \otimes 1]$  and a new shift operator  $[1 \otimes \xi_5]$ , which respect the same recursion relation as the  $\gamma_5$  operator in the continuum. Then we show that matrix elements of the chirality operator sandwiched between two eigenstates of the staggered Dirac operator are related to those of the shift operator by the Ward identity of the conserved  $U(1)_A$  symmetry of staggered fermion actions. We perform a numerical study in quenched QCD using HYP staggered quarks to demonstrate the Ward identity. We introduce a new concept of leakage patterns which collectively represent the matrix elements of the chirality operator and the shift operator sandwiched between two eigenstates of the staggered Dirac operator. The leakage pattern provides a new method to identify zero modes and nonzero modes in the Dirac eigenvalue spectrum. This method is as robust as the spectral flow method but requires much less computing power. Analysis using a machine learning technique confirms that the leakage pattern is universal, since the staggered Dirac eigenmodes on normal gauge configurations respect it. In addition, the leakage pattern can be used to determine a ratio of renormalization factors as a by-product. We conclude that it might be possible and realistic to measure the topological charge  $Q$  using the Atiya-Singer index theorem and the leakage pattern of the chirality operator in the staggered fermion formalism.

DOI: [10.1103/PhysRevD.104.014508](https://doi.org/10.1103/PhysRevD.104.014508)

## I. INTRODUCTION

It is important to understand the low-lying eigenvalue spectrum of the Dirac operator, which exhibits the topological Ward identity of the Atiya-Singer index theorem [1], the Banks-Casher relationship [2], and the universality of the distribution of the near-zero modes for fixed topological charge sectors [3,4]. Study on the eigenvalue spectrum of the Dirac operator is, by nature, highly non-perturbative. Hence, numerical tools available in lattice

gauge theory provide a perfect playground to study diverse properties of the Dirac eigenvalue spectrum.

In lattice QCD, there are a number of popular methods to implement a discrete version of the continuum Dirac operator. We are interested in one particular class of lattice fermions that are widely used in the lattice QCD community: improved staggered quarks [5–7]. Here we study the eigenvalue spectrum of staggered Dirac operators in quenched QCD to show that the small eigenvalues near zero modes of the staggered Dirac operators reproduce the continuum properties very closely, which was originally noticed in Refs. [8–10]. To reach this conclusion, the authors of Refs. [8,9] performed a number of tests, verifying consistency of lattice data with (1) the Atiya-Singer index theorem that describes the chiral Ward identity relating the zero modes to the topological charge; (2) the Banks-Casher relationship that relates the chiral condensate

\*wlee@snu.ac.kr

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP<sup>3</sup>.

to the density of eigenvalues at the zero modes; and (3) the universality of the small eigenvalue spectrum in the  $\varepsilon$ -regime predicted by random matrix theory. In addition, the authors of Refs. [11,12] used the spectral flow method of Adams [13] to identify the zero modes from a mixture with nonzero modes. The spectral flow method is robust but highly expensive in a computational sense.

Here, we introduce a new, advanced chirality operator  $[\gamma_5 \otimes 1]$ , which respects the continuum algebra of  $\gamma_5$ . We show that matrix elements of this chirality operator between eigenstates are related to those of the shift operator  $[1 \otimes \xi_5]$  through the Ward identity of the conserved  $U(1)_A$  symmetry of staggered fermions. In addition, we introduce a new concept of leakage patterns to distinguish zero modes from nonzero modes. Using the leakage pattern of the chirality and shift operators, we show that one can measure the zero modes as reliably as when using the spectral flow method. Hence, one could determine the topological charge  $Q$  using the leakage pattern with much smaller computational cost than by using the spectral flow. We also show that it is possible to determine the ratio of renormalization constants  $Z_{P \times S}/Z_{P \times P}$  using the leakage pattern.

In Sec. II, we briefly review the continuum theory of the eigenvalue spectrum and its relation to the quark condensate  $\langle \bar{\psi}\psi \rangle$ . We also review the Atiya-Singer index theorem in brief. In Sec. III, we briefly review the eigenvalue spectrum of staggered Dirac operators that is obtained using the Lanczos algorithm. In Sec. IV, we briefly review the conserved  $U(1)_A$  symmetry in the staggered fermion formalism and explain its role in the eigenvalue spectrum of staggered Dirac operators. We also present numerical examples to help readers to understand basic concepts and notation. In Sec. V, we define the chirality operator  $[\gamma_5 \otimes 1]$  and the shift operator  $[1 \otimes \xi_5]$ . We show that they respect the continuum recursion relation of  $\gamma_5$ . Then we derive the chiral Ward identity of the  $U(1)_A$  symmetry to show that the matrix elements of the chirality operator are related to those of the shift operator through the Ward identity. We discuss the eigenvalue spectrum in the continuum limit and introduce a new notation of quartet indices. Then we introduce the concept of leakage patterns for the chirality operator and the shift operator. We also present numerical examples to demonstrate that the leakage patterns of zero modes are completely different from those of nonzero modes. In Sec. VI, we review a machine learning technique and describe how to apply it to extract efficiently the quartet structure of nonzero modes using leakage patterns. In Sec. VII, we explain how the leakage pattern of the zero modes can be used to determine the ratio of the renormalization factors nonperturbatively. In Sec. VIII, we conclude. The appendixes contain technical details on Lanczos algorithms and mathematical proofs, and more plots of leakage patterns for diverse topological charge values.

Preliminary results of this paper are published in Refs. [14–16].

## II. QUARK CONDENSATE IN THE CONTINUUM

In the continuum the quark condensate is given by

$$\langle \bar{\psi}\psi \rangle = \frac{1}{N_f} \sum_f \langle 0 | \bar{\psi}_f \psi_f | 0 \rangle \quad (1)$$

$$= -\frac{1}{VN_f} \int d^4x \text{Tr} \left( \frac{1}{D+m} \right), \quad (2)$$

where  $D$  is the Dirac operator,  $m$  is the quark mass,  $x$  is the space-time coordinate,  $V$  is the volume, and  $N_f$  is the number of flavors with the same mass  $m$ . The trace is a sum over spin and color. Let us think of the eigenvalues of the Dirac operator.  $D$  is anti-Hermitian, so its eigenvalues are purely imaginary or zero.

$$D^\dagger = -D \quad (3)$$

$$D u_\lambda(x) = i\lambda u_\lambda(x) \quad (4)$$

where  $\lambda$  is a real eigenvalue, and  $u_\lambda(x)$  is the corresponding eigenvector.

By spectral decomposition [4],

$$S_f(x, y) = \langle \psi_f(x) \bar{\psi}_f(y) \rangle = \sum_\lambda \frac{1}{i\lambda + m} u_\lambda(x) u_\lambda^\dagger(y) \quad (5)$$

$$\langle \bar{\psi}\psi \rangle = -\frac{1}{V} \sum_\lambda \frac{1}{i\lambda + m} \int d^4x \text{Tr}(u_\lambda(x) u_\lambda^\dagger(x)) \quad (6)$$

$$= -\frac{1}{V} \sum_\lambda \frac{1}{i\lambda + m} \quad (7)$$

where we adopt the normalization convention

$$\langle u_a | u_b \rangle = \int d^4x u_a^\dagger(x) u_b(x) = \delta_{ab}. \quad (8)$$

Thanks to the chiral symmetry,

$$\gamma_5 D = -D \gamma_5 \quad (9)$$

$$D \gamma_5 | u_\lambda \rangle = -i\lambda \gamma_5 | u_\lambda \rangle. \quad (10)$$

Let us define  $u_{-\lambda} \equiv \gamma_5 u_\lambda$ , so that  $D u_{-\lambda} = -i\lambda u_{-\lambda}$ . Hence, if there exists  $u_\lambda$  for  $\lambda \neq 0$ , then the parity partner eigenstate  $u_{-\lambda}$  with negative eigenvalue  $-i\lambda$  must also exist.

Now let us separate the zero mode contribution from the spectral decomposition.

$$\langle \bar{\psi}\psi \rangle = -\frac{1}{V} \sum_{\lambda>0} \frac{2m}{\lambda^2 + m^2} - \frac{n_+ + n_-}{mV}. \quad (11)$$

Here,  $n_+$  ( $n_-$ ) is the number of right-handed (left-handed) zero modes per flavor. Let us define the subtracted quark condensate  $\langle \bar{\psi}\psi \rangle_{\text{sub}}$ :

$$\langle \bar{\psi}\psi \rangle_{\text{sub}} = \langle \bar{\psi}\psi \rangle + \frac{n_+ + n_-}{mV} = -\frac{1}{V} \sum_{\lambda>0} \frac{2m}{\lambda^2 + m^2}. \quad (12)$$

$$= -\frac{1}{V} \sum_n \frac{2m}{\lambda_n^2 + m^2} \quad \text{with } \lambda_n > 0 \quad (13)$$

$$= -\int_{-\infty}^{+\infty} d\lambda \frac{m}{\lambda^2 + m^2} \rho_s(\lambda), \quad (14)$$

where the spectral density  $\rho_s(\lambda)$  is

$$\rho_s(\lambda) = \frac{1}{V} \sum_n \delta(\lambda - \lambda_n). \quad (15)$$

Here,  $\rho_s$  is a spectral density on a single gauge configuration with volume  $V$ . Now let us average over a full ensemble of gauge field configurations and take the limit of infinite volume ( $V \rightarrow \infty$ ). Then, in that limit, the spectral density  $\rho(\lambda) = \langle \rho_s(\lambda) \rangle$  has a well defined (smooth and continuous) value as  $\lambda \rightarrow 0$ . We can define the chiral condensate as

$$\begin{aligned} \Sigma &= -\langle 0 | \bar{\psi}\psi | 0 \rangle_{\text{sub}}(m=0) \\ &= \lim_{m \rightarrow 0} \int_{-\infty}^{+\infty} d\lambda \frac{m}{\lambda^2 + m^2} \rho(\lambda) = \pi\rho(0), \end{aligned} \quad (16)$$

which is the Banks-Casher relation. The subtracted quark condensate  $\langle \bar{\psi}\psi \rangle_{\text{sub}}$  is expected to behave well in the chiral limit, even though the contribution from the zero modes is divergent as a simple pole in the chiral limit. Hence, in the numerical study on the lattice, it is important to identify the would-be zero modes which correspond to the zero modes in the continuum limit, and to remove them in the calculation of the quark condensate.

Before proceeding, let us briefly go through the index theorem. In the continuum theory in Euclidean space, the axial Ward identity [17] is

$$\partial_\mu A_\mu(x) = 2mP(x) - 2N_f q(x). \quad (17)$$

Here  $A_\mu \equiv \bar{\psi}\gamma_\mu\gamma_5\psi$  is the axial vector current in the flavor singlet representation,  $P \equiv \bar{\psi}\gamma_5\psi$  is the corresponding pseudo-scalar operator, and  $q \equiv \frac{1}{32\pi^2} \text{Tr}[F_{\mu\nu}\tilde{F}_{\mu\nu}]$  is the topological charge density (=winding number density). Now the topological charge  $Q$  is

$$Q \equiv \int d^4x \langle q(x) \rangle \quad (18)$$

$$= -\frac{1}{2N_f} \int d^4x \langle \partial_\mu A_\mu(x) - 2mP(x) \rangle \quad (19)$$

$$= \frac{m}{N_f} \int d^4x \langle \bar{\psi}\gamma_5\psi \rangle. \quad (20)$$

Using the spectral decomposition, we can rewrite  $Q$  as follows,

$$Q = -m \sum_\lambda \frac{1}{i\lambda + m} \int d^4x [u_\lambda^\dagger(x) \gamma_5 u_\lambda(x)]. \quad (21)$$

Noting that  $\gamma_5 u_\lambda(x) = u_{-\lambda}(x)$  for  $\lambda \neq 0$ ,

$$\int d^4x [u_\lambda^\dagger(x) \gamma_5 u_\lambda(x)] = \langle u_\lambda | u_{-\lambda} \rangle = 0. \quad (22)$$

Hence, only zero modes contribute to  $Q$ . For the zero modes, it is convenient to choose the helicity eigenstates as the basis vectors so that  $\langle u_0^L | \gamma_5 | u_0^L \rangle = -1$  and  $\langle u_0^R | \gamma_5 | u_0^R \rangle = +1$ , where the superscripts  $L, R$  represent left-handed and right-handed helicity, respectively. Then deriving the index theorem is straightforward [1]:

$$Q = n_- - n_+, \quad (23)$$

where  $n_+$  ( $n_-$ ) is the number of the right-handed (left-handed) zero modes.

### III. SPECTRAL DECOMPOSITION WITH STAGGERED FERMIONS

A number of improved versions of staggered fermions exist, such as HYP-smearred staggered fermions [5], asqtad improved staggered fermions [18], and highly improved staggered quarks (HISQ) [7]. Here we refer to all of them collectively as ‘‘staggered fermions.’’ Staggered fermions have four tastes per flavor by construction [19]. Hence, the quark condensate for staggered fermions is defined as

$$\langle \bar{\chi}\chi \rangle = -\frac{1}{VN_t} \left\langle \text{Tr} \left( \frac{1}{D_s + m} \right) \right\rangle_U, \quad (24)$$

where  $\chi$  represents a staggered quark field,  $D_s$  is the staggered Dirac operator for a single valence flavor,  $V$  is the lattice volume, and  $N_t$  is the number of tastes. We measure the quark condensate using a stochastic method.

$$(D_s + m)_{x,y} \chi(y) = \xi(x) \quad (25)$$

$$\chi(x) = \left[ \frac{1}{D_s + m} \right]_{x,y} \xi(y) \quad (26)$$

$$\text{Tr}\left(\frac{1}{D_s + m}\right) = \lim_{N_\xi \rightarrow \infty} \frac{1}{N_\xi} \sum_\xi \sum_y \xi^\dagger(y) \chi(y), \quad (27)$$

where  $x, y$  are representative indices which represent the space-time coordinates, taste, and color indices collectively. Here  $\xi(x)$  represents either Gaussian random numbers or  $U(1)$  noise random numbers which satisfy a simple identity:

$$\lim_{N_\xi \rightarrow \infty} \frac{1}{N_\xi} \sum_\xi \xi^\dagger(x) \xi(y) = \delta_{xy},$$

where  $N_\xi$  is the number of random vector samples.

Staggered fermions have a taste symmetry  $SU(4)_L \otimes SU(4)_R \otimes U(1)_V$  in the continuum limit at  $a = 0$  [20]. However, this symmetry breaks down to a subgroup  $U(1)_V \otimes U(1)_A$  on the lattice with  $a \neq 0$  [19,20]. The remaining axial symmetry  $U(1)_A$  plays an important role in protecting the quark mass from receiving an additive renormalization. In addition, it does not have any axial anomaly.

The Dirac operator ( $D_s$ ) of staggered fermions is anti-Hermitian:  $D_s^\dagger = -D_s$ . Hence, its eigenvalues are purely imaginary:

$$D_s |f_\lambda^s\rangle = i\lambda |f_\lambda^s\rangle, \quad (28)$$

where  $\lambda$  is real. Here, the subscript  $s$  and superscript  $s$  represent staggered quarks.

In practice, when we obtain eigenvalues of  $D_s$  numerically, we use the following relationship instead of Eq. (28):

$$D_s^\dagger D_s |g_{\lambda^2}^s\rangle = \lambda^2 |g_{\lambda^2}^s\rangle \quad (29)$$

where the  $|g_{\lambda^2}^s\rangle$  state is a mixture of the two eigenvectors  $|f_{+\lambda}^s\rangle$  and  $|f_{-\lambda}^s\rangle$ . In other words,

$$|g_{\lambda^2}^s\rangle = c_1 |f_{+\lambda}^s\rangle + c_2 |f_{-\lambda}^s\rangle \quad (30)$$

where the  $c_i$  are complex numbers that satisfy the normalization condition

$$|c_1|^2 + |c_2|^2 = 1. \quad (31)$$

The numerical algorithm is a variation of a Lanczos algorithm adapted for lattice QCD [21]. Details on the numerical algorithms as well as comprehensive references are given in Appendix A.

Why do we obtain  $\lambda^2$  instead of  $i\lambda$ ? The first reason is that doing so allows us to use even-odd preconditioning [22], which makes Lanczos run on only even or odd sites on the lattice. This leads to two benefits: One is that there is a substantial gain in the speed of the code, and the other is that the code uses only half the memory otherwise required.

Details on even-odd preconditioning are described in Appendix B. The second reason is that obtaining  $\lambda^2$  instead of  $i\lambda$  allows us to implement polynomial acceleration algorithms [23] into Lanczos more easily, since the eigenvalues of  $D_s^\dagger D_s$  are positive definite and have a lower bound  $\lambda^2 > 0$ . Note that staggered fermions have would-be zero modes whose eigenvalues are small and positive ( $\lambda^2 > 0$ ) on rough gauge configurations. There are no exact zero modes ( $\lambda = 0$ ) with staggered fermions on rough gauge configurations [24]. Details of our implementation of polynomial acceleration are described in Appendix A.

Hence, we use the Lanczos algorithm to solve Eq. (29) for the eigenvector  $|g_{\lambda^2}^s\rangle$  as well as the corresponding eigenvalue  $\lambda^2$ . We obtain  $|f_{+\lambda}^s\rangle$  and  $|f_{-\lambda}^s\rangle$  by using projection operators defined as

$$P_+ = (D_s + i\lambda) \quad (32)$$

$$P_- = (D_s - i\lambda), \quad (33)$$

where  $P_+$  is the projection operator that selects only the  $|f_{+\lambda}^s\rangle$  component and removes the  $|f_{-\lambda}^s\rangle$  component. Then

$$|\chi_+\rangle = P_+ |g_{\lambda^2}^s\rangle \quad (34)$$

$$|\chi_-\rangle = P_- |g_{\lambda^2}^s\rangle \quad (35)$$

and the orthonormal eigenvectors are

$$|f_{+\lambda}^s\rangle = \frac{|\chi_+\rangle}{\sqrt{\langle \chi_+ | \chi_+ \rangle}} \quad (36)$$

$$|f_{-\lambda}^s\rangle = \frac{|\chi_-\rangle}{\sqrt{\langle \chi_- | \chi_- \rangle}}. \quad (37)$$

#### IV. CHIRAL SYMMETRY OF STAGGERED FERMIONS

The two vectors  $|f_{\pm\lambda}^s\rangle$  are related to each other through a chiral Ward identity of staggered fermions. Here we address this issue of the chiral symmetry of staggered fermions and its consequences.

##### A. Notation and definitions

Let us begin with notation and definitions. For staggered fermions, there are two independent methods to transcribe operators to the lattice: One is the Golterman method [19,25,26], and the other is the Kluberg-Stern method [27–30]. In Appendix D, we explain how to construct chirality operators using both the Golterman method and the Kluberg-Stern method, and we compare the two methods. The comparison is summarized in Table XI of Appendix D. Since the Kluberg-Stern method respects the recursion relationship, uniqueness of chirality, and the

Ward identity while the Golterman method does not, we adopt the former method to construct bilinear operators. Accordingly we define staggered bilinear operators as

$$\begin{aligned} \mathcal{O}_{S \times T}(x) &\equiv \sum_{A,B} \bar{\chi}(x_A) [\gamma_S \otimes \xi_T]_{AB} \chi(x_B) \\ &= \sum_{A,B} \bar{\chi}_a(x_A) \overline{(\gamma_S \otimes \xi_T)_{AB}} U(x_A, x_B)_{ab} \chi_b(x_B) \end{aligned} \quad (38)$$

where  $\chi_b$  are staggered quark fields, and  $a, b$  are color indices. Here the coordinate  $x_A = 2x + A$ ,  $x$  is a coordinate of the hypercube, and  $A, B$  are hypercubic vectors with  $A_\mu, B_\mu \in \{0, 1\}$ . The spin-taste matrices are

$$\overline{(\gamma_S \otimes \xi_T)_{AB}} \equiv \frac{1}{4} \text{Tr}(\gamma_A^\dagger \gamma_S \gamma_B \gamma_T^\dagger) \quad (39)$$

where  $\gamma_S$  represents the Dirac spin matrix, and  $\xi_T$  represents the  $4 \times 4$  taste matrix. In addition,

$$\begin{aligned} U(x_A, x_B) &\equiv \mathbb{P}_{SU(3)} \left[ \sum_{p \in \mathcal{C}} V(x_A, x_{p_1}) V(x_{p_1}, x_{p_2}) \right. \\ &\quad \left. \cdots V(x_{p_n}, x_B) \right] \end{aligned} \quad (40)$$

where  $\mathbb{P}_{SU(3)}$  represents the  $SU(3)$  projection, and  $\mathcal{C}$  represents the complete set of shortest paths from  $x_A$  to  $x_B$ .  $V(x, y)$  represents the HYP-smear fat link [5,6] for HYP staggered fermions, the Fat7 fat link [6,31–33] for asqtad staggered fermions or HISQ, and the thin gauge link for unimproved staggered fermions.

The conserved  $U(1)_A$  axial symmetry transformation is

$$\begin{aligned} \Gamma_\epsilon(A, B, a, b) &\equiv [\gamma_S \otimes \xi_S]_{AB;ab} \\ &= \overline{(\gamma_S \otimes \xi_S)_{AB}} \cdot \delta_{ab} \\ &= \epsilon(A) \cdot \delta_{AB} \cdot \delta_{ab} \end{aligned} \quad (41)$$

where  $\Gamma_\epsilon$  is often called ‘‘distance parity,’’ and

$$\epsilon(A) \equiv (-1)^{S_A} \quad (42)$$

$$S_A \equiv \sum_{\mu=1}^4 A_\mu. \quad (43)$$

Under the  $U(1)_A$  transformation, the staggered Dirac operator transforms as follows,

$$\Gamma_\epsilon D_s \Gamma_\epsilon = D_s^\dagger = -D_s \quad (44)$$

$$\Gamma_\epsilon D_s = -D_s \Gamma_\epsilon. \quad (45)$$

Therefore,

$$\begin{aligned} D_s |f_{+\lambda}^s\rangle &= +i\lambda |f_{+\lambda}^s\rangle \\ D_s \Gamma_\epsilon |f_{+\lambda}^s\rangle &= -i\lambda \Gamma_\epsilon |f_{+\lambda}^s\rangle, \end{aligned} \quad (46)$$

and  $f_{-\lambda}^s$  can be obtained from  $f_{+\lambda}^s$  through  $\Gamma_\epsilon$  transformation as follows,

$$\begin{aligned} \Gamma_\epsilon |f_{+\lambda}^s\rangle &= e^{+i\theta} |f_{-\lambda}^s\rangle \\ \Gamma_\epsilon |f_{-\lambda}^s\rangle &= e^{-i\theta} |f_{+\lambda}^s\rangle. \end{aligned} \quad (47)$$

In general, there is no constraint for the real phase  $\theta$ , so we expect its probability distribution to be random. In practice, however, we make use of even-odd preconditioning, and we obtain the odd site fermion fields ( $|g_o\rangle$ ) from the even site fermion fields ( $|g_e\rangle$ ) with the relation  $|g_o\rangle = \eta D_{oe} |g_e\rangle$ , where  $D_{oe}$  is a sector of  $D_s$  that connects even site fields to odd site fields, and  $\eta$  is a random complex number. Hence the distribution of  $\theta$  depends on our choice of  $\eta$ . In our numerical study, we set  $\eta = 1$ . Then  $\theta$  is given by

$$\theta = \pi + 2\beta, \quad \beta = \arctan(\lambda). \quad (48)$$

Details on the even-odd preconditioning and the derivation of Eq. (48) are explained in Appendix B.

We expect that if there exists an eigenvector  $|f_{+\lambda}^s\rangle$ , there must be a corresponding parity partner  $|f_{-\lambda}^s\rangle$  due to the exact chiral symmetry  $\Gamma_\epsilon$ . In other words, the Ward identity of Eq. (47) comes directly from the conserved  $U(1)_A$  axial symmetry.

## B. Numerical examples

We now use numerical examples to demonstrate how the above theory works in quenched QCD. In Table I, details of the gauge configurations are presented.

We measure the topological charge  $Q$  using gauge links. We use the  $Q(5\text{Li})$  operator defined in Refs. [40,41] after 10 ~ 30 iterations of APE smearing with  $\alpha = 0.45$  [42–44]. We show an example of the eigenvalue spectrum for  $Q = 0$  in Fig. 1. Since  $Q = 0$ , we do not expect to find any zero

TABLE I. Input parameters for numerical study in quenched QCD. For more details, refer to Ref. [9]. The relationship between sample sizes in our study and the number of the gauge configurations is nontrivial and discussed later.

Parameters	Values
Glue action	Tree level Symanzik [34–36]
Tadpole improvement	Yes
$\beta$	5.0
Geometry	20 <sup>4</sup>
$a$	0.077(1) fm [37]
1/ $a$	2.6 GeV
Valence quarks	HYP staggered fermions [6,38,39]
$N_f$	$N_f = 0$ (quenched QCD)

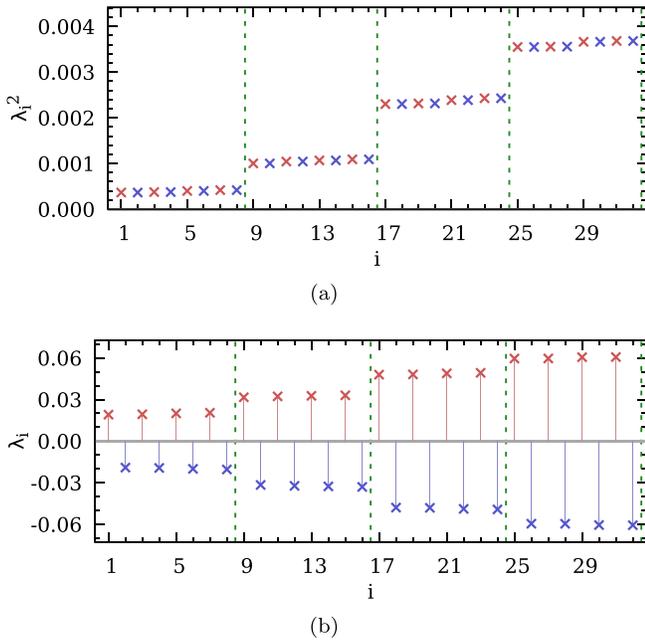


FIG. 1. Eigenvalue spectrum of staggered Dirac operator on a  $Q = 0$  gauge configuration. (a)  $\lambda_i^2$  and (b)  $\lambda_i$ .

modes for this gauge configuration. In Fig. 1(a), we show eigenvalues  $\lambda^2$  for the eigenvectors  $|g_{\lambda^2}^s\rangle$  defined in Eq. (29). Here we observe eight-fold degeneracy for nonzero eigenmodes due to the conserved  $U(1)_A$  axial symmetry. Here  $\lambda_2 = -\lambda_1$  and, in general,  $\lambda_{2n} = -\lambda_{2n-1}$  for integer  $n > 0$ . In other words,  $\lambda_{2n}$  is the parity partner of  $\lambda_{2n-1}$ . For each  $\lambda_i$ , there exists four-fold degeneracy due to approximate  $SU(4)$  taste symmetry. For each of these four-fold degenerate eigenvalues [for example  $\lambda_1, \lambda_3, \lambda_5, \lambda_7$  in Fig. 1(a)], there exists a parity partner eigenvalue due to the  $U(1)_A$  symmetry:  $\lambda_2 = -\lambda_1, \lambda_4 = -\lambda_3, \lambda_6 = -\lambda_5$ , and  $\lambda_8 = -\lambda_7$  [refer to Fig. 1(b)].

Let us turn to an example with  $Q = -1$ . Since  $Q = -1$ , we expect to observe four-fold would-be zero modes. The gauge configurations are so rough that we expect to observe not exact zero modes but would-be zero modes. In Fig. 2, we demonstrate how the would-be zero modes behave on a gauge configuration with  $Q = -1$ . As one can see in Figs. 2(a) and 2(b), we find four-fold degenerate would-be zero modes:  $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ . Thanks to the  $U(1)_A$  chiral Ward identity in Eq. (47), we find that  $\lambda_2 = -\lambda_1$  and  $\lambda_4 = -\lambda_3$ . As in the case with  $Q = 0$ , we find that the nonzero eigenmodes are eight-fold degenerate. This pattern of four-fold degeneracy for would-be zero modes and eight-fold degeneracy for nonzero modes is also observed in the cases with  $Q = -2$  and  $Q = -3$ , which are presented in Appendix C.

At this point, the reader may have already concluded that we can distinguish would-be zero modes of staggered quarks from nonzero modes by counting the degeneracy of the eigenvalues [8,9,45]. This conclusion is true but can

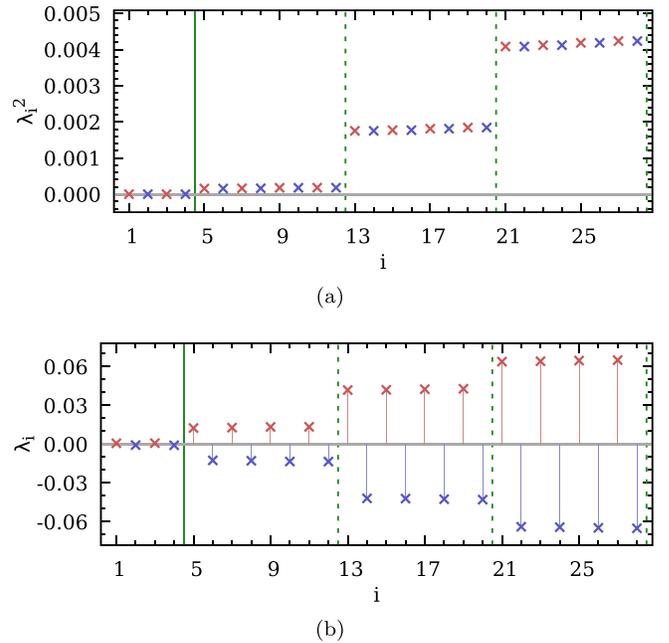


FIG. 2. Eigenvalue spectrum of staggered Dirac operator on a  $Q = -1$  gauge configuration. (a)  $\lambda_i^2$  and (b)  $\lambda_i$ .

lead to wrong answers in practice. The reason is that, on large lattices, the eigenvalues are so dense that visually distinguishing between 4-fold and 8-fold degeneracies is typically impossible. Hence, we need a practical and robust method to identify would-be zero modes and nonzero modes of staggered fermions. The introduction of such a method is the main subject of the next section, Sec. V.

Using the chiral Ward identity of Eq. (47), we can measure the phase  $\theta$  numerically. In Fig. 3, we show numerical results (red circles) for  $\theta$ . Here, the blue line represents the theoretical prediction given in Eq. (48). We find the results are consistent with the prediction within numerical precision.

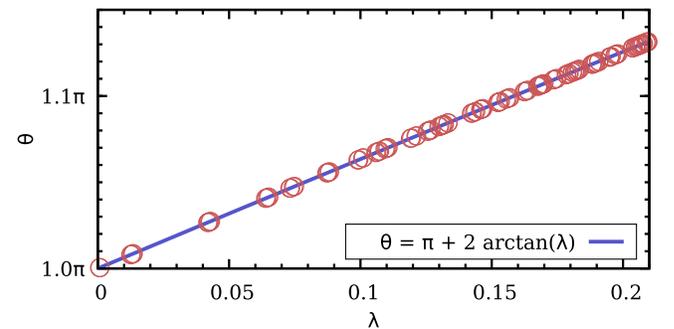


FIG. 3. The phase  $\theta$  as a function of  $\lambda$ . Red circles represent numerical results for  $\theta$ . The blue line represents the prediction from the theory. Here we use a gauge configuration with  $Q = -1$  for the measurement.

## V. CHIRALITY MEASUREMENT

To simplify the notation, we introduce the following convention for eigenvalue indices,

$$D_s|f_j\rangle = i\lambda_j|f_j\rangle, \quad (49)$$

where  $|f_j\rangle = |f_{\lambda_j}^s\rangle$  is defined in Eq. (28). Using the Kluberg-Stern method explained in Appendix D, we define the chirality operator and abbreviations as follows.

$$\begin{aligned} \Gamma_5(\lambda_i, \lambda_j) &\equiv \langle f_i | [\gamma_5 \otimes 1] | f_j \rangle \\ &\equiv \sum_x \sum_{A,B} [f_{\lambda_i}^s(x_A)]^\dagger [\gamma_5 \otimes 1]_{x,AB} f_{\lambda_j}^s(x_B), \end{aligned} \quad (50)$$

$$(\Gamma_5)_j^i \equiv \Gamma_5(\lambda_i, \lambda_j), \quad (51)$$

$$|\Gamma_5|_j^i \equiv |\Gamma_5(\lambda_i, \lambda_j)|, \quad (52)$$

where  $x_A, x_B$  and  $[\gamma_5 \otimes 1]$  are defined in Eqs. (38)–(40) in Sec. IV A, and  $\lambda_i$  and  $\lambda_j$  represent eigenvalues of  $D_s$ . The chirality operator  $[\gamma_5 \otimes 1]$  satisfies the same relationships as the continuum chirality operator  $\gamma_5$ .

$$[\gamma_5 \otimes 1]^{2n+1} = [\gamma_5 \otimes 1], \quad (53)$$

$$[\gamma_5 \otimes 1]^{2n} = [1 \otimes 1], \quad (54)$$

$$\left[ \frac{1}{2}(1 \pm \gamma_5) \otimes 1 \right]^n = \left[ \frac{1}{2}(1 \pm \gamma_5) \otimes 1 \right], \quad (55)$$

$$\left[ \frac{1}{2}(1 + \gamma_5) \otimes 1 \right] \left[ \frac{1}{2}(1 - \gamma_5) \otimes 1 \right] = 0, \quad (56)$$

where  $n$  is a non-negative integer. A rigorous proof of Eqs. (53)–(56) is given in Appendix E.

Our definition of the chirality operator  $[\gamma_5 \otimes 1]$  uses the Kluberg-Stern method, and is different from that used in Refs. [8,13,24], which adopt the Golterman method. The old chirality operator (the Golterman method) of Refs. [8,13,24] does not satisfy the recursion relations of Eqs. (53)–(56). In addition, it does not satisfy the chiral Ward identities of Eqs. (62)–(64). This difference is addressed in Appendices D and E. The bottom line is that the conventional chirality operator (the Golterman method) does not satisfy the recursion relationships in Eqs. (53)–(56), even though it is classified according to the true irreducible representations of the lattice rotational symmetry group [25,26,28].

For further discussion we define another operator  $[1 \otimes \xi_5]$ , which we call the “(maximal) shift operator,”

$$\begin{aligned} \Xi_5(\lambda_i, \lambda_j) &\equiv \langle f_i | [1 \otimes \xi_5] | f_j \rangle \\ &\equiv \sum_x \sum_{A,B} [f_{\lambda_i}^s(x_A)]^\dagger [1 \otimes \xi_5]_{x,AB} f_{\lambda_j}^s(x_B), \end{aligned} \quad (57)$$

$$(\Xi_5)_j^i \equiv \Xi_5(\lambda_i, \lambda_j), \quad (58)$$

$$|\Xi_5|_j^i \equiv |\Xi_5(\lambda_i, \lambda_j)|. \quad (59)$$

where  $x_A, x_B$  and  $[1 \otimes \xi_5]$  are defined in Eqs. (38)–(40) in Sec. IV A, and  $\lambda_i$  and  $\lambda_j$  represent eigenvalues of  $D_s$ . This shift operator satisfies the following recursion relations:

$$[1 \otimes \xi_5]^{2n+1} = [1 \otimes \xi_5], \quad (60)$$

$$[1 \otimes \xi_5]^{2n} = [1 \otimes 1], \quad (61)$$

where  $n$  is a non-negative integer. The conserved  $U(1)_A$  symmetry transformation can be expressed in terms of the chirality operator and the shift operator as follows,

$$\begin{aligned} \Gamma_e &\equiv [\gamma_5 \otimes \xi_5] \\ &= [\gamma_5 \otimes 1][1 \otimes \xi_5] \\ &= [1 \otimes \xi_5][\gamma_5 \otimes 1]. \end{aligned} \quad (62)$$

A rigorous proof of Eq. (62) is given in Appendix E. In addition, the chirality and shift operators satisfy the following relations:

$$\Gamma_e[\gamma_5 \otimes 1] = [\gamma_5 \otimes 1]\Gamma_e = [1 \otimes \xi_5], \quad (63)$$

$$\Gamma_e[1 \otimes \xi_5] = [1 \otimes \xi_5]\Gamma_e = [\gamma_5 \otimes 1]. \quad (64)$$

A rigorous proof of Eqs. (63)–(64) is also given in Appendix E. Therefore, we obtain the Ward identities:

$$\begin{aligned} e^{+i\theta}[\gamma_5 \otimes 1]|f_{-i}\rangle &= [1 \otimes \xi_5]|f_{+i}\rangle, \\ e^{-i\theta}[\gamma_5 \otimes 1]|f_{+i}\rangle &= [1 \otimes \xi_5]|f_{-i}\rangle, \end{aligned} \quad (65)$$

where

$$|f_{\pm i}\rangle \equiv |f_{\pm\lambda_i}^s\rangle. \quad (66)$$

Hence, we define the spectral decomposition by

$$[\gamma_5 \otimes 1]|f_j\rangle = \sum_i (\Gamma_5)_j^i |f_i\rangle \quad (67)$$

where we use Eqs. (50) and (51). Similarly,

$$[1 \otimes \xi_5]|f_j\rangle = \sum_i (\Xi_5)_j^i |f_i\rangle \quad (68)$$

where we use Eqs. (57) and (58). Thanks to the Ward identities of Eq. (65), we obtain

$$\begin{aligned}
e^{-i\theta}\Gamma_5(\lambda_i, +\lambda_j) &= \Xi_5(\lambda_i, -\lambda_j) \\
e^{-i\theta}(\Gamma_5)_{+j}^i &= (\Xi_5)_{-j}^i \\
|\Gamma_5|_{+j}^i &= |\Xi_5|_{-j}^i.
\end{aligned} \tag{69}$$

Similarly,

$$\begin{aligned}
e^{+i\theta}\Gamma_5(\lambda_i, -\lambda_j) &= \Xi_5(\lambda_i, +\lambda_j) \\
e^{+i\theta}(\Gamma_5)_{-j}^i &= (\Xi_5)_{+j}^i \\
|\Gamma_5|_{-j}^i &= |\Xi_5|_{+j}^i.
\end{aligned} \tag{70}$$

Applying  $\Gamma_\epsilon$  to both sides of Eq. (67), we obtain

$$[1 \otimes \xi_5]|f_j\rangle = \sum_{\ell} (\Gamma_5)_j^{\ell} e^{i\theta_{\ell}} |f_{-\ell}\rangle \tag{71}$$

$$= \sum_i (\Xi_5)_j^i |f_i\rangle. \tag{72}$$

Hence, we obtain another Ward identity:

$$|\Gamma_5|_j^{-i} = |\Xi_5|_j^{+i}. \tag{73}$$

Similarly, we can obtain the Ward identities:

$$|\Gamma_5|_{-j}^{-i} = |\Xi_5|_{-j}^{+i}, \tag{74}$$

$$|\Gamma_5|_j^{+i} = |\Xi_5|_j^{-i}. \tag{75}$$

We can summarize all the results of Eqs. (69)–(75) in the following form:

$$|\Gamma_5|_j^i = |\Xi_5|_j^{-i} = |\Xi_5|_{-j}^i = |\Gamma_5|_{-j}^{-i}, \tag{76}$$

$$\begin{aligned}
\Leftrightarrow |\Gamma_5(\lambda_i, \lambda_j)| &= |\Xi_5(-\lambda_i, \lambda_j)| = |\Xi_5(\lambda_i, -\lambda_j)| \\
&= |\Gamma_5(-\lambda_i, -\lambda_j)|.
\end{aligned} \tag{77}$$

In addition, Hermiticity ensures that we can interchange  $\lambda_i$  and  $\lambda_j$ , which gives the final form of the chiral Ward identities.

$$\begin{aligned}
|\Gamma_5|_j^i &= |\Xi_5|_j^{-i} = |\Xi_5|_{-j}^i = |\Gamma_5|_{-j}^{-i} = |\Gamma_5|_i^j = |\Xi_5|_i^{-j} \\
&= |\Xi_5|_{-i}^j = |\Gamma_5|_{-i}^{-j}
\end{aligned} \tag{78}$$

$$\begin{aligned}
\Leftrightarrow |\Gamma_5(\lambda_i, \lambda_j)| &= |\Xi_5(-\lambda_i, \lambda_j)| = |\Xi_5(\lambda_i, -\lambda_j)| \\
&= |\Gamma_5(-\lambda_i, -\lambda_j)| = |\Gamma_5(\lambda_j, \lambda_i)| = |\Xi_5(-\lambda_j, \lambda_i)| \\
&= |\Xi_5(\lambda_j, -\lambda_i)| = |\Gamma_5(-\lambda_j, -\lambda_i)|.
\end{aligned} \tag{79}$$

The quantity  $(|\Gamma_5|_j^i)^2$  for  $i \neq j$  represents the leakage probability of the chirality operator. We call  $|\Gamma_5|_j^i$  the leakage parameter for the chirality operator. Similarly, the quantity  $(|\Xi_5|_j^i)^2$  for  $i \neq j$  represents the leakage

probability of the shift operator, and we call  $|\Xi_5|_j^i$  the leakage parameter for the shift operator. By examining the leakage pattern, we can distinguish zero modes and nonzero modes, which is the main subject of the next subsections.

### A. Eigenvalue spectrum of $D_s$ in the continuum

Here we consider staggered quark actions at  $a = 0$ . We define a general form of the shift operator corresponding to a generator of the  $SU(4)$  taste symmetry:

$$\Xi_F = [1 \otimes \xi_F], \tag{80}$$

$$\xi_F \in \{\xi_5, \xi_\mu, \xi_{\mu 5}, \xi_{\mu\nu}\} \quad \text{for } \mu \neq \nu, \tag{81}$$

where  $\xi_\mu$  respects the Clifford algebra  $\{\xi_\mu, \xi_\nu\} = 2\delta_{\mu\nu}$  in Euclidean space.

Let us consider the following quantity  $W_1$  in the continuum:

$$W_1 \equiv \langle f_\ell | \Xi_F D_s | f_n \rangle \tag{82}$$

$$D_s | f_n \rangle = i\lambda_n | f_n \rangle. \tag{83}$$

Since the  $SU(4)$  taste symmetry is exactly conserved in the continuum, we know that

$$[\Xi_F, D_s] = 0 \tag{84}$$

Hence, we find the following Ward identity:

$$W_1 = \langle f_\ell | \Xi_F D_s | f_n \rangle = i\lambda_n \langle f_\ell | \Xi_F | f_n \rangle \tag{85}$$

$$= \langle f_\ell | D_s \Xi_F | f_n \rangle = i\lambda_\ell \langle f_\ell | \Xi_F | f_n \rangle \tag{86}$$

and therefore

$$i(\lambda_\ell - \lambda_n) \cdot \langle f_\ell | \Xi_F | f_n \rangle = 0. \tag{87}$$

Hence, in the continuum, we find the following properties of the eigenvalue spectrum:

- (i) If  $\lambda_\ell \neq \lambda_n$ ,  $(\Xi_F)_n^\ell = \langle f_\ell | \Xi_F | f_n \rangle = 0$ . In other words, if the eigenvalues are different, there is no leakage ( $(\Xi_F)_n^\ell = 0$ ) between the two eigenmodes.
- (ii) If  $\lambda_j \equiv \lambda_\ell = \lambda_n$ ,  $(\Xi_F)_n^\ell \neq 0$  is possible. In this case, the eigenvalues are degenerate and belong to a quartet such that

$$D_s | f_{j,m} \rangle = i\lambda_j | f_{j,m} \rangle \tag{88}$$

$$|f_\ell\rangle, |f_n\rangle \in \{|f_{j,m}\rangle \text{ with } m = 1, 2, 3, 4\}, \tag{89}$$

where  $m$  is a taste index which represents the four-fold degeneracy for the eigenvalue  $\lambda_j$ ,  $|f_\ell\rangle$  and  $|f_n\rangle$  are linear combinations of the quartet  $\{|f_{j,m}\rangle\}$ , and

the eigenvectors for different  $m$  are orthogonal to each other by construction due to the Lanczos algorithm.

- (iii) We know that the staggered fermion field  $\chi^c(x_A)$  is mapped into the continuum fermion field  $\psi_{\alpha,t}^c(x)$ , where  $\alpha$  represents a Dirac spinor index,  $c$  represents a color index, and  $t = 1, 2, 3, 4$  represents a taste index. Hence, for a given eigenvalue  $\lambda_j$ , there remain four degrees of freedom which come from the taste index. Accordingly, for a given eigenvalue  $\lambda_j$ , there are four degenerate eigenstates  $|f_{j,m}\rangle$  with  $m = 1, 2, 3, 4$ .
- (iv) If we know all four eigenstates  $\{|f_{j,m}\rangle\}$  for a certain eigenvalue  $\lambda_j$ , we find that

$$\begin{aligned} \text{Tr}(\Xi_F) &= \sum_{m=1}^4 (\Xi_F)_{j,m}^{j,m} \\ &= \sum_{m=1}^4 \langle f_{j,m} | \Xi_F | f_{j,m} \rangle = 0 \end{aligned} \quad (90)$$

This is because the  $SU(4)$  group generators are traceless in the fundamental representation.

However, on the lattice at  $a \neq 0$ , the taste symmetry is broken by terms of order  $a^2 \alpha_s^n$  with  $n \geq 1$ , which is explained in Ref. [20]. In addition, for  $a \neq 0$ ,

$$D_s |f_{j,m}\rangle = i\lambda_{j,m} |f_{j,m}\rangle \quad (91)$$

and  $\lambda_{j,m} \neq \lambda_{j,m'}$  in general for  $m \neq m'$ , which reflects the taste symmetry breaking effect at  $a \neq 0$ . We know that  $\lambda_{j,m} = \lambda_{j,m'}$  for all  $m, m'$  in the continuum ( $a = 0$ ) due to the exact taste symmetry. Hence, on the finite lattice, we expect a small deviation from the above continuum properties. A good barometer to measure this effect is to monitor  $T_5$

$$T_5 \equiv \frac{1}{4} \text{Tr}(\Xi_5) = \frac{1}{4} \sum_m (\Xi_5)_{j,m}^{j,m} \quad (92)$$

and measure how much it deviates from zero (the continuum value). Another direct barometer to measure effects of taste symmetry breaking is the leakage  $S_5$  from one quartet ( $\lambda_\ell$ ) to another quartet ( $\lambda_j$ ) with  $\lambda_\ell \neq \lambda_j$ .

$$S_5 \equiv \frac{1}{16} \sum_{m,m'} |\Xi_5|_{j,m'}^{\ell,m} = \frac{1}{16} \sum_{m,m'} |\langle f_{\ell,m} | \Xi_5 | f_{j,m'} \rangle| \quad (93)$$

The size of the leakage  $S_5$  indicates directly how much the taste symmetry is broken at  $a \neq 0$ , since  $S_5 = 0$  in the continuum. We present numerical results for  $T_5$  and  $S_5$  in the next subsection.

## B. Numerical study on chirality and leakage

Here we use dual notations for the eigenmodes: One is the serial index  $i$  for  $\lambda_i$  and the other is the quartet index  $j$  with taste index  $m$  for  $\lambda_{j,m}$ . The serial index is convenient for the plots, tables, and leakage patterns such as  $|\Gamma_5|_b^a$ , while the quartet index is convenient to explain the eigenstates classified by the taste symmetry group. The one-to-one mapping from the serial index  $i$  to the quartet indices  $j, m$  is given in Table II for the quartet index  $j = 0, \pm 1$  when  $Q = \pm 1$ . The mapping for the quartet index  $j = \pm 2$  (nonzero modes) is similar.

In Fig. 4, we present the leakage pattern of the zero mode of  $\lambda_1$  and its parity partner  $\lambda_2 = -\lambda_1$ . Since  $Q = -1$  in Fig. 4, we expect to observe four-fold degenerate would-be zero modes within a single quartet (quartet index  $j = 0$ ).

$$\lim_{a \rightarrow 0} \lambda_i = 0 \quad \text{for } i = 1, 2, 3, 4. \quad (94)$$

In the continuum limit ( $a = 0$ ), the  $SU(4)$  taste symmetry becomes exactly conserved and so would-be zero modes become exact zero modes. However, at finite lattice spacing  $a \neq 0$ , the gauge configuration is sufficiently rough that would-be zero modes have nonzero eigenvalues:  $\lambda_2 = -\lambda_1$ ,  $\lambda_4 = -\lambda_3$ , and  $\lambda_1 \neq \lambda_3$  for  $\lambda_1, \lambda_3 > 0$ .

In Fig. 4(a), we show the leakage pattern of  $|\Gamma_5|_1^i = |\Gamma_5(\lambda_i, \lambda_1)| = |\langle f_i | \Gamma_5 | f_1 \rangle|$ . We find that there is, in practice, no leakage, and so the only nonzero component is  $|\Gamma_5|_1^1 = |\Gamma_5(\lambda_1, \lambda_1)|$ . The other components are practically zero. In Figs. 4(b), 4(c), and 4(d), we find that the Ward identity of Eqs. (78) and (79) is well-respected by the numerical results. In other words, the Ward identity  $|\Gamma_5|_1^1 = |\Xi_5|_1^2 = |\Xi_5|_2^2 = |\Gamma_5|_2^2$  is satisfied within the numerical precision of the computer. Please refer to Table III for numerical details. This confirms that the theoretical

TABLE II. One-to-one mapping of serial index  $i$  of the  $\lambda_i$  eigenstate into a quartet index  $j$  and taste index  $m$  for  $\lambda_{j,m} = \lambda_i$ . Here  $\lambda_{2n} = -\lambda_{2n-1}$  and  $\lambda_{-j,m} = -\lambda_{+j,m}$ . Here we assume  $Q = \pm 1$ .

$\lambda_i$	$\lambda_{j,m}$	$i$	$j$	$m$	Mode
$\lambda_1$	$\lambda_{0,1}$	1	0	1	Zero
$\lambda_2$	$\lambda_{0,2}$	2	0	2	Zero
$\lambda_3$	$\lambda_{0,3}$	3	0	3	Zero
$\lambda_4$	$\lambda_{0,4}$	4	0	4	Zero
$\lambda_5$	$\lambda_{+1,1}$	5	+1	1	Nonzero
$\lambda_7$	$\lambda_{+1,2}$	7	+1	2	Nonzero
$\lambda_9$	$\lambda_{+1,3}$	9	+1	3	Nonzero
$\lambda_{11}$	$\lambda_{+1,4}$	11	+1	4	nonzero
$\lambda_6$	$\lambda_{-1,1}$	6	-1	1	Nonzero
$\lambda_8$	$\lambda_{-1,2}$	8	-1	2	Nonzero
$\lambda_{10}$	$\lambda_{-1,3}$	10	-1	3	Nonzero
$\lambda_{12}$	$\lambda_{-1,4}$	12	-1	4	Nonzero

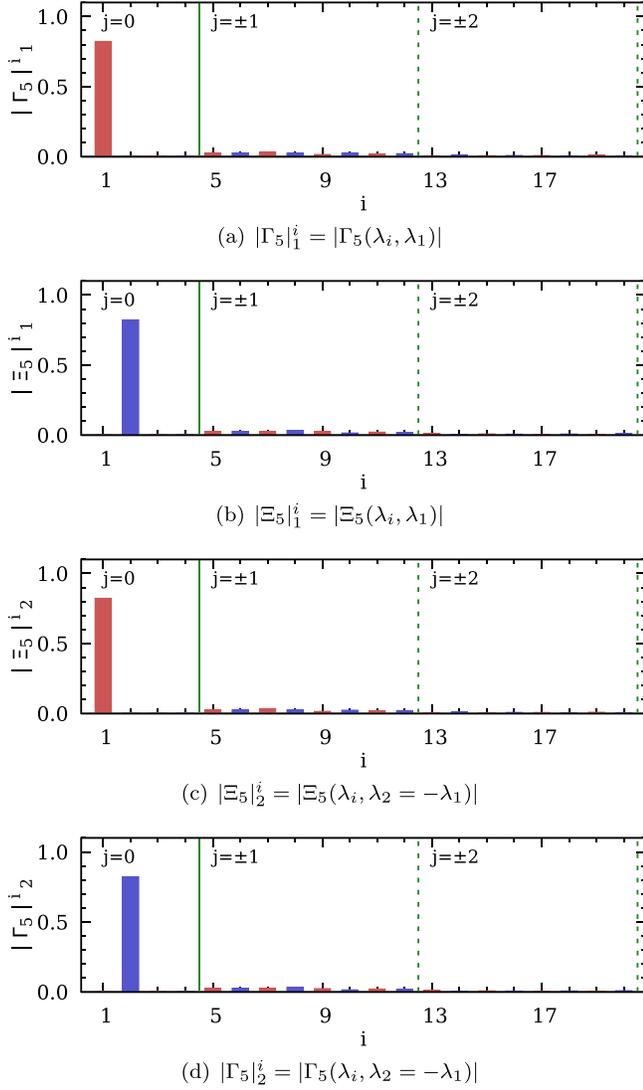


FIG. 4. Leakage pattern for would-be zero modes at  $Q = -1$ . Here, the red bar represents leakage to  $\lambda_{i=2n-1} > 0$  with  $i$  odd, and the blue bar represents leakage to its parity partner  $\lambda_{i=2n} = -\lambda_{2n-1}$  with  $i$  even.

prediction from the Ward identity in Eqs. (78) and (79) is correct.

In Fig. 4(a), we find that there is small leakage into other quartets ( $j = \pm 1, \pm 2$ ). The typical size of leakage between off-diagonal elements of the would-be zero modes, the  $j = 0$  quartet, (e.g.,  $|\Gamma_5|_1^3$ ) is of order  $10^{-3}$ . We also observe small leakage patterns of order  $10^{-2} \sim 10^{-3}$  from the would-be zero modes, the  $j = 0$  quartet, to the nonzero modes, the  $j = \pm 1, \pm 2$  quartets (e.g.,  $|\Gamma_5|_1^5$ ).

Now let us consider nonzero modes in the  $j = +1$  quartet. In Fig. 5, we present the leakage pattern for the nonzero modes of  $\lambda_5$  and its parity partner  $\lambda_6 = -\lambda_5$ . Even in the continuum limit ( $a = 0$ ),  $\lambda_5 \neq 0$ , and so it is a nonzero mode. Thanks to the approximate  $SU(4)$  taste symmetry and the exact  $U(1)_A$  axial symmetry, there will

TABLE III. Numerical values for leakage patterns from the  $\lambda_1$  eigenstate to the  $\lambda_i$  eigenstate in Fig. 4. Here,  $j$  represents a quartet index for the  $\lambda_i$  eigenstate. The leakage represents  $|\mathcal{O}|_1^i = |\mathcal{O}(\lambda_i, \lambda_1)| = |\langle f_i | \mathcal{O} | f_1 \rangle|$  for  $\mathcal{O} = \Gamma_5, \Xi_5$ .

$j$	Leakage	Value	Ward identity
0	$ \Gamma_5 _1^1$	0.82382566818582	$= \Xi_5 _1^2$
0	$ \Xi_5 _1^2$	0.82382566818581	$= \Xi_5 _1^2$
0	$ \Xi_5 _2^1$	0.82382566818580	$= \Gamma_5 _2^2$
0	$ \Gamma_5 _2^2$	0.82382566818579	$= \Gamma_5 _1^1$
0	$ \Gamma_5 _1^2$	$6.67 \times 10^{-4}$	
0	$ \Gamma_5 _1^3$	$1.34 \times 10^{-3}$	
0	$ \Gamma_5 _1^4$	$1.79 \times 10^{-3}$	
+1	$ \Gamma_5 _1^5$	$2.56 \times 10^{-2}$	
-1	$ \Gamma_5 _1^6$	$2.54 \times 10^{-2}$	
+2	$ \Gamma_5 _1^3$	$5.77 \times 10^{-3}$	
-2	$ \Gamma_5 _1^4$	$1.18 \times 10^{-2}$	

be eight-fold degeneracy in the family of eight eigenstates composed of the  $j = +1$  quartet, to which  $\lambda_5$  belongs, and the  $j = -1$  quartet (the parity partners). These eight-fold degenerate modes are designated together as  $j = \pm 1$  quartets in Fig. 5, where they are a set of  $\{\lambda_i\}$  with  $5 \leq i \leq 12$ .

Let us scrutinize the leakage pattern of the nonzero mode  $\lambda_5 = \lambda_{j=+1, m=1}$ . In Fig. 5(a), first note that there is practically no leakage in the  $\Gamma_5$  chirality measurement from  $\lambda_5$  into  $\lambda_{2n-1}$  with  $n > 0$ . In other words,  $|\Gamma_5|_5^{2n-1} = |\Gamma_5(\lambda_{2n-1}, \lambda_5)| \cong 0$ . This implies that the chirality operator on the nonzero mode with  $\lambda > 0$  leaks into only the parity partner modes with  $\lambda < 0$ . Second, note that the nontrivial leakage goes to those eigenstates in the  $j = -1$  quartet such as  $\{\lambda_6, \lambda_8, \lambda_{10}, \lambda_{12}\} = \{\lambda_{j,m} | j = -1, m = 1, 2, 3, 4\}$ . In addition, we find that the Ward identity of Eqs. (78) and (79) is well-respected within the numerical precision in Figs. 5(a), 5(b), 5(c), and 5(d). In Table IV, we present numerical values of the  $|\Gamma_5|_5^i$  shown in Fig. 5(a).

Let us examine the  $\Gamma_5 = [\gamma_5 \otimes 1]$  leakage pattern of the  $j = +1$  quartet of the nonzero modes  $\{\lambda_5, \lambda_7, \lambda_9, \lambda_{11}\}$ . In Fig. 6, we find that the chirality measurement vanishes;  $(\Gamma_5)_i^i = \Gamma_5(\lambda_i, \lambda_i) = 0$  for  $\lambda_i$  in the  $j = +1$  quartet of nonzero modes. We also find that the  $\Gamma_5$  leakage of  $\lambda_{+1, m} > 0$  of the  $j = +1$  quartet goes to the parity partners with  $\lambda_{-1, m'} < 0$  of the  $j = -1$  quartet, and the leakage to other quartets such as  $j = \pm 2$  is negligibly small compared to the leakage to the  $j = -1$  quartet. The numerical values of  $|\Gamma_5|_{+1, m'}^{-1, m}$  are summarized in Table V.

Let us examine the  $\Xi_5 = [1 \otimes \xi_5]$  leakage pattern of the  $j = +1$  quartet of the nonzero modes  $\{\lambda_5, \lambda_7, \lambda_9, \lambda_{11}\}$ . In Fig. 7, we find that the  $\Xi_5$  leakage from the  $j = +1$  quartet to the  $j = -1$  quartet (parity partners) vanishes in practice. Since the leakage pattern of  $\Xi_5$  is related to the leakage pattern of  $\Gamma_5$  by the Ward identity

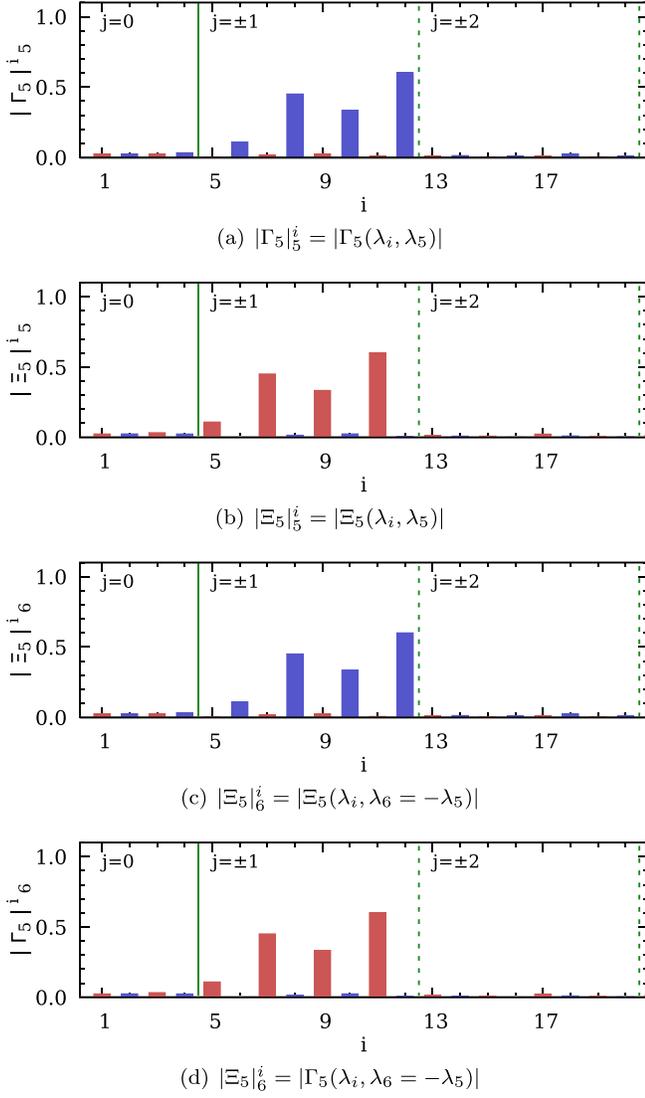
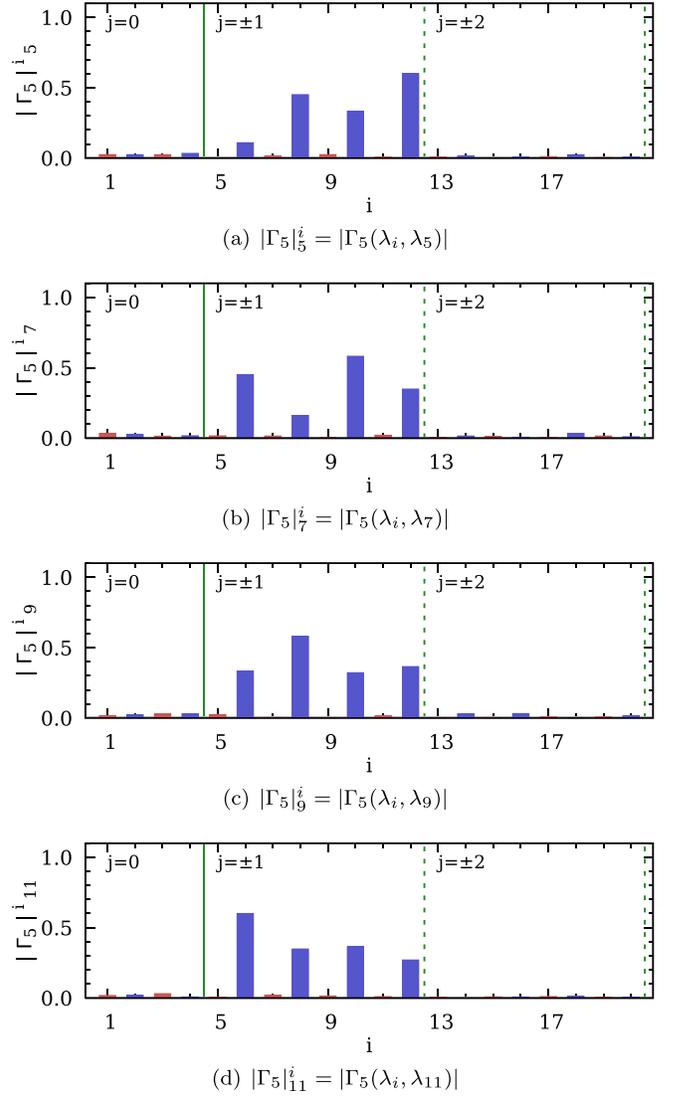
FIG. 5. Leakage pattern for nonzero modes at  $Q = -1$ .FIG. 6.  $[\gamma_5 \otimes 1]$  leakage pattern for nonzero modes at  $Q = -1$ .

TABLE IV. Numerical values for data in Fig. 5.

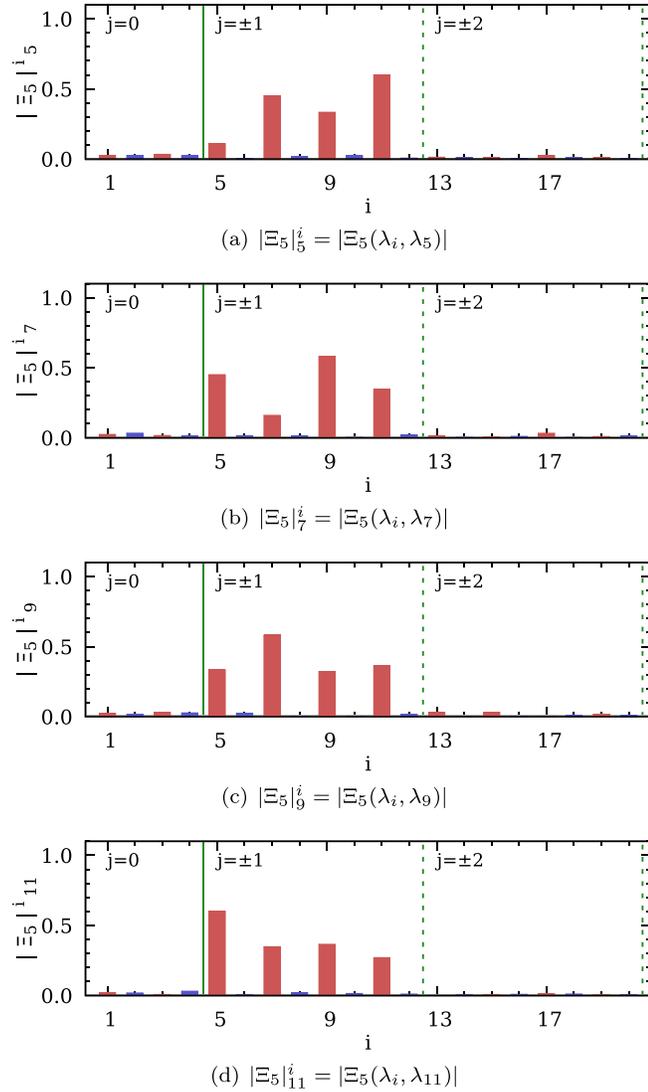
$j$	Leakage	Value	Ward identities
-1	$ \Gamma_5 _5^6$	0.110	$= \Xi_5 _5^5 =  \Xi_5 _6^6 =  \Gamma_5 _5^5$
-1	$ \Gamma_5 _5^8$	0.452	$= \Xi_5 _5^7 =  \Xi_5 _6^8 =  \Gamma_5 _6^7 =  \Gamma_5 _8^5 =  \Xi_5 _7^5 =  \Xi_5 _8^6 =  \Gamma_5 _7^6$
-1	$ \Gamma_5 _5^{10}$	0.334	$= \Xi_5 _5^9 =  \Xi_5 _6^{10} =  \Gamma_5 _6^9 =  \Gamma_5 _{10}^5 =  \Xi_5 _9^5 =  \Xi_5 _{10}^6 =  \Gamma_5 _9^6$
-1	$ \Gamma_5 _5^{12}$	0.601	$= \Xi_5 _5^{11} =  \Xi_5 _6^{12} =  \Gamma_5 _6^{11} =  \Gamma_5 _{12}^5 =  \Xi_5 _{11}^5 =  \Xi_5 _{12}^6 =  \Gamma_5 _{11}^6$
+1	$ \Gamma_5 _5^5$	$2.05 \times 10^{-3}$	$= \Xi_5 _5^6 =  \Xi_5 _6^5 =  \Gamma_5 _6^6$
+1	$ \Gamma_5 _5^7$	$16.7 \times 10^{-3}$	$= \Xi_5 _5^8 =  \Xi_5 _6^7 =  \Gamma_5 _6^8 =  \Gamma_5 _7^5 =  \Xi_5 _8^5 =  \Xi_5 _7^6 =  \Gamma_5 _8^6$
+1	$ \Gamma_5 _5^9$	$25.6 \times 10^{-3}$	$= \Xi_5 _5^{10} =  \Xi_5 _6^9 =  \Gamma_5 _6^{10} =  \Gamma_5 _9^5 =  \Xi_5 _{10}^5 =  \Xi_5 _9^6 =  \Gamma_5 _{10}^6$
+1	$ \Gamma_5 _5^{11}$	$7.32 \times 10^{-3}$	$= \Xi_5 _5^{12} =  \Xi_5 _6^{11} =  \Gamma_5 _6^{12} =  \Gamma_5 _{11}^5 =  \Xi_5 _{12}^5 =  \Xi_5 _{11}^6 =  \Gamma_5 _{12}^6$
0	$ \Gamma_5 _5^3$	$2.52 \times 10^{-2}$	
0	$ \Gamma_5 _5^4$	$3.43 \times 10^{-2}$	
+2	$ \Gamma_5 _5^{13}$	$1.02 \times 10^{-2}$	
-2	$ \Gamma_5 _5^{14}$	$1.38 \times 10^{-2}$	

TABLE V.  $|\Gamma_5|_{+1,m}^{-1,m}$  values in Fig. 6.

$\lambda_j$	$\lambda_i$			
	$\lambda_5$	$\lambda_7$	$\lambda_9$	$\lambda_{11}$
$\lambda_6$	0.110	0.452	0.334	0.601
$\lambda_8$	0.452	0.161	0.582	0.349
$\lambda_{10}$	0.334	0.582	0.323	0.366
$\lambda_{12}$	0.601	0.349	0.366	0.271

$$|\Xi_5|_{j',m'}^{j,m} = |\Gamma_5|_{j',m'}^{-j,m}, \quad (95)$$

Fig. 7 can be obtained from Fig. 6 using the Ward identity. We find that the  $\Xi_5$  leakage from the  $j = +1$  quartet to other quartets such as  $j = \pm 2$  quartets is negligibly small compared to leakage to itself (the  $j = +1$  quartet). Leakage


 FIG. 7.  $[1 \otimes \xi_5]$  leakage pattern for nonzero modes at  $Q = -1$ .

patterns of the  $\Gamma_5$  chirality and  $\Xi_5$  shift operators for diverse topological charges are shown in Appendix G.

Let us summarize the leakage pattern for would-be zero modes and that for nonzero modes. We first begin with the leakage pattern for the zero modes.

1. A zero mode of staggered fermions appears as a four-fold degenerate quartet. In other words, for the topological charge  $Q$ , the number of zero modes is  $4 \times (n_+ + n_-)$ , and  $Q = n_- - n_+$  (Atiyah-Singer index theorem), where  $n_+$  ( $n_-$ ) is the number of right-handed (left-handed) zero mode quartets.
2. In the chirality  $\Gamma_5 = [\gamma_5 \otimes 1]$  measurement, the zero mode has practically no leakage to other eigenstates.
3. In the shift  $\Xi_5 = [1 \otimes \xi_5]$  measurement, the zero mode with eigenvalue  $\lambda$  has a full (100%) leakage into its parity partner mode with eigenvalue  $-\lambda$ , and no leakage into any other eigenmodes.

The leakage pattern for nonzero modes is

1. A nonzero mode of staggered fermions appears as an eight-fold degeneracy composed of a quartet ( $+j$  quartet) and its parity partner quartet ( $-j$  quartet). In other words, nonzero eigenmodes can be grouped into sets with eight elements in each set. This is due to the approximate  $SU(4)$  taste symmetry and the conserved  $U(1)_A$  axial symmetry.
2. In the chirality  $\Gamma_5 = [\gamma_5 \otimes 1]$  measurement, the nonzero mode with eigenvalue  $\lambda_{j,m}$  has no leakage to its own quartet ( $j$  quartet), but has leakage only to the parity partner ( $-j$  quartet) with  $\lambda_{-j,m'}$ . It has no leakage to any eigenmode which belongs to other quartets with  $\ell \neq \pm j$ .
3. In the shift  $\Xi_5 = [1 \otimes \xi_5]$  measurement, the nonzero mode with  $\lambda_{j,m}$  has no leakage to its parity partner ( $-j$  quartet) at all. But it has leakage only to the eigenstates in its own ( $+j$ ) quartet. This pattern comes directly from the Ward identity. In other words, the  $\Xi_5$  leakage pattern is a mirror image reflecting  $\Gamma_5$  through the mirror of the Ward identity.  $\Xi_5$  has no leakage to any eigenmode which belongs to other quartets with  $\ell \neq \pm j$ .
4. Thanks to the Ward identity of the conserved  $U(1)_A$  symmetry, the leakage pattern of  $|\Gamma_5|_{\ell,m'}^{-j,m}$  is identical to that of  $|\Xi_5|_{\ell,m'}^{+j,m}$ .

In Appendix F, we provide more examples to demonstrate our claim that the leakage pattern for zero modes holds in general. In Appendix G, we give more examples to demonstrate our claim that the leakage pattern for nonzero modes is valid in general. We have repeated numerical tests over hundreds of zero modes and tens of thousands of nonzero modes. We performed the numerical study on hundreds of gauge configurations and find that the above leakage pattern is valid for all of them except for those gauge configurations with unstable topological charge.

TABLE VI. Numerical results for  $T_5$ . To obtain the results, we use 292 gauge configurations with the input parameters in Table I.  $N_q$  represents the number of quartets used to obtain the statistical error. Here  $j = 0$  represents would-be zero modes, and  $j > 0$  represents nonzero modes.

$j$	$ \text{Re}(T_5) $	$ \text{Im}(T_5) $	$N_q$
$j = 0$	$7.2(130) \times 10^{-4}$	$5.9(46) \times 10^{-12}$	490
$j > 0$	$6.2(120) \times 10^{-3}$	$3.3(25) \times 10^{-12}$	7034

1. We find a number of gauge configurations which do not have a stable topological charge.
2. We find about 10 gauge configurations with unstable topological charge among 100 gauge configurations with  $12^4$  lattice geometry at  $\beta = 4.6$ .
3. We find about 8 gauge configurations with unstable topological charge among 300 gauge configurations with  $20^4$  lattice geometry at  $\beta = 5.0$ .

In Table VI, we present results for  $T_5$  defined in Eq. (92), which is a direct barometer to estimate the effect of taste symmetry breaking. If the taste symmetry is exactly conserved, then  $T_5$  must vanish. Hence, a nontrivial value of  $T_5$  indicates the size of taste symmetry breaking. In Table VI, we find that  $|\text{Re}(T_5)|$  is of order  $10^{-3}$ , while  $|\text{Im}(T_5)|$  is essentially zero. This indicates that the effect of taste symmetry breaking is very small (in the subpercent level within each quartet).

In Fig. 8, we present  $S_5$  defined in Eq. (93) as a function of  $|\ell - j|$  with  $\ell, j \geq 0$ . Here  $|\ell - j| = 1$  represents a pair of nearest neighbor quartets,  $|\ell - j| = 2$  represents a pair of next-to-nearest neighbor quartets, and so on. The values of  $S_5$  are the same size as the statistical errors. This indicates that the taste symmetry breaking results in simply random noise added to the physical signal ( $S_5 = 0$ ). For  $|\ell - j| = 1$ , the noise is  $\approx 7\%$ , and for  $|\ell - j| = 2$ , the noise is  $\approx 3\%$ . We find that the noise decreases as  $|\ell - j|$  increases. The numerical values of  $S_5$  in Fig. 8 are presented in Table VII.

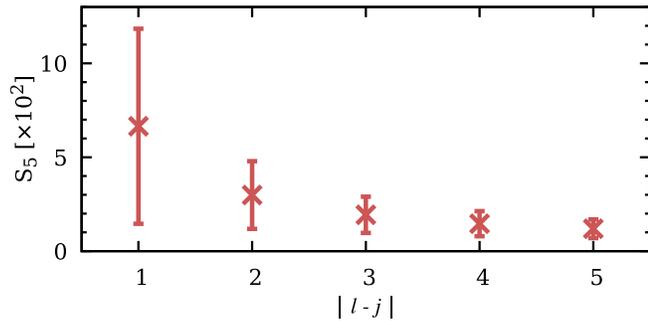


FIG. 8.  $S_5$  as a function of  $|\ell - j|$ . Numerical values are given in Table VII.

TABLE VII. Numerical results for  $S_5$ . Here, we measure  $S_5$  between two different quartets ( $\ell \neq j$  and  $\ell, j \geq 0$ ).  $N_p$  represents the number of  $(\ell, j)$  pairs with  $\ell \neq j$ .

$ \ell - j $	$S_5$	$N_p$
1	$6.6(52) \times 10^{-2}$	7185
2	$3.0(18) \times 10^{-2}$	6893
3	$1.9(10) \times 10^{-2}$	6601
4	$1.5(7) \times 10^{-2}$	6309
5	$1.2(5) \times 10^{-2}$	6017

## VI. MACHINE LEARNING

In previous sections, we have shown that the  $U(1)_A$  symmetry of staggered fermions induces the chiral Ward identities in Eq. (78), and we have also noted that the approximate  $SU(4)$  taste symmetry brings in the quartet behavior of the eigenvalue spectrum. Furthermore, a combined effect of those symmetries gives us distinctive leakage patterns for the chirality operator  $\Gamma_5$  and the shift operator  $\Xi_5$ . In this section, we apply a machine learning technique to the following tasks.

1. We want to know how much the nonzero modes respect the quartet classification rules, which come from the  $SU(4)$  taste symmetry.
2. We want to know how efficiently we can measure the topological charge  $Q$  using the index theorem from the quartet structure of the nonzero modes.
3. We want to detect any anomalous behavior of the eigenvalue spectrum, which does not follow the standard leakage pattern of the nonzero modes.
4. We want to figure out what causes the anomalous behavior of the eigenvalue spectrum.

Let us explain our sampling method for the machine learning. In Fig. 9, we show matrix elements  $|\Gamma_5|_j^i$  on a gauge configuration with  $Q = 2$ . Figure 9(a) is for the 200 lowest eigenmodes, and Fig. 9(b) is a zoomed-in version of Fig. 9(a) for the 32 lowest eigenmodes. Here the depth of the blue color represents the size of the matrix element  $|\Gamma_5|_j^i$ , and  $i, j$  run from zero to 199. We identify two zero mode quartets (red boxes) by looking at the magnitude of the diagonal components. These two quartets have the same chirality ( $n_- = 2$ ), which is consistent with the topological charge  $Q = 2$ . Excluding the would-be zero modes, we randomly choose a  $15 \times 15$  sub-matrix of  $|\Gamma_5|_j^i$  along the diagonal line of  $|\Gamma_5|_j^i$  matrix elements. This  $15 \times 15$  sub-matrix is the smallest square sub-matrix of  $|\Gamma_5|$  which contains all elements of only one quartet of nonzero modes and its parity partner quartet.

In Fig. 10, we present 8 different classes for arbitrary samples. Our purpose for the machine learning is to find borders (black lines) of the nonzero mode quartets (or octets when the parity partners are included) in each sample. We classify arbitrary samples into eight different

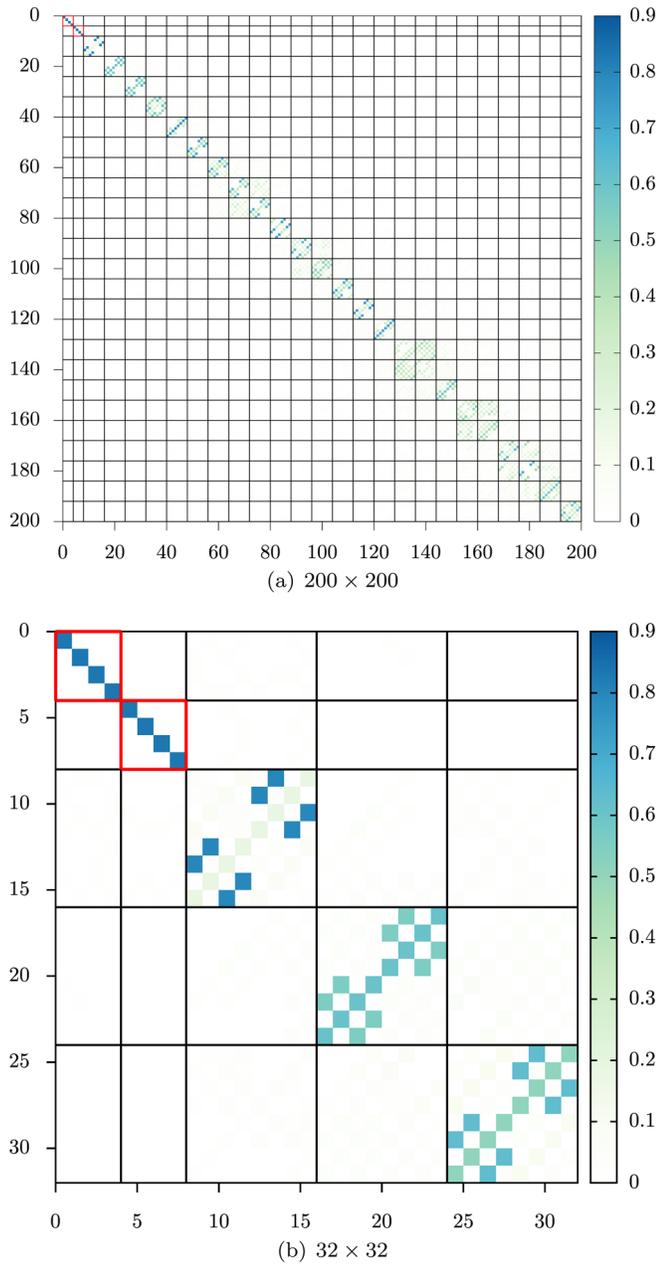


FIG. 9. Matrix elements of  $|\Gamma_5|$  for 200 and 32 of the lowest eigenmodes on a gauge configuration with  $Q = 2$ . Here, indices on both axes are the eigenvalue index. The color of each square represents the magnitude of the corresponding matrix element. Black lines indicate borders of nonzero mode quartets, and red lines are those of zero mode quartets.

classes according to the location of the border lines. Each class is labeled as in Fig. 10.

We use a deep learning model which combines the multilayer perceptron (MLP) [46] and the convolutional neural network (CNN) [46]. In Table VIII, we present our basic setup for the machine learning. We use the gauge configuration ensemble described in Table I. The data measured over 292 gauge configurations are distributed over a training set, validation set, and test set as in

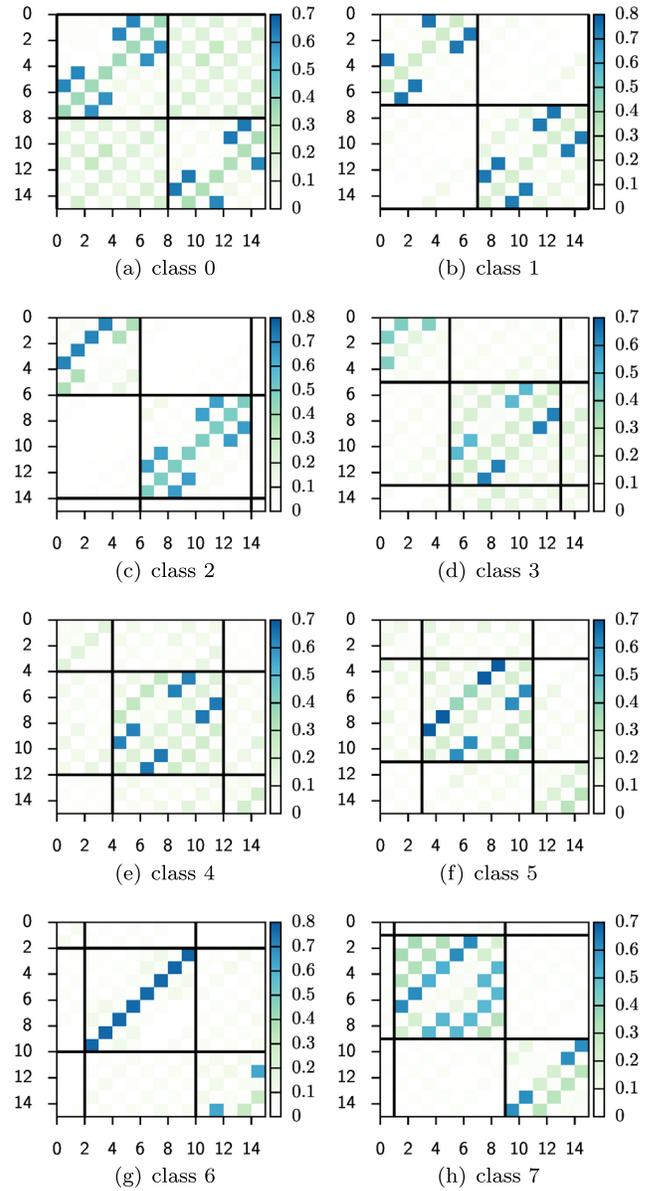


FIG. 10. Examples for our samples. Every sample contains only one nonzero mode quartet. There are eight kinds of classes according to the location of the borders of the quartet.

Table VIII. For each gauge configuration, we generate around ten  $15 \times 15$  matrix samples from the 200 lowest eigenmodes without overlapping. We make popular and suitable choices for the loss function [47], optimization method [48], and activation functions [50] relevant to our purpose, which are summarized in Table VIII. The best hyperparameters such as the number of layers and the number of units for each layer are determined by using the Keras Tuner [51].

The accuracy of classification per gauge configuration is obtained by averaging the accuracies of the machine learning (ML) prediction for all the samples on a single gauge configuration. Our best model achieves an average accuracy of 96.5(156)% for 142 test gauge configurations.

TABLE VIII. Parameters for machine learning.

Parameters	Values
Number of training configurations	120
Number of training samples	1223
Number of validation configurations	30
Number of validation samples	308
Number of test configurations	142
Number of test samples	1448
Loss function	Categorical Cross-entropy [46,51]
Optimization method	Adam [49]
Activation function for hidden layers	ReLU [46]
Activation function for output layer	Softmax [46]

TABLE IX. Hyper-parameters for neural networks. Here, we show one of the examples of best performance, in which we use only MLP but not CNN.

Layer	Type	Number of units	Activation
Input	...	225	...
Hidden #1	MLP	160	ReLU
Hidden #2	MLP	1210	ReLU
Hidden #3	MLP	1490	ReLU
Output	MLP	8	Softmax

The hyperparameters which represent the structure of the neural network model are given in Table IX. Among the test set, we find five gauge configurations on which the average accuracy per gauge configuration is lower than 50%. Data show that some ghost (unphysical) eigenvectors are present in the eigenvalue spectrum on these gauge configurations, so that the ML prediction gives a wrong answer not due to failure of the ML algorithm but due to human mistakes in labeling quartet samples based on the eigenvalue index. Excluding these five gauge configurations, we achieve the average accuracy of 99.4(23)%. Considering that all samples generated on the same gauge configuration are connected by the eigenvalue index (or quartet index), this average accuracy of 99.4% implies that one can in the end find completely correct quartet groups for all the normal gauge configurations of the test set. It also demonstrates our claim that the leakage pattern is universal over all normal gauge configuration ensembles. Details of the results of this ML research will be reported separately in Ref. [52].

## VII. ZERO MODES AND RENORMALIZATION

As explained in Sec. V, we know that there is practically no leakage for the zero modes in the chirality measurement. Hence, it is possible to determine the renormalization factor  $\kappa_P$  by imposing the index theorem as follows. For  $Q \neq 0$ ,

TABLE X. Numerical results for  $\kappa_P$ .

Topological charge	Number of samples	$\kappa_P$
$ Q  = 1$	72	1.26(13)
$ Q  = 2$	68	1.22(3)
$ Q  = 3$	45	1.23(2)
Weighted average	241	1.23(2)

$$4 \times Q = -\kappa_P \times \sum_{\lambda \in S_0} \langle f_{\lambda}^s | [\gamma_5 \otimes 1] | f_{\lambda}^s \rangle \quad (96)$$

$$\kappa_P = -\frac{4Q}{C_0} \quad (97)$$

$$C_0 = \sum_{\lambda \in S_0} \Gamma_5(\lambda, \lambda) \quad (98)$$

where  $S_0$  is the entire set of zero modes, and

$$\kappa_P = \frac{Z_{P \times S}(\mu)}{Z_{P \times P}(\mu)}, \quad (99)$$

where

$$O_S = \bar{\chi} [\gamma_5 \otimes 1] \chi \quad (100)$$

$$O_P = \bar{\chi} [\gamma_5 \otimes \xi_5] \chi \quad (101)$$

$$[O_S]_R(\mu) = Z_{P \times S}(\mu) [O_S]_B \quad (102)$$

$$[O_P]_R(\mu) = Z_{P \times P}(\mu) [O_P]_B, \quad (103)$$

and the subscript  $[\dots]_R$  ( $[\dots]_B$ ) represents a renormalized (bare) operator. The  $Z_{P \times S}$  and  $Z_{P \times P}$  are the renormalization factors for the bilinear operators  $O_S$  and  $O_P$ , respectively. One advantage of this scheme is that  $\kappa_P$  is independent of valence quark masses, even though we perform the measurement with arbitrary masses for valence quarks. Numerical results for  $\kappa_P$  are summarized in Table X.

There are a few key issues in the physical interpretation of  $\kappa_P$ .

- (i) Since the topological charge  $Q$  and sum  $C_0$  are independent of renormalization scale,  $\kappa_P$  must be independent of the renormalization scale  $\mu$ .
- (ii) This means that the scale dependence of  $Z_{P \times S}(\mu)$  must cancel off that of  $Z_{P \times P}(\mu)$ .
- (iii) It would be nice to cross-check this property of  $\kappa_P$  in the RI-MOM scheme [53] and in the RI-SMOM scheme [54].

## VIII. CONCLUSION

We study general properties of the eigenvalue spectrum of Dirac operators in the staggered fermion formalism. As an example, we use the Dirac operator for HYP staggered

quarks. In Sec. V, we introduce a new chirality operator  $\Gamma_5$  and a new shift operator  $\Xi_5$  and prove that they respect the continuum recursion relationships, as given in Eqs. (53)–(56) and Eqs. (60)–(61). Using these operators with nice chiral properties, we find that the leakage pattern of  $|\Gamma_5|_{\ell,m}^{-j,m}$  is related to that of  $|\Xi_5|_{\ell,m'}^{j,m}$  through the Ward identity of the conserved  $U(1)_A$  symmetry.

We find that the leakage pattern of  $\Gamma_5$  and  $\Xi_5$  for the zero modes is quite different from that for the nonzero modes. This difference in leakage pattern allows us to distinguish the zero modes from the nonzero modes even though we do not know *a priori* about the topological charge. We find that using the leakage pattern of  $\Gamma_5$  and  $\Xi_5$ , one can determine the topological charge as reliably as when using standard field theoretical methods such as the cooling method.

We use a machine learning (ML) technique to check the universality of this leakage pattern over the entire ensemble of gauge configurations (refer to Table I). Our best-trained deep learning model identifies the quartet of nonzero modes with 98.7(34)% accuracy using a single normal gauge configuration. Choosing the highest probability prediction of the ML and comparing the prediction with the known answer, we find that the ML can identify all quartet groups on an eigenvalue spectrum correctly. In addition, the ML technique detects wrong answers resulting from human input mistakes since the ML prediction disagrees with a wrong answer by giving the prediction with low accuracy ( $< 50\%$ ). This reassures us that the ML technique is highly reliable at identifying anomalous gauge configurations with defects such as violation of the index theorem and ghost eigenmodes.

Once we identify the zero modes, it is also possible to determine the ratio of renormalization factors  $\kappa_P = Z_{P \times S}(\mu)/Z_{P \times P}(\mu)$  from the chirality measurement of  $\Gamma_5$ .

The leakage pattern is a new concept introduced in this paper. It can be used to study the low lying eigenvalue spectrum of staggered Dirac operators systematically. It helps us understand how to extract the taste symmetry and chiral symmetry from the staggered eigenvalue spectrum. The leakage pattern will help us to dig out related physics more efficiently, such as topological charge, index theorem, Banks–Casher relation, and nonperturbative renormalization.

## ACKNOWLEDGMENTS

We would like to express sincere gratitude to Eduardo Follana for helpful discussion and providing his code to cross-check results of our code. We also thank Jon A. Bailey for helpful comments on the manuscript. We also thank Stephen R. Sharpe for helpful discussion. The research of W. Lee is supported by the Mid-Career Research Program (Grant No. NRF-2019R1A2C2085685) of the NRF grant funded by the Korean government (MOE). This work was supported by Seoul National University Research Grant

No. 0409-20190221. W. Lee would like to acknowledge the support from the KISTI supercomputing center through the strategic support program for the supercomputing application research [No. KSC-2016-C3-0072, No. KSC-2017-G2-0009, No. KSC-2017-G2-0014, No. KSC-2018-G2-0004, No. KSC-2018-CHA-0010, No. KSC-2018-CHA-0043]. Computations were carried out in part on the DAVID supercomputer at Seoul National University.

## APPENDIX A: LANCZOS ALGORITHM

Lanczos is a numerical algorithm for calculating eigenvalues and eigenvectors of a Hermitian matrix [21]. It transforms an  $n \times n$  Hermitian matrix  $H$  to a tridiagonal matrix  $T$  through a unitary transformation  $Q$ , which is represented by

$$T = Q^\dagger H Q. \quad (\text{A1})$$

Here columns of  $Q$  are composed of basis vectors of the  $n$ th Krylov subspace  $\mathcal{K}_n(H, b)$  generated by  $H$  and a starting vector  $b$  of our choice. Each iteration of Lanczos computes a column of  $Q$  and  $T$  in sequence. At the end, diagonalizing the tridiagonal matrix  $T$  yields eigenvalues and eigenvectors of  $H$ .

In principle, Lanczos is a direct method that takes  $n$  iterations to construct the  $n \times n$  tridiagonal matrix  $T$ . However, since these columns of  $T$  are computed in order, a sequence of  $m < n$  iterations also constructs an  $m \times m$  tridiagonal matrix  $T'$  which is a submatrix of  $T$ . In practice, the real benefit of Lanczos is that eigenvalues of  $T'$  approximate eigenvalues of  $T$ . As iteration continues, and the size of the submatrix  $T'$  increases, eigenvalues of  $T'$  converge to eigenvalues of  $T$ . The convergence behavior is somewhat complicated. The eigenvalues converge to the largest, the smallest, or the most sparse eigenvalue first. The speed of convergence depends on the density of eigenvalues. The less dense, the faster the convergence.

In this paper, we make use of two popular improvement techniques of Lanczos: (1) implicit restart [55], and (2) polynomial acceleration with Chebyshev polynomials [56]. The implicit restart method gets rid of converged eigenvalues in the middle of the Lanczos iteration. It takes effect as if we restarted the Lanczos with a shifted matrix  $H'$  given by

$$H' \equiv H - \sum_i \lambda_i \mathcal{I}, \quad (\text{A2})$$

where  $\lambda_i$  are eigenvalues we want to remove. Then  $H'$  is still Hermitian but does not have such eigenvalues  $\lambda_i$ . Hence, Lanczos with  $H'$  converges to remaining eigenvalues faster. The implicit restarting procedure gives us a new submatrix, which has a dimension  $((m - r) \times (m - r))$  reduced by the number of eigenvalues we have removed ( $r$ ). Then we

iterate Lanczos  $r$  times to refill the submatrix and restore the structure of the  $m \times m$  matrix. We repeat the implicit restart to obtain a new submatrix of dimension  $(m - r) \times (m - r)$ , and so on. This procedure allows us to control the size of the submatrix, the computational cost, and the memory usage, while the submatrix  $T'$  contains  $(m - r)$  eigenmodes that are more precise (much closer to the true eigenmodes of the full matrix  $H$ ) for each iteration.

A polynomial operation on a matrix changes the eigenvalue spectrum accordingly while retaining the eigenvectors. Since the polynomial of a Hermitian matrix is also Hermitian, Lanczos is still available to calculate its eigenvalues and eigenvectors. By choosing a proper polynomial, one can manipulate the density of the eigenvalue spectrum so that the convergence to the desired eigenvalues is accelerated. A Chebyshev polynomial is a popular choice for this purpose. Using the Chebyshev polynomial, we want to map the first region of eigenmodes of no interest to  $[-1, 1]$  and map the second region of eigenmodes of our interest to  $[-\infty, -1]$ . In the interval  $[-1, 1]$ , the eigenvalues are dense enough that Lanczos does not converge. In addition, the Chebyshev polynomial rapidly changes in the second region so that the density of eigenmodes is low enough to more quickly accelerate the convergence of Lanczos. Here we apply the Chebyshev polynomial for  $D_s^\dagger D_s$ , whose eigenvalues are  $\lambda^2 \geq 0$ . We set the lower bound of the first region to a value somewhat greater than the largest eigenvalue of interest. This strategy will not only suppress high unwanted eigenmodes, but also accelerate the speed of Lanczos for the low eigenmodes of interest.

Numerical stability is essential for the Lanczos algorithm. Each Lanczos iteration generates Lanczos vectors, which are column vectors of the unitary matrix  $Q$  in Eq. (A1). After several iterations, however, Lanczos vectors lose their orthogonality due to gradual loss of numerical precision. If not addressed, this loss would induce spurious ghost eigenvalues [57]. A straightforward prescription to solve the problem is performing a reorthogonalization for every calculation of Lanczos vectors. There are also alternative approaches to eliminate the ghost eigenvalues without reorthogonalization, such as the Cullum-Willoughby method [58,59]. Here we choose the first solution and perform the full reorthogonalization for each Lanczos iteration.

For a large scale simulation using Lanczos, multigrid Lanczos [60] and block Lanczos [61] are available. Multigrid Lanczos is also based on the implicit restart and Chebyshev acceleration. In addition, multigrid Lanczos reduces the memory requirement significantly by compressing the eigenvectors using their local coherence [62]. A spatially-blocked deflation subspace is constructed from some of the lowest eigenvectors of the Dirac operator. Then the coherence of eigenvectors allows us to represent other eigenvectors on this subspace and to run Lanczos with much less memory. Meanwhile, block Lanczos utilizes the

split grid method [61]. This algorithm deals with multiple starting vectors for Lanczos, where the split grid method divides the domain of the Dirac operator application into multiple smaller domains so that each partial domain runs in parallel on a partial grid (lattice) with a lower surface to volume ratio than that of the full grid. Hence, one can optimize the off-node communication by adjusting the block (grid) size. This approach would give a significant speed-up compared with our method. We plan to implement multigrid Lanczos and block Lanczos in the near future.

## APPENDIX B: EVEN-ODD PRECONDITIONING AND PHASE AMBIGUITY

Even-odd preconditioning reorders a fermion field  $\chi(x)$  so that even site fermion fields are obtained first, and odd site fermion fields are obtained from them:

$$\chi(x) = \begin{pmatrix} \chi_e \\ \chi_o \end{pmatrix}, \quad (\text{B1})$$

where  $\chi_e$  ( $\chi_o$ ) is the fermion field collection on even (odd) sites. On this basis, the massless staggered Dirac operator  $D_s$  can be represented as a block matrix:

$$D_s = \begin{pmatrix} 0 & D_{eo} \\ D_{oe} & 0 \end{pmatrix}, \quad (\text{B2})$$

where  $D_{oe}$  ( $D_{eo}$ ) relates even (odd) site fermion fields to odd (even) site fermion fields. Since  $D_s^\dagger = -D_s$ , we also find that  $D_{oe}^\dagger = -D_{eo}$  and  $D_{eo}^\dagger = -D_{oe}$ .

On this basis,  $D_s^\dagger D_s$  is expressed as

$$D_s^\dagger D_s = \begin{pmatrix} 0 & -D_{eo} \\ -D_{oe} & 0 \end{pmatrix} \begin{pmatrix} 0 & D_{eo} \\ D_{oe} & 0 \end{pmatrix} \quad (\text{B3})$$

$$= \begin{pmatrix} -D_{eo}D_{oe} & 0 \\ 0 & -D_{oe}D_{eo} \end{pmatrix}. \quad (\text{B4})$$

Hence, the eigenvalue equation of  $D_s^\dagger D_s$  [Eq. (29)] can be divided into two eigenvalue equations as follows,

$$-D_{eo}D_{oe}|g_e\rangle = \lambda^2|g_e\rangle, \quad (\text{B5})$$

$$-D_{oe}D_{eo}|g_o\rangle = \lambda^2|g_o\rangle, \quad (\text{B6})$$

where  $|g_{e(o)}\rangle$  is the collection of even (odd) site components of  $|g_{j_2}^s\rangle$ . Here we omit the superscript  $s$  and the subscript  $\lambda^2$  for notational simplicity. Now let us multiply  $D_{oe}$  from the left on both sides of Eq. (B5). Then we find that

$$-D_{oe}D_{eo}(D_{oe}|g_e\rangle) = \lambda^2(D_{oe}|g_e\rangle), \quad (\text{B7})$$

which is identical to Eq. (B6). Hence, we find that  $|g_o\rangle = \eta D_{oe}|g_e\rangle$  where  $\eta = re^{i\alpha}$  is an arbitrary complex number with  $r > 0$  and  $0 \leq \alpha < 2\pi$ . Here  $r$  represents the scaling behavior and  $\alpha$  represents a random phase. Since  $-D_{eo}D_{oe}(=D_{oe}^\dagger D_{oe})$  is Hermitian and positive semidefinite, one can solve Eq. (B5) using the Lanczos algorithm introduced in Appendix A. From the result for  $|g_e\rangle$ , it is straightforward to obtain the eigenvector  $|g_{\lambda^2}^s\rangle$  of Eq. (29) since

$$|g_{\lambda^2}^s\rangle = \begin{pmatrix} |g_e\rangle \\ \eta D_{oe}|g_e\rangle \end{pmatrix}. \quad (\text{B8})$$

Now we apply the projection operator  $P_+$ , defined in Eq. (32), to  $|g_{\lambda^2}^s\rangle$ . Using Eq. (B5), we find that

$$\begin{aligned} |\chi_+\rangle &= P_+|g_{\lambda^2}^s\rangle = \begin{pmatrix} i\lambda & D_{eo} \\ D_{oe} & i\lambda \end{pmatrix} \begin{pmatrix} |g_e\rangle \\ \eta D_{oe}|g_e\rangle \end{pmatrix} \\ &= (1+i\eta\lambda) \begin{pmatrix} i\lambda|g_e\rangle \\ D_{oe}|g_e\rangle \end{pmatrix}. \end{aligned} \quad (\text{B9})$$

Similarly, for the projection operator  $P_-$ , defined in Eq. (33), we find that

$$|\chi_-\rangle = P_-|g_{\lambda^2}^s\rangle = (1-i\eta\lambda) \begin{pmatrix} -i\lambda|g_e\rangle \\ D_{oe}|g_e\rangle \end{pmatrix}. \quad (\text{B10})$$

Since  $\eta$  only appears in the overall factor for both cases, it gives only the relative phase difference between the normalized eigenvectors  $|f_{\pm\lambda}^s\rangle$  defined in Eqs. (36) and (37).

We can proceed further to obtain the eigenvectors  $|f_{\pm\lambda}^s\rangle$ . The norm of  $|\chi_+\rangle$  is given by

$$\langle\chi_+|\chi_+\rangle = [(1-i\eta^*\lambda)(1+i\eta\lambda)] \cdot 2\lambda^2\langle g_e|g_e\rangle. \quad (\text{B11})$$

Hence,  $|f_{+\lambda}^s\rangle$  is

$$|f_{+\lambda}^s\rangle = \frac{1}{N} \sqrt{\frac{1+i\eta\lambda}{1-i\eta^*\lambda}} \begin{pmatrix} i\lambda|g_e\rangle \\ D_{oe}|g_e\rangle \end{pmatrix}, \quad (\text{B12})$$

where

$$N \equiv \sqrt{2\lambda^2\langle g_e|g_e\rangle}. \quad (\text{B13})$$

Similarly,

$$|f_{-\lambda}^s\rangle = \frac{1}{N} \sqrt{\frac{1-i\eta\lambda}{1+i\eta^*\lambda}} \begin{pmatrix} -i\lambda|g_e\rangle \\ D_{oe}|g_e\rangle \end{pmatrix}. \quad (\text{B14})$$

These results for  $|f_{\pm\lambda}^s\rangle$  indicate that the phase difference  $\theta$  for the  $\Gamma_\epsilon$  transformation defined in Eq. (47) depends on the value of  $\eta$ .

In our numerical study, we set  $\eta$  to  $\eta = re^{i\alpha} = 1$ :  $r = 1$  and  $\alpha = 0$ . Hence, the relative random phase between  $|f_{\pm\lambda}^s\rangle$  states is removed by hand. Therefore, our value of  $\theta$  defined in Eq. (47) includes a bias from our choice of  $\eta = 1$ .

For  $\eta = 1$  (our choice),  $\Gamma_\epsilon|f_{+\lambda}^s\rangle$  is

$$\Gamma_\epsilon|f_{+\lambda}^s\rangle = \frac{1}{N} \sqrt{\frac{1+i\lambda}{1-i\lambda}} \begin{pmatrix} i\lambda|g_e\rangle \\ -D_{oe}|g_e\rangle \end{pmatrix}, \quad (\text{B15})$$

while  $|f_{-\lambda}^s\rangle$  is

$$|f_{-\lambda}^s\rangle = \frac{1}{N} \sqrt{\frac{1-i\lambda}{1+i\lambda}} \begin{pmatrix} -i\lambda|g_e\rangle \\ D_{oe}|g_e\rangle \end{pmatrix}. \quad (\text{B16})$$

Then we obtain  $e^{i\theta}$  from the following matrix element,

$$\begin{aligned} \langle f_{-\lambda}^s|\Gamma_\epsilon|f_{+\lambda}^s\rangle &= \frac{1}{N^2} \sqrt{\left(\frac{1-i\lambda}{1+i\lambda}\right)^* \frac{1+i\lambda}{1-i\lambda}} \cdot (-N^2) \\ &= -\frac{1+i\lambda}{1-i\lambda} \\ &= e^{i(\pi+2\beta)} = e^{i\theta}, \end{aligned} \quad (\text{B17})$$

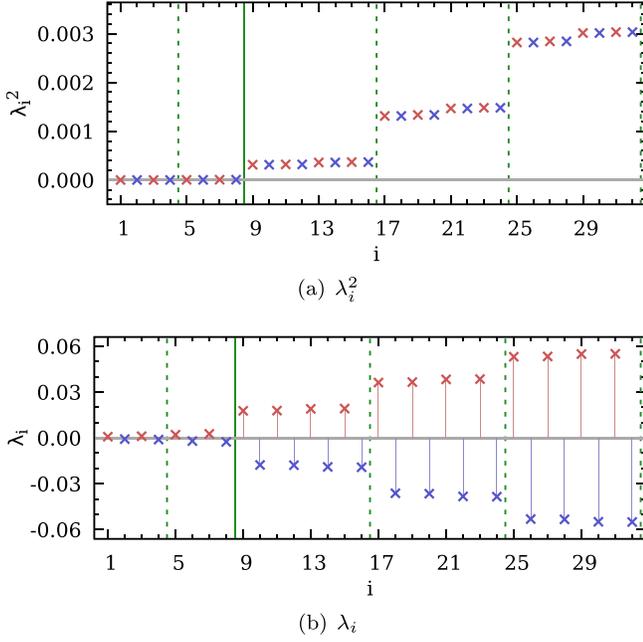
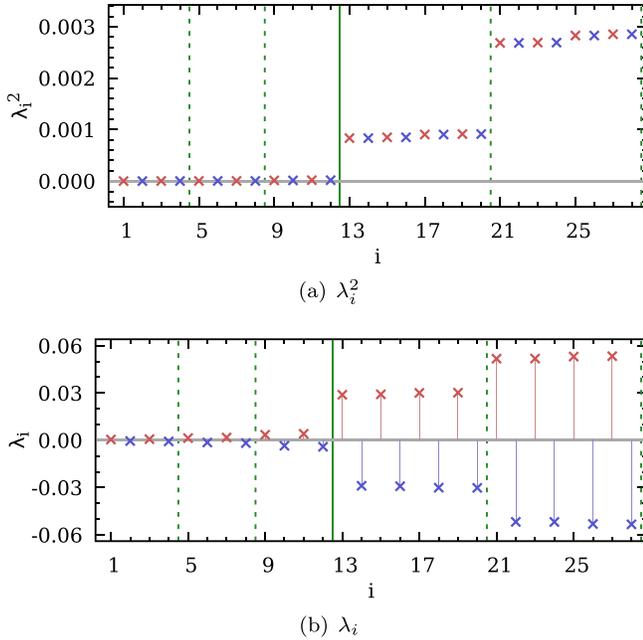
where  $\beta \equiv \arctan(\lambda)$ . From Eqs. (47) and (B17), we find that

$$\theta = \pi + 2\beta. \quad (\text{B18})$$

In Fig. 3, we show the measurements of the phase  $\theta$  for hundreds of eigenvectors on a gauge configuration with  $Q = -1$ . The results for  $\theta$  are consistent with our theoretical prediction Eq. (B18) within numerical precision.

### APPENDIX C: EIGENVALUE SPECTRUM FOR $Q = -2$ AND $Q = -3$

In Figs. 11 and 12, we present examples of the eigenvalue spectrum for  $Q = -2$  and  $Q = -3$ , respectively. Figures 11(a) and 12(a) show eigenvalues  $\lambda^2$  for the eigenvectors  $|g_{\lambda^2}^s\rangle$  defined in Eq. (29). In Fig. 11, we find two sets of four-fold degenerate eigenstates, corresponding to  $\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$  and  $\{\lambda_5, \lambda_6, \lambda_7, \lambda_8\}$ . Each set of four eigenvalues indicates a quartet of would-be zero modes. The number of quartets is related to the topological charge  $Q = -2$  by the index theorem of Eq. (23) when all would-be zero modes have the same chirality ( $n_- = 0$  and  $n_+ = 2$ ). Apart from the would-be zero modes, we observe that nonzero modes are eight-fold degenerate, as in the cases  $Q = 0$  (Fig. 1) and  $Q = -1$  (Fig. 2).

FIG. 11. The same as Fig. 1 except for  $Q = -2$ .FIG. 12. The same as Fig. 1 except for  $Q = -3$ .

Similarly, in Fig. 12, we find three quartets of would-be zero modes with  $n_- = 0$  and  $n_+ = 3$  ( $Q = -3$ ):  $\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$ ,  $\{\lambda_5, \lambda_6, \lambda_7, \lambda_8\}$ , and  $\{\lambda_9, \lambda_{10}, \lambda_{11}, \lambda_{12}\}$ . Because the number of quartets equals the absolute value of the topological charge  $|Q| = 3$ , it is possible to deduce that all the would-be zero modes have the same chirality in accordance with the index theorem of Eq. (23). For nonzero modes, we observe the pattern of eight-fold degeneracy as

in other examples for  $Q = 0$  in Fig. 1,  $Q = -1$  in Fig. 2, and  $Q = -2$  in Fig. 11.

#### APPENDIX D: COMPARISON OF THE GOLTERMAN AND KLUBERG-STERN METHODS

Using the Golterman method, the chirality operator is defined as follows.

$$\begin{aligned} \mathcal{O}_{\gamma_5 \otimes 1}^{\text{Gol}}(x) &= \sum_A \bar{\chi}(x_A) [\gamma_5 \otimes 1]^{\text{Gol}} \chi(x_A) \\ &= \sum_A \bar{\chi}(x_A) [\rho_{\gamma_5 \otimes 1}(A) M_{\gamma_5 \otimes 1}] \chi(x_A) \end{aligned} \quad (\text{D1})$$

where the coordinate  $x_A = 2x + A$ ,  $x$  is a coordinate of the hypercube, and  $A$  is a hypercubic vector with  $A_\mu \in \{0, 1\}$  with  $\mu = 1, 2, 3, 4$ .  $M_{\gamma_5 \otimes 1}$  is defined as

$$M_{\gamma_5 \otimes 1} \chi(x_A) \equiv \frac{1}{4!} \sum_{\{p_\mu\} \in \mathcal{P}} \tilde{D}_{p_4} \tilde{D}_{p_3} \tilde{D}_{p_2} \tilde{D}_{p_1} \chi(x_A) \quad (\text{D2})$$

where  $p_1 \neq p_2 \neq p_3 \neq p_4$ , and  $\mathcal{P}$  is the set of all permutations of  $\{1, 2, 3, 4\}$ . The symmetric shift operator  $\tilde{D}_\mu$  is defined as

$$\tilde{D}_\mu \chi(y) = \frac{1}{2} [V_\mu(y) \chi(y + \hat{\mu}) + V_\mu^\dagger(y - \hat{\mu}) \chi(y - \hat{\mu})] \quad (\text{D3})$$

where  $\hat{\mu}$  is a unit vector in the  $\mu$  direction of Euclidean space.  $V_\mu(y)$  is a (smeared) gauge link used in Eq. (40).  $\rho_{\gamma_5 \otimes 1}$  is defined as

$$\rho_{\gamma_5 \otimes 1}(A) = \frac{1}{4} \text{Tr}(\gamma_A^\dagger \gamma_5 \gamma_{\bar{A}}) = (-1)^{A_1 + A_3} \quad (\text{D4})$$

where  $\bar{A}_\mu = (A_\mu + 1) \bmod 2$ . In the Golterman method, the chirality operator  $[\gamma_5 \otimes 1]^{\text{Gol}}$  connects a site  $A = (0, 0, 0, 0)$  of  $\bar{\chi}$  to the 16 sites  $B = (1, 1, 1, 1), (-1, 1, 1, 1), \dots, (-1, -1, -1, -1)$  of  $\chi$ . As a consequence,

$$([\gamma_5 \otimes 1]^{\text{Gol}})^2 \neq 1 \quad (\text{D5})$$

unlike the continuum chirality operator which satisfies  $[\gamma_5 \otimes 1]^2 = [1 \otimes 1] = 1$ .

Using the definition of bilinear operators in Eq. (38), obtained with the Kluberg-Stern method, we define the chirality operator as

$$\begin{aligned} \mathcal{O}_{\gamma_5 \times 1}^{\text{Klu}}(x) &\equiv \sum_A \bar{\chi}(x_A) [\gamma_5 \otimes 1]_{AA}^{\text{Klu}} \chi(x_A) \\ &= \sum_A \bar{\chi}_a(x_A) (\gamma_5 \otimes 1)_{A\bar{A}} U(x_A, x_{\bar{A}})_{ab} \chi_b(x_{\bar{A}}). \end{aligned} \quad (\text{D6})$$

Using the definition in Eq. (39) with  $\gamma_S = \gamma_5$  and  $\xi_T = 1$ , we find

$$\begin{aligned} \overline{(\gamma_5 \otimes 1)_{A\bar{A}}} &\equiv \frac{1}{4} \text{Tr}(\gamma_A^\dagger \gamma_5 \gamma_{\bar{A}} 1^\dagger) \\ &= \rho_{\gamma_5 \otimes 1}(A) = (-1)^{A_1 + A_3} \end{aligned} \quad (\text{D7})$$

Hence, the phase of the Kluberg-Stern operator is identical to that of the Golterman operator, which is in general true for the whole set of bilinear operators. We have freedom to choose  $U(x_A, x_{\bar{A}})_{ab}$  to make the chirality operator gauge-invariant. Here, we set  $U(x_A, x_{\bar{A}})_{ab}$  to

$$\begin{aligned} U(x_A, x_{\bar{A}}) &\equiv \mathbb{P}_{SU(3)} \left[ \sum_{p \in \mathcal{C}} V(x_A, x_{p_1}) V(x_{p_1}, x_{p_2}) \right. \\ &\quad \left. \cdots V(x_{p_n}, x_{\bar{A}}) \right] \end{aligned} \quad (\text{D8})$$

where  $\mathbb{P}_{SU(3)}$  represents the  $SU(3)$  projection, and  $\mathcal{C}$  represents the complete set of shortest paths from  $x_A$  to  $x_{\bar{A}}$ . Here, the  $SU(3)$  projection is crucial to make the chirality operator satisfy the continuum recursion relation:

$$([\gamma_5 \otimes 1]^{\text{Klu}})^2 = [1 \otimes 1] = 1 \quad (\text{D9})$$

A rigorous proof of Eq. (D9) is given in Appendix E, Theorem E. 1. In the Kluberg-Stern method, the chirality operator  $[\gamma_5 \otimes 1]^{\text{Klu}}$  connects a site  $A = (0, 0, 0, 0)$  of  $\tilde{\chi}$  to a single site  $\bar{A} = (1, 1, 1, 1)$ , which makes it possible to satisfy the recursion relation of Eq. (D9).

The Kluberg-Stern operator without the  $SU(3)$  projection  $\mathbb{P}_{SU(3)}$  contains the Golterman operator as a leading term:

$$\mathcal{O}_{\gamma_5 \times 1}^{\text{Klu}}(x) = \mathcal{O}_{\gamma_5 \times 1}^{\text{Gol}}(x) + \mathcal{O}_{\text{irrel}}(x) \quad (\text{D10})$$

where  $\mathcal{O}_{\text{irrel}}$  represents irrelevant operators of higher dimension. For example,  $\mathcal{O}_{\text{irrel}}$  includes a 4-dimensional operator:

$$\mathcal{O}_{\gamma_\mu \times \xi_{\mu 5}}^{\text{irrel}} = \sum_A \tilde{\chi}(x_A) [\rho_{\gamma_\mu \otimes \xi_{\mu 5}}(A) M_{\gamma_\mu \otimes \xi_{\mu 5}}] \chi(x_A) \quad (\text{D11})$$

$$\rho_{\gamma_\mu \otimes \xi_{\mu 5}}(A) = \frac{1}{4} \text{Tr}(\gamma_A^\dagger \gamma_\mu \gamma_{\bar{A}} \gamma_{\mu 5}^\dagger) \quad (\text{D12})$$

$$\begin{aligned} M_{\gamma_\mu \otimes \xi_{\mu 5}} &= \frac{1}{4!} \sum_{\{p_a\} \in \mathcal{P}_\mu} [D_\mu \tilde{D}_{p_1} \tilde{D}_{p_2} \tilde{D}_{p_3} \\ &\quad + \tilde{D}_{p_1} D_\mu \tilde{D}_{p_2} \tilde{D}_{p_3} + \tilde{D}_{p_1} \tilde{D}_{p_2} D_\mu \tilde{D}_{p_3} \\ &\quad + \tilde{D}_{p_1} \tilde{D}_{p_2} \tilde{D}_{p_3} D_\mu] \end{aligned} \quad (\text{D13})$$

$$D_\mu \chi(y) = \frac{1}{2} [V_\mu(y) \chi(y + \hat{\mu}) - V_\mu^\dagger(y - \hat{\mu}) \chi(y - \hat{\mu})] \quad (\text{D14})$$

where  $p_1 \neq p_2 \neq p_3 \neq \mu$ , and  $\mathcal{P}_\mu$  is the set of all permutations of  $\{p_a | p_a \neq \mu\}$ . Technical details of the derivation of a complete set of irrelevant operators are explained in Ref. [63]. All the irrelevant operators have tastes different from 1 ( $\xi_T \neq 1$ ), and they contain at least one derivative  $D_\mu$ , which leads to higher dimension operators. As a consequence, their contribution to the chirality vanishes in the continuum limit  $a \rightarrow 0$ .

The recursion relation in Eq. (D9) is essential to define the chirality value uniquely for the staggered fermion formulation.

$$([\gamma_5 \otimes 1]^{\text{Klu}})^{2n+1} = [\gamma_5 \otimes 1]^{\text{Klu}} \quad (\text{D15})$$

for all positive  $n \in \mathbb{Z}$ . Hence, in the case of the Kluberg-Stern operators with the  $SU(3)$  projection, we can define the chirality value uniquely without any ambiguity. However, in the case of the Golterman operators, it is not possible to define the chirality value uniquely due to the following ambiguity:

$$\mathcal{O}_{\gamma_5 \times 1}^{\text{Gol}, n}(x) \equiv \sum_A \tilde{\chi}(x_A) ([\gamma_5 \otimes 1]^{\text{Gol}})^{2n+1} \chi(x_A) \quad (\text{D16})$$

$$[\gamma_5 \otimes 1]^{\text{Gol}} \neq ([\gamma_5 \otimes 1]^{\text{Gol}})^{2n+1} \quad (\text{D17})$$

$$\mathcal{O}_{\gamma_5 \times 1}^{\text{Gol}, n} \neq \mathcal{O}_{\gamma_5 \times 1}^{\text{Gol}, m} \quad \text{if } n \neq m. \quad (\text{D18})$$

In addition, the Golterman operator does not satisfy the Ward identity, while the Kluberg-Stern operator respects it,

$$[\gamma_5 \otimes 1]^{\text{Gol}} [1 \otimes \xi_5]^{\text{Gol}} \neq [\gamma_5 \otimes \xi_5] = \Gamma_\epsilon \quad (\text{D19})$$

$$[\gamma_5 \otimes 1]^{\text{Klu}} [1 \otimes \xi_5]^{\text{Klu}} = [\gamma_5 \otimes \xi_5] = \Gamma_\epsilon \quad (\text{D20})$$

A rigorous proof is given in Theorem E. 3. However, in the continuum limit, they converge to a unique value:

$$\begin{aligned} \lim_{a \rightarrow 0} \mathcal{O}_{\gamma_5 \times 1}^{\text{Gol}, n} &= \lim_{a \rightarrow 0} \mathcal{O}_{\gamma_5 \times 1}^{\text{Gol}, 1} \quad \forall \text{ positive } n \in \mathbb{Z} \\ &= \lim_{a \rightarrow 0} \mathcal{O}_{\gamma_5 \times 1}^{\text{Klu}} \end{aligned} \quad (\text{D21})$$

TABLE XI. Comparison between the Golterman and Kluberg-Stern methods. Here Gol (Klu) represents the Golterman (Kluberg-Stern) method. Recursion represents the recursion relationship. Uniqueness represents the uniqueness of the chirality operator value. The  $\circ$  ( $\times$ ) indicates that a given property is (is not) respected by a specific transcription. Ref. represents key equations given for reference.

Property	Gol	Klu	Ref.
Recursion	$\times$	$\circ$	Eqs. (D5) and (D15)
Uniqueness	$\times$	$\circ$	Eqs. (D18) and (D15)
Ward identity	$\times$	$\circ$	Eqs. (D19) and (D20)

since the contribution from all irrelevant operators vanishes in the continuum. We summarize the differences between the Golterman method and the Kluberg-Stern method in Table XI.

### APPENDIX E: RECURSION RELATIONSHIPS FOR CHIRALITY OPERATORS

We define the chirality operator

$$\langle f_\alpha^s | [\gamma_5 \otimes 1] | f_\beta^s \rangle \equiv \sum_x \sum_{A,B} [f_\alpha^s(x_A)]^\dagger \overline{(\gamma_5 \otimes 1)}_{AB} U(x_A, x_B) f_\beta^s(x_B) \quad (\text{E1})$$

$$\overline{(\gamma_5 \otimes \xi_T)}_{AB} = \frac{1}{4} \text{Tr}(\gamma_A^\dagger \gamma_S \gamma_B \gamma_T^\dagger) \quad (\text{E2})$$

$$U(x_A, x_B) = \mathbb{P}_{SU(3)} \left[ \sum_{p \in \mathcal{C}} V(x_A, x_{p_1}) V(x_{p_1}, x_{p_2}) \times V(x_{p_2}, x_{p_3}) V(x_{p_3}, x_B) \right] \quad (\text{E3})$$

First let us prove the following theorem.

#### Theorem E.1.

$$[\gamma_5 \otimes 1][\gamma_5 \otimes 1] = [1 \otimes 1] \quad (\text{E4})$$

*Proof.*—Let us first rewrite  $[\gamma_5 \otimes 1]^2$  as follows,

$$\begin{aligned} [\gamma_5 \otimes 1]_{AC}^2 &= \sum_B \overline{(\gamma_5 \otimes 1)}_{AB} U(x_A, x_B) \cdot \overline{(\gamma_5 \otimes 1)}_{BC} U(x_B, x_C) \\ &= \sum_B [\overline{(\gamma_5 \otimes 1)}_{AB} \overline{(\gamma_5 \otimes 1)}_{BC}] \\ &\quad \cdot [U(x_A, x_B) U(x_B, x_C)] \end{aligned} \quad (\text{E5})$$

We know that

$$\begin{aligned} \overline{(\gamma_5 \otimes 1)}_{AB} &= \frac{1}{4} \text{Tr}(\gamma_A^\dagger \gamma_5 \gamma_B) \\ &= \delta_{B\bar{A}} [\eta_1(A) \eta_2(A) \eta_3(A) \eta_4(A)] \\ &= \delta_{B\bar{A}} \eta_5(A), \end{aligned} \quad (\text{E6})$$

where  $\bar{A}_\mu = (A_\mu + 1) \bmod 2$ , and

$$\eta_\mu(A) = (-1)^{X_\mu}, \quad \text{for } \mu = 1, 2, 3, 4, \quad (\text{E7})$$

$$X_\mu = \sum_{\nu < \mu} A_\nu, \quad (\text{E8})$$

$$\eta_5(A) = \eta_1(A) \eta_2(A) \eta_3(A) \eta_4(A) = (-1)^{A_1 + A_3}. \quad (\text{E9})$$

Similarly, we find that

$$\overline{(\gamma_5 \otimes 1)}_{BC} = \delta_{C\bar{B}} \eta_5(B). \quad (\text{E10})$$

Hence, we can rewrite Eq. (E5) as follows,

$$\begin{aligned} [\gamma_5 \otimes 1]_{AC}^2 &= \sum_B [\delta_{B\bar{A}} \eta_5(A) \delta_{C\bar{B}} \eta_5(B)] \cdot [U(x_A, x_B) U(x_B, x_C)] \\ &= \delta_{AC} [U(x_A, x_{\bar{A}}) U(x_{\bar{A}}, x_A)], \end{aligned} \quad (\text{E11})$$

where we use the helpful identity  $\eta_5(\bar{A}) = \eta_5(A)$ . Thanks to the  $SU(3)$  projection in Eq. (E3),  $U(x_{\bar{A}}, x_A) = [U(x_A, x_{\bar{A}})]^\dagger \in SU(3)$ . Hence,  $[U(x_A, x_{\bar{A}}) U(x_{\bar{A}}, x_A)] = 1$ . Therefore, we can rewrite Eq. (E11) as follows,

$$[\gamma_5 \otimes 1]_{AC}^2 = \delta_{AC} = [1 \otimes 1]_{AC}. \quad (\text{E12})$$

Hence, we have just proven that  $[\gamma_5 \otimes 1]^2 = [1 \otimes 1]$ . (Q.E.D.) ■

Using the results of Eq. (E4), we can prove the recursion relationship as follows,

$$[\gamma_5 \otimes 1]^{2n+1} = ([\gamma_5 \otimes 1]^{2n}) \cdot [\gamma_5 \otimes 1] \quad (\text{E13})$$

$$= ([1 \otimes 1])^n \cdot [\gamma_5 \otimes 1] \quad (\text{E14})$$

$$= [1 \otimes 1] \cdot [\gamma_5 \otimes 1] \quad (\text{E15})$$

$$= [\gamma_5 \otimes 1]. \quad (\text{E16})$$

Using the results of Eq. (E4), we can prove another recursion relationship as follows,

$$[\gamma_5 \otimes 1]^{2n} = ([\gamma_5 \otimes 1]^{2n}) \quad (\text{E17})$$

$$= ([1 \otimes 1])^n \quad (\text{E18})$$

$$= [1 \otimes 1]. \quad (\text{E19})$$

Finally, we can prove the following theorem.

#### Theorem E.2.

$$\left[ \frac{1 + \gamma_5}{2} \otimes 1 \right] \left[ \frac{1 + \gamma_5}{2} \otimes 1 \right] = \left[ \frac{1 + \gamma_5}{2} \otimes 1 \right] \quad (\text{E20})$$

*Proof.*—

$$\begin{aligned} \left[ \frac{1+\gamma_5}{2} \otimes 1 \right]^2 &= \frac{1}{4} ([1 \otimes 1] + [\gamma_5 \otimes 1])^2 \\ &= \frac{1}{4} ([1 \otimes 1] + 2[\gamma_5 \otimes 1] + [\gamma_5 \otimes 1]^2) \\ &= \frac{1}{2} ([1 \otimes 1] + [\gamma_5 \otimes 1]) \\ &= \left[ \frac{1+\gamma_5}{2} \otimes 1 \right]. \end{aligned} \quad (\text{E21})$$

(Q.E.D.) ■

Using Eq. (E20), we can prove that for integer  $n > 0$ ,

$$\left[ \frac{1+\gamma_5}{2} \otimes 1 \right]^n = \left[ \frac{1+\gamma_5}{2} \otimes 1 \right] \quad (\text{E22})$$

by induction.

At this stage, it will be trivial to prove that

$$\left[ \frac{1+\gamma_5}{2} \otimes 1 \right] \left[ \frac{1-\gamma_5}{2} \otimes 1 \right] = 0. \quad (\text{E23})$$

The next two theorems concern the chiral Ward identities.

**Theorem E.3.**

$$[\gamma_5 \otimes \xi_5] = [\gamma_5 \otimes 1][1 \otimes \xi_5] = [1 \otimes \xi_5][\gamma_5 \otimes 1] \quad (\text{E24})$$

*Proof.*—Using the results of Eq. (E6), we find that

$$\overline{(\gamma_5 \otimes 1)}_{AB} = \delta_{B\bar{A}} \eta_5(A), \quad (\text{E25})$$

where  $\bar{A}_\mu = (A_\mu + 1) \bmod 2$ . Let us rewrite  $\overline{(1 \otimes \xi_5)}_{AB}$  as follows,

$$\begin{aligned} \overline{(1 \otimes \xi_5)}_{AB} &= \frac{1}{4} \text{Tr}(\gamma_A^\dagger 1 \gamma_B \gamma_5^\dagger) \\ &= \delta_{B\bar{A}} [\zeta_1(\bar{A}) \zeta_2(\bar{A}) \zeta_3(\bar{A}) \zeta_4(\bar{A})] \\ &= \delta_{B\bar{A}} \zeta_5(A), \end{aligned} \quad (\text{E26})$$

where  $\bar{A}_\mu = (A_\mu + 1) \bmod 2$ , and

$$\zeta_\mu(A) = (-1)^{Y_\mu}, \quad \text{for } \mu = 1, 2, 3, 4, \quad (\text{E27})$$

$$Y_\mu = \sum_{\nu > \mu} A_\nu, \quad (\text{E28})$$

$$\zeta_5(A) = \zeta_1(A) \zeta_2(A) \zeta_3(A) \zeta_4(A) = (-1)^{A_2 + A_4}, \quad (\text{E29})$$

$$\zeta_5(\bar{A}) = \zeta_5(A). \quad (\text{E30})$$

Hence, we find that

$$\begin{aligned} [\gamma_5 \otimes 1][1 \otimes \xi_5]_{AC} &= \sum_B [\gamma_5 \otimes 1]_{AB} [1 \otimes \xi_5]_{BC} \\ &= \sum_B \{ \overline{(\gamma_5 \otimes 1)}_{AB} \overline{(1 \otimes \xi_5)}_{BC} \} [U(x_A, x_B) U(x_B, x_C)] \\ &= \sum_B \{ \delta_{B\bar{A}} \eta_5(A) \delta_{C\bar{B}} \zeta_5(C) \} [U(x_A, x_{\bar{A}}) U(x_{\bar{A}}, x_A)] \\ &= \delta_{AC} \eta_5(A) \zeta_5(A) \\ &= \delta_{AC} \epsilon(A) = [\gamma_5 \otimes \xi_5]_{AC} \end{aligned} \quad (\text{E31})$$

This is a proof of the first part of the theorem. Similarly,

$$\begin{aligned} [1 \otimes \xi_5][\gamma_5 \otimes 1]_{AC} &= \sum_B [1 \otimes \xi_5]_{AB} [\gamma_5 \otimes 1]_{BC} \\ &= \sum_B \{ \overline{(1 \otimes \xi_5)}_{AB} \overline{(\gamma_5 \otimes 1)}_{BC} \} [U(x_A, x_B) U(x_B, x_C)] \\ &= \sum_B \{ \delta_{B\bar{A}} \zeta_5(A) \delta_{C\bar{B}} \eta_5(C) \} [U(x_A, x_{\bar{A}}) U(x_{\bar{A}}, x_A)] \\ &= \delta_{AC} \eta_5(A) \zeta_5(A) \\ &= \delta_{AC} \epsilon(A) = [\gamma_5 \otimes \xi_5]_{AC} \end{aligned} \quad (\text{E32})$$

This is a proof of the second part of the theorem. (Q.E.D.) ■

We prove the Ward identities in Eqs. (63)–(64) as follows.

**Theorem E.4.**

$$[\gamma_5 \otimes \xi_5][\gamma_5 \otimes 1] = [\gamma_5 \otimes 1][\gamma_5 \otimes \xi_5] = [1 \otimes \xi_5], \quad (\text{E33})$$

$$[\gamma_5 \otimes \xi_5][1 \otimes \xi_5] = [1 \otimes \xi_5][\gamma_5 \otimes \xi_5] = [\gamma_5 \otimes 1]. \quad (\text{E34})$$

*Proof.*—Using the results of Theorem E.3, we know the following Ward identity:

$$[\gamma_5 \otimes \xi_5] = [1 \otimes \xi_5][\gamma_5 \otimes 1]. \quad (\text{E35})$$

Let us multiply  $[\gamma_5 \otimes 1]$  on both sides of Eq. (E35). Then,

$$[\gamma_5 \otimes \xi_5][\gamma_5 \otimes 1] = [1 \otimes \xi_5][\gamma_5 \otimes 1]^2 = [1 \otimes \xi_5]. \quad (\text{E36})$$

Here we use the recursion relationship in Theorem E.1. Similarly, from the results of Theorem E.3, we know that

$$[\gamma_5 \otimes \xi_5] = [\gamma_5 \otimes 1][1 \otimes \xi_5]. \quad (\text{E37})$$

Let us multiply  $[\gamma_5 \otimes 1]$  on both sides of Eq. (E37).

$$[\gamma_5 \otimes 1][\gamma_5 \otimes \xi_5] = [\gamma_5 \otimes 1]^2[1 \otimes \xi_5] = [1 \otimes \xi_5]. \quad (\text{E38})$$

Here, we use the recursion relation in Theorem E.1. This completes a proof of the first part of Theorem E.4.

Let us multiply  $[1 \otimes \xi_5]$  on both sides of Eq. (E37).

$$[\gamma_5 \otimes \xi_5][1 \otimes \xi_5] = [\gamma_5 \otimes 1][1 \otimes \xi_5]^2 = [\gamma_5 \otimes 1]. \quad (\text{E39})$$

Similarly, let us multiply  $[1 \otimes \xi_5]$  on both sides of Eq. (E35).

$$[1 \otimes \xi_5][\gamma_5 \otimes \xi_5] = [1 \otimes \xi_5]^2[\gamma_5 \otimes 1] = [\gamma_5 \otimes 1]. \quad (\text{E40})$$

This completes a proof of the second part of Theorem E.4. (Q.E.D.) ■

## APPENDIX F: EXAMPLES FOR THE LEAKAGE PATTERN OF ZERO MODES

Let us begin with the case  $Q = -2$ . In Fig. 13, we show leakage patterns of the chirality operator for the first set of zero modes at  $Q = -2$ . In Fig. 14, we present the leakage patterns of the shift operator for the first set of zero modes at  $Q = -2$ . By comparing Fig. 13 with Fig. 14, we find that the chiral Ward identities of Eqs. (78) and (79) are well respected.

In Fig. 15, we show leakage patterns of the chirality operator for the second set of zero modes at  $Q = -2$ . In Fig. 16, we present the leakage patterns of the shift operator for the second set of zero modes at  $Q = -2$ . By comparing

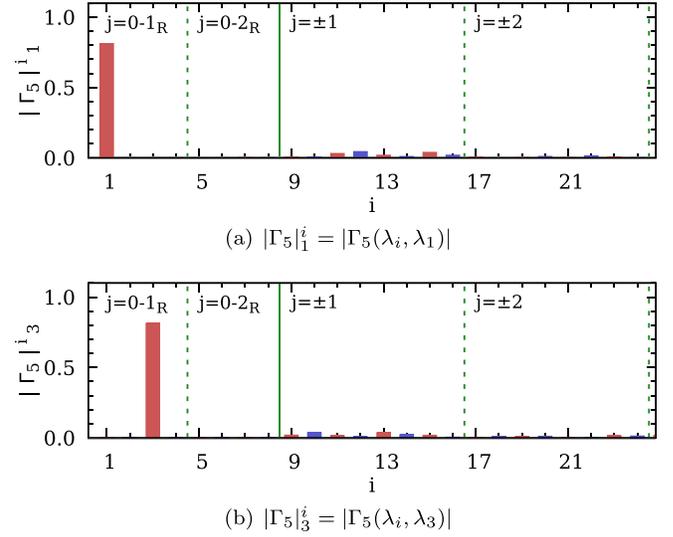


FIG. 13.  $[\gamma_5 \otimes 1]$  leakage pattern for the first quartet of would-be zero modes at  $Q = -2$ .

Fig. 15 with Fig. 16, we find that the chiral Ward identities of Eqs. (78) and (79) are well-preserved.

Now let us consider an example with  $Q = -3$ . The leakage patterns for the first and second sets of zero modes are similar to those at  $Q = -2$ . Hence, we choose the third set of zero modes as our example. In Fig. 17, we show leakage patterns of the chirality operator for the third set of zero modes at  $Q = -3$ . In Fig. 18, we present the leakage pattern of the shift operator for the third set of zero modes at  $Q = -3$ . By comparing Fig. 17 with Fig. 18, we find that the chiral Ward identities of Eqs. (78) and (79) are well-preserved.

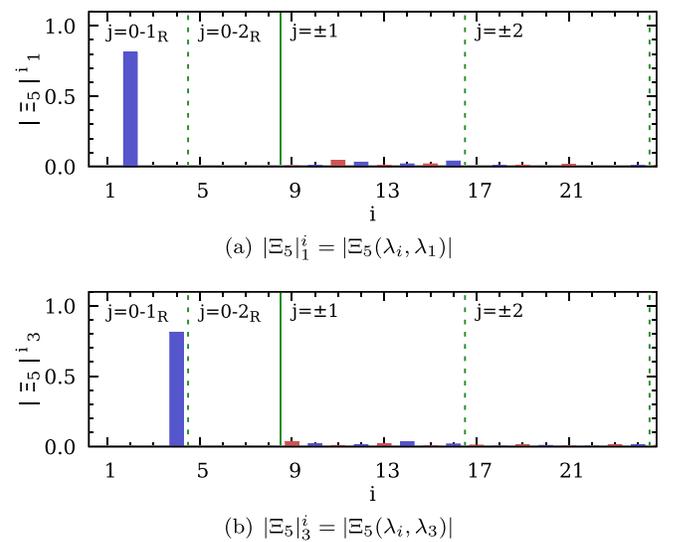


FIG. 14.  $[1 \otimes \xi_5]$  leakage pattern for the first quartet of would-be zero modes at  $Q = -2$ .

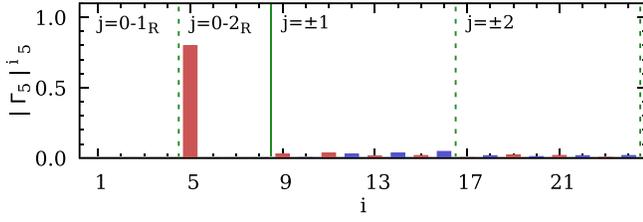
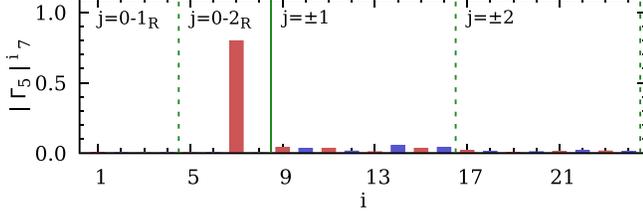

 (a)  $|\Gamma_5|_5^i = |\Gamma_5(\lambda_i, \lambda_5)|$ 

 (b)  $|\Gamma_5|_7^i = |\Gamma_5(\lambda_i, \lambda_7)|$ 

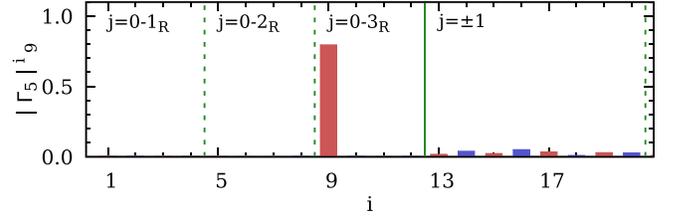
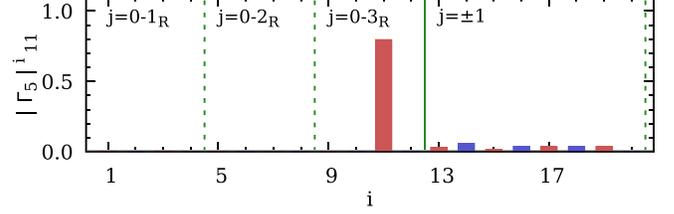
 FIG. 15.  $[\gamma_5 \otimes 1]$  leakage pattern for the second quartet of would-be zero modes at  $Q = -2$ .

 (a)  $|\Gamma_5|_9^i = |\Gamma_5(\lambda_i, \lambda_9)|$ 

 (b)  $|\Gamma_5|_{11}^i = |\Gamma_5(\lambda_i, \lambda_{11})|$ 

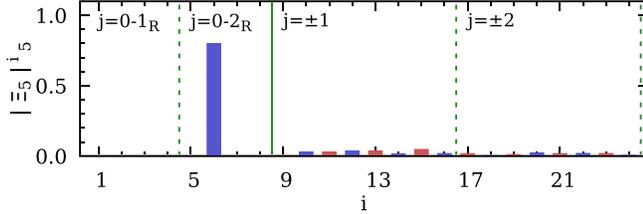
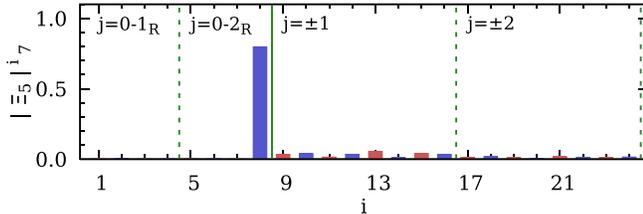
 FIG. 17.  $[\gamma_5 \otimes 1]$  leakage pattern for the third quartet of would-be zero modes at  $Q = -3$ .

 (a)  $|\Xi_5|_5^i = |\Xi_5(\lambda_i, \lambda_5)|$ 

 (b)  $|\Xi_5|_7^i = |\Xi_5(\lambda_i, \lambda_7)|$ 

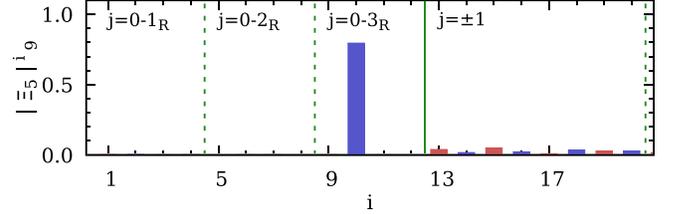
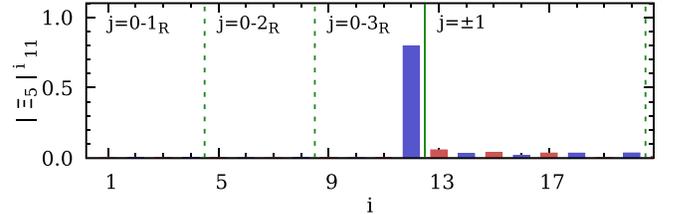
 FIG. 16.  $[1 \otimes \xi_5]$  leakage pattern for the second quartet of would-be zero modes at  $Q = -2$ .

 (a)  $|\Xi_5|_9^i = |\Xi_5(\lambda_i, \lambda_9)|$ 

 (b)  $|\Xi_5|_{11}^i = |\Xi_5(\lambda_i, \lambda_{11})|$ 

 FIG. 18.  $[1 \otimes \xi_5]$  leakage pattern for the third quartet of would-be zero modes at  $Q = -3$ .

## APPENDIX G: EXAMPLES FOR THE LEAKAGE PATTERN OF NONZERO MODES

Let us begin with an example with  $Q = 0$ . Since the gauge configuration with  $Q = 0$  usually has no zero mode ( $n_- = n_+ = 0$ ), it is relatively easy to study nonzero modes. In Fig. 19, we present leakage patterns of the chirality operator  $\Gamma_5 = [\gamma_5 \otimes 1]$  for nonzero modes  $\{\lambda_1, \lambda_3, \lambda_5, \lambda_7\} = \{\lambda_{j,m} | j = +1, m = 1, 2, 3, 4\}$  in the  $j = +1$  quartet when  $Q = 0$ . The results show that the  $\Gamma_5$  leakages for nonzero modes  $\lambda_{+1,m}$  mostly go into their parity partners  $\{\lambda_2, \lambda_4, \lambda_6, \lambda_8\} = \{\lambda_{j,m} | j = -1, m = 1, 2, 3, 4\}$  in the  $j = -1$  quartet. Meanwhile, the leakages

to other quartets such as  $j = \pm 2, \pm 3$  are negligibly small compared to those of the  $j = -1$  quartet elements. This observation is consistent with that for  $Q = -1$  in Fig. 6.

In Fig. 20, we present leakage patterns of the shift operator  $\Xi_5 = [1 \otimes \xi_5]$  for the nonzero modes  $\{\lambda_1, \lambda_3, \lambda_5, \lambda_7\}$  of  $\lambda_{+1,m}$  in the  $j = +1$  quartet when  $Q = 0$ . For the  $\Xi_5$  operator, we find the great part of leakages are from nonzero modes  $\lambda_{+1,m}$  to other elements within the  $j = +1$  quartet. Meanwhile, there are only negligible leakages to parity partner quartet elements ( $j = -1$ ) and other quartets with  $j = \pm 2, \pm 3$ , and so on. This observation corresponds to the case  $Q = -1$  in Fig. 7.

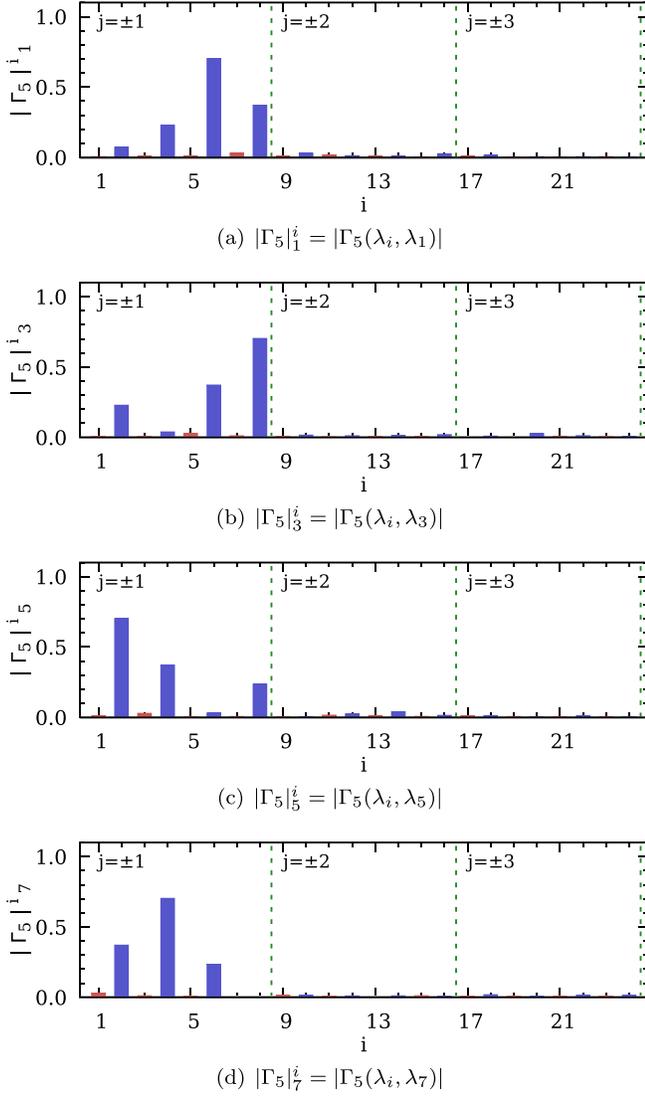


FIG. 19.  $[\gamma_5 \otimes 1]$  leakage pattern for the first quartet of nonzero modes at  $Q = 0$ .

We also find that the leakages of  $\Gamma_5$  in Fig. 19 and  $\Xi_5$  in Fig. 20 are related to each other by the Ward identity of Eq. (95).

In Figs. 21 and 22, we present leakage patterns of the  $\Gamma_5$  and  $\Xi_5$  operators, respectively, for nonzero modes  $\{\lambda_9, \lambda_{11}, \lambda_{13}, \lambda_{15}\} = \{\lambda_{j,m} | j = +2, m = 1, 2, 3, 4\}$  in the  $j = +2$  quartet when  $Q = 0$ . Similar to the above cases for  $j = +1$ ,  $\Gamma_5$  leakages for nonzero modes of  $j = +2$  mostly go to their parity partner quartet elements of  $j = -2$ :  $\{\lambda_{10}, \lambda_{12}, \lambda_{14}, \lambda_{16}\} = \{\lambda_{j,m} | j = -2, m = 1, 2, 3, 4\}$ , and  $\Xi_5$  leakages mostly go to members within the  $j = +2$  quartet:  $\{\lambda_9, \lambda_{11}, \lambda_{13}, \lambda_{15}\}$ . There are only negligible leakages to other quartets for both operators.

Now let us examine the leakage patterns when would-be zero modes exist ( $Q \neq 0$ ). In Figs. 23 and 24, we present leakage patterns of the  $\Gamma_5$  and  $\Xi_5$  operators, respectively, for nonzero modes  $\{\lambda_9, \lambda_{11}, \lambda_{13}, \lambda_{15}\}$  in the  $j = +1$  quartet

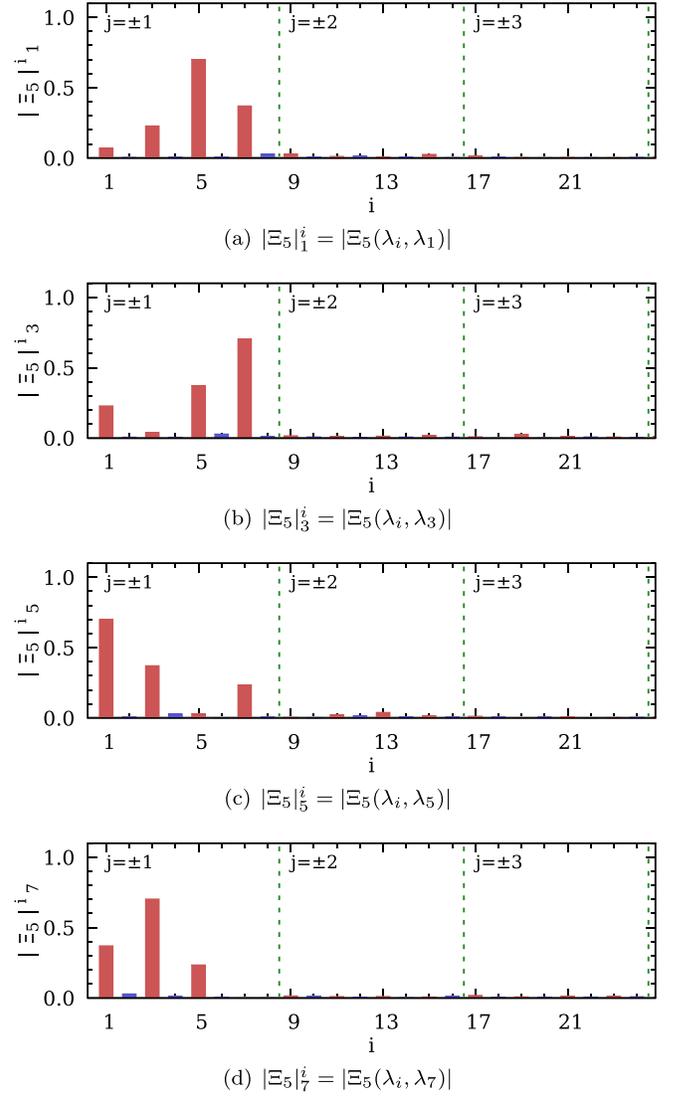
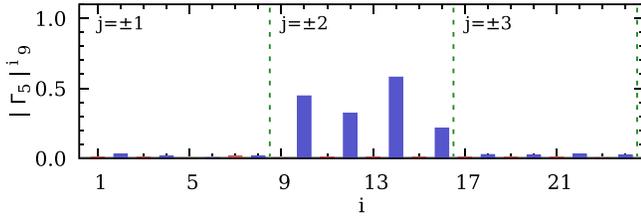


FIG. 20.  $[1 \otimes \xi_5]$  leakage pattern for the first quartet of nonzero modes at  $Q = 0$ .

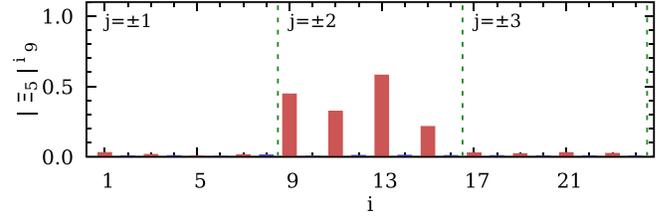
when  $Q = -2$ . There are two quartets of right-handed would-be zero modes where  $j = 0 - 1_R$  and  $0 - 2_R$ , which corresponds to  $n_- = 0$  and  $n_+ = 2$  with  $Q = -2$  by the index theorem ( $Q = n_- - n_+$ ).

As in the cases  $Q = -1$  (Figs. 6 and 7) and  $Q = 0$  (Figs. 19 and 20),  $\Gamma_5$  leakages from nonzero modes of  $j = +1$  mostly go to the parity partner  $j = -1$  quartet, and  $\Xi_5$  leakages from nonzero modes of  $j = +1$  mostly go within the  $j = +1$  quartet itself. Leakages to other nonzero mode quartets and would-be zero mode quartets are negligibly small. We also find that the Ward identity between the two leakage patterns holds.

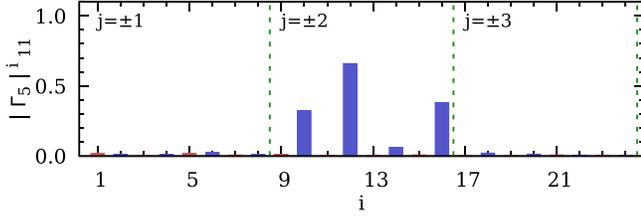
In Figs. 25 and 26, we present leakage patterns for the  $\Gamma_5$  and  $\Xi_5$  operators, respectively, for nonzero modes  $\{\lambda_{13}, \lambda_{15}, \lambda_{17}, \lambda_{19}\}$  in the  $j = +1$  quartet when  $Q = -3$ . Their leakage patterns are also consistent with those for  $Q = 0, -1, -2$  in our previous discussion.



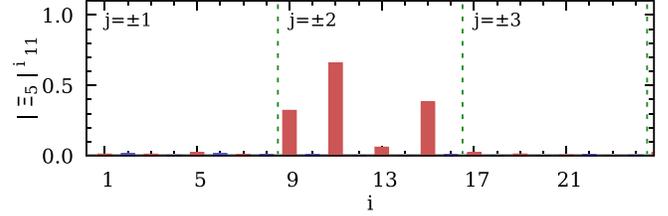
(a)  $|\Gamma_5|_9^i = |\Gamma_5(\lambda_i, \lambda_9)|$



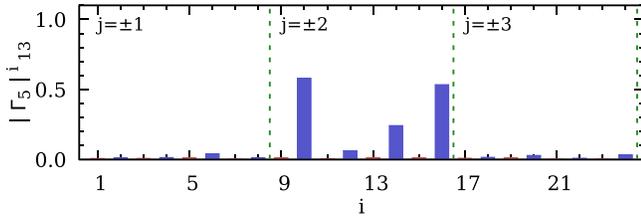
(a)  $|\Xi_5|_9^i = |\Xi_5(\lambda_i, \lambda_9)|$



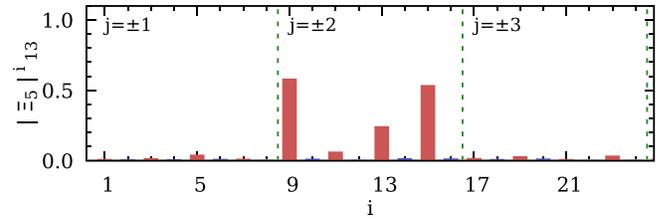
(b)  $|\Gamma_5|_{11}^i = |\Gamma_5(\lambda_i, \lambda_{11})|$



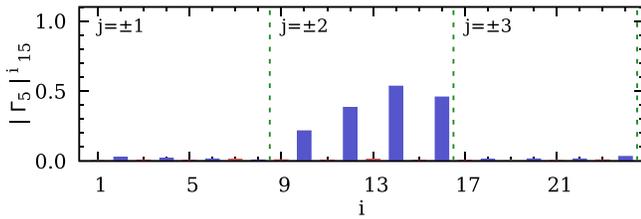
(b)  $|\Xi_5|_{11}^i = |\Xi_5(\lambda_i, \lambda_{11})|$



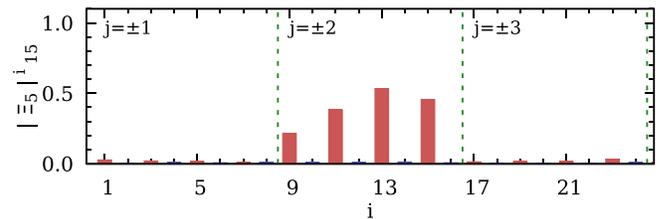
(c)  $|\Gamma_5|_{13}^i = |\Gamma_5(\lambda_i, \lambda_{13})|$



(c)  $|\Xi_5|_{13}^i = |\Xi_5(\lambda_i, \lambda_{13})|$



(d)  $|\Gamma_5|_{15}^i = |\Gamma_5(\lambda_i, \lambda_{15})|$



(d)  $|\Xi_5|_{15}^i = |\Xi_5(\lambda_i, \lambda_{15})|$

FIG. 21.  $[\gamma_5 \otimes 1]$  leakage pattern for the second quartet of nonzero modes at  $Q = 0$ .

FIG. 22.  $[1 \otimes \xi_5]$  leakage pattern for the second quartet of nonzero modes at  $Q = 0$ .

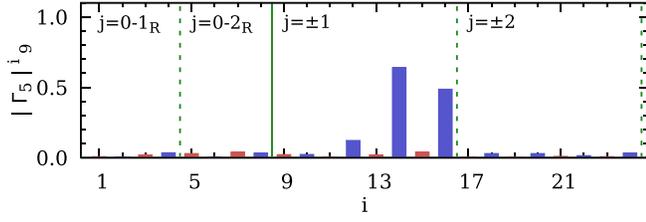
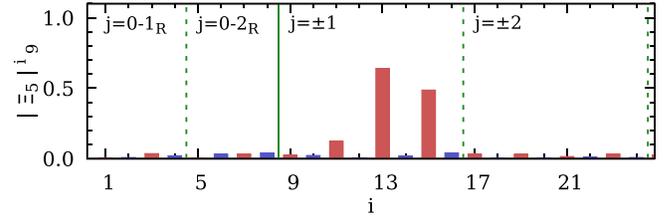
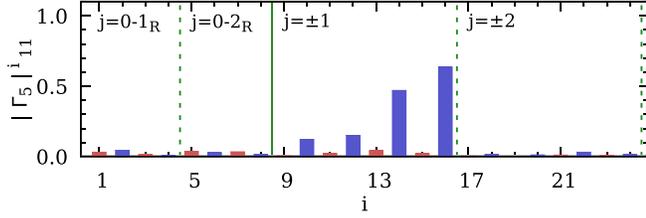
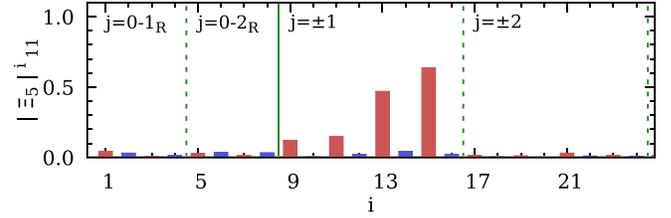
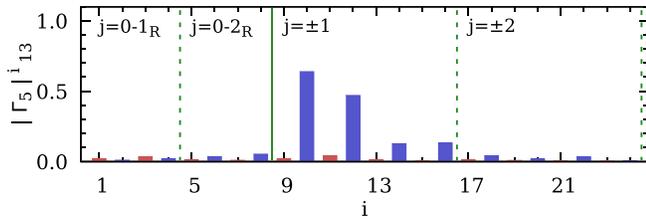
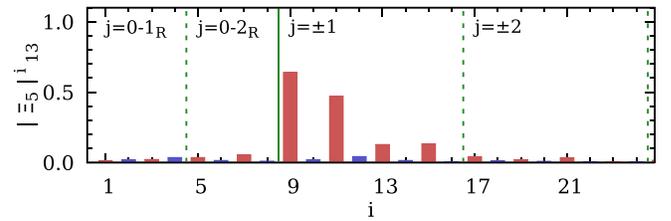
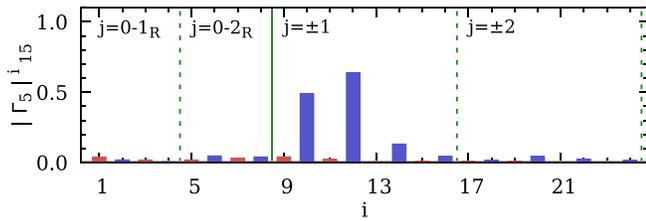
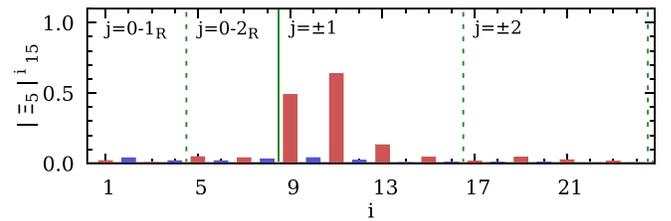
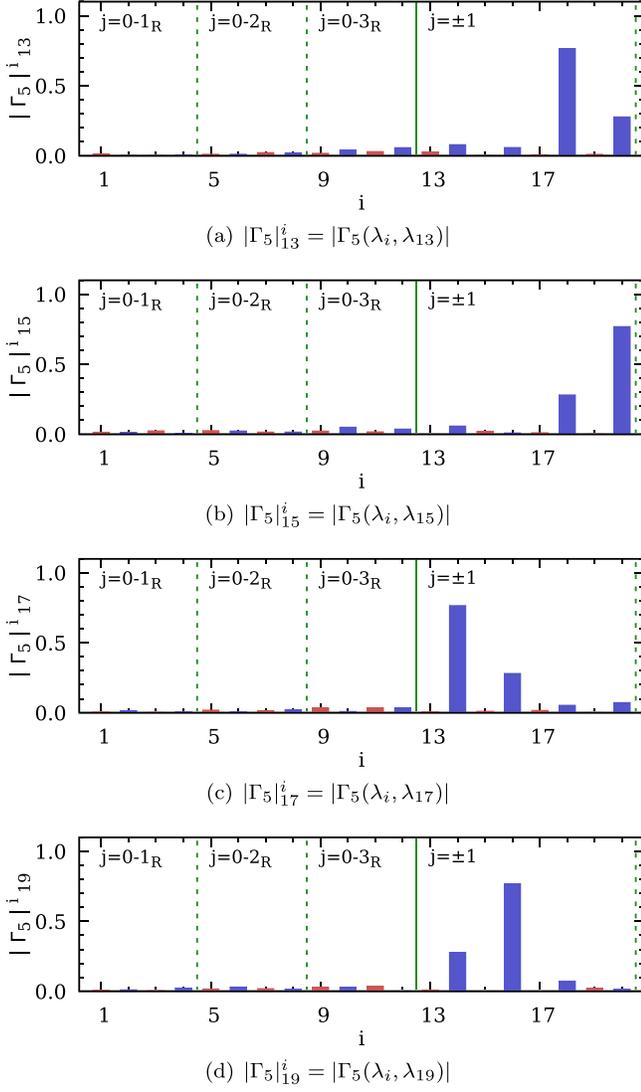
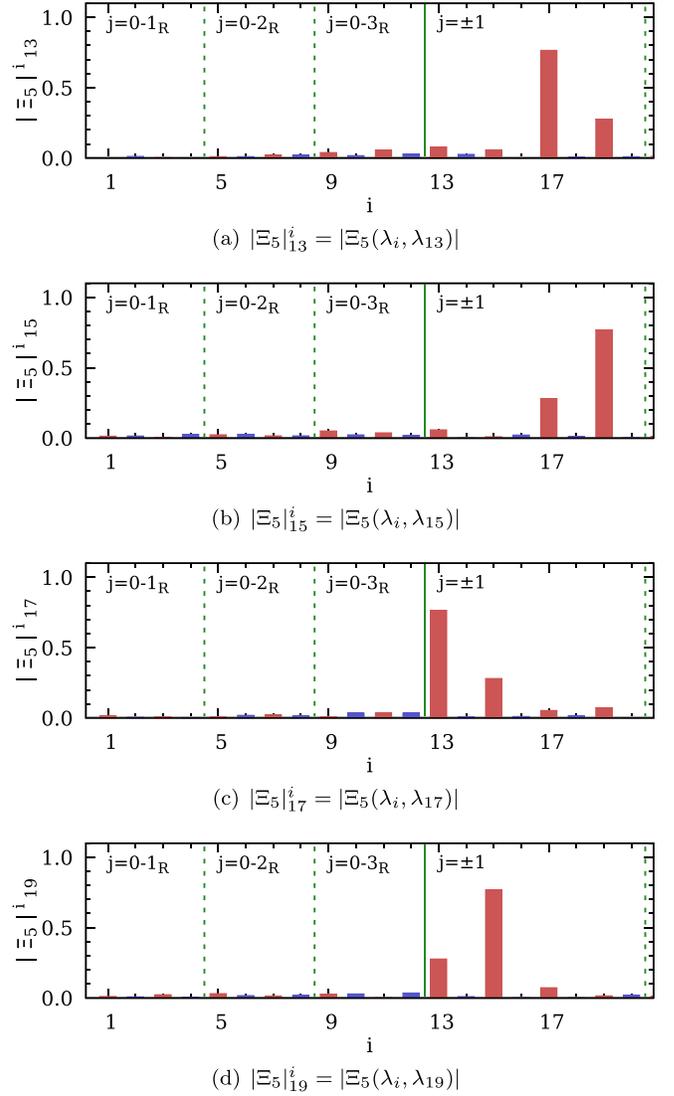

 (a)  $|\Gamma_5|_9^i = |\Gamma_5(\lambda_i, \lambda_9)|$ 

 (a)  $|\Xi_5|_9^i = |\Xi_5(\lambda_i, \lambda_9)|$ 

 (b)  $|\Gamma_5|_{11}^i = |\Gamma_5(\lambda_i, \lambda_{11})|$ 

 (b)  $|\Xi_5|_{11}^i = |\Xi_5(\lambda_i, \lambda_{11})|$ 

 (c)  $|\Gamma_5|_{13}^i = |\Gamma_5(\lambda_i, \lambda_{13})|$ 

 (c)  $|\Xi_5|_{13}^i = |\Xi_5(\lambda_i, \lambda_{13})|$ 

 (d)  $|\Gamma_5|_{15}^i = |\Gamma_5(\lambda_i, \lambda_{15})|$ 

 (d)  $|\Xi_5|_{15}^i = |\Xi_5(\lambda_i, \lambda_{15})|$ 

 FIG. 23.  $[\gamma_5 \otimes 1]$  leakage pattern for the first quartet of nonzero modes at  $Q = -2$ .

 FIG. 24.  $[1 \otimes \xi_5]$  leakage pattern for the first quartet of nonzero modes at  $Q = -2$ .


 FIG. 25.  $[\gamma_5 \otimes 1]$  leakage pattern for the first quartet of nonzero modes at  $Q = -3$ .

 FIG. 26.  $[1 \otimes \xi_5]$  leakage pattern for the first quartet of nonzero modes at  $Q = -3$ .

[1] M. F. Atiyah and I. M. Singer, The index of elliptic operators on compact manifolds, *Bull. Am. Math. Soc.* **69**, 422 (1963).

[2] T. Banks and A. Casher, Chiral symmetry breaking in confining theories, *Nucl. Phys.* **B169**, 103 (1980).

[3] E. V. Shuryak and J. J. M. Verbaarschot, Random matrix theory and spectral sum rules for the Dirac operator in QCD, *Nucl. Phys.* **A560**, 306 (1993).

[4] H. Leutwyler and A. V. Smilga, Spectrum of Dirac operator and role of winding number in QCD, *Phys. Rev. D* **46**, 5607 (1992).

[5] A. Hasenfratz and F. Knechtli, Flavor symmetry and the static potential with hypercubic blocking, *Phys. Rev. D* **64**, 034504 (2001).

[6] W.-J. Lee and S. R. Sharpe, One loop matching coefficients for improved staggered bilinears, *Phys. Rev. D* **66**, 114501 (2002).

[7] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. P. Lepage, J. Shigemitsu, H. Trotter, and K. Wong (HPQCD and UKQCD Collaborations), Highly improved staggered quarks on the lattice, with applications to charm physics, *Phys. Rev. D* **75**, 054502 (2007).

[8] E. Follana, A. Hart, and C. T. H. Davies (HPQCD and UKQCD Collaborations), The Index Theorem and Universality Properties of the Low-Lying Eigenvalues of Improved Staggered Quarks, *Phys. Rev. Lett.* **93**, 241601 (2004).

[9] E. Follana, A. Hart, C. T. H. Davies, and Q. Mason (HPQCD and UKQCD Collaborations), The low-lying

- Dirac spectrum of staggered quarks, *Phys. Rev. D* **72**, 054501 (2005).
- [10] S. Durr, C. Hoelbling, and U. Wenger, Staggered eigenvalue mimicry, *Phys. Rev. D* **70**, 094502 (2004).
- [11] V. Azcoiti, G. Di Carlo, E. Follana, and A. Vaquero, Topological index theorem on the lattice through the spectral flow of staggered fermions, *Phys. Lett. B* **744**, 303 (2015).
- [12] S. Durr, Taste-split staggered actions: Eigenvalues, chiralities and Symanzik improvement, *Phys. Rev. D* **87**, 114501 (2013).
- [13] D. H. Adams, Theoretical Foundation for the Index Theorem on the Lattice with Staggered Fermions, *Phys. Rev. Lett.* **104**, 141602 (2010).
- [14] N. D. Cundy, H. Jeong, and W. Lee, Calculation of strange and light quark condensate using improved staggered fermions and overlap fermions, *Proc. Sci., LATTICE2015* (2016) 066.
- [15] H. Jeong, S. Jwa, J. Kim, S. Kim, S. Lee, W. Lee, and J. Pak (SWME Collaboration), How to identify zero modes for improved staggered fermions, [arXiv:1711.01826](https://arxiv.org/abs/1711.01826).
- [16] H. Jeong, C. Jung, S. Kim, W. Lee, and J. Pak (SWME Collaboration), Chiral ward identities for Dirac eigenmodes with staggered fermions, *Proc. Sci., LATTICE2019* (2019) 031 [[arXiv:2001.06568](https://arxiv.org/abs/2001.06568)].
- [17] S. Weinberg, *The Quantum Theory of Fields. Vol. 2: Modern Applications* (Cambridge University Press, Cambridge, England, 2013).
- [18] A. Bazavov *et al.* (MILC Collaboration), Nonperturbative QCD simulations with 2 + 1 flavors of improved staggered quarks, *Rev. Mod. Phys.* **82**, 1349 (2010).
- [19] M. F. L. Golterman and J. Smit, Selfenergy and flavor interpretation of staggered fermions, *Nucl. Phys.* **B245**, 61 (1984).
- [20] W.-J. Lee and S. R. Sharpe, Partial flavor symmetry restoration for chiral staggered fermions, *Phys. Rev. D* **60**, 114503 (1999).
- [21] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, *J. Res. Natl. Bur. Stand. B Math. Sci.* **45**, 255 (1950).
- [22] T. A. DeGrand and P. Rossi, Conditioning techniques for dynamical fermions, *Comput. Phys. Commun.* **60**, 211 (1990).
- [23] Y. Saad, Chebyshev acceleration techniques for solving nonsymmetric eigenvalue problems, *Math. Comput.* **42**, 567 (1984).
- [24] J. Smit and J. C. Vink, Remnants of the index theorem on the lattice, *Nucl. Phys.* **B286**, 485 (1987).
- [25] M. F. L. Golterman, Staggered mesons, *Nucl. Phys.* **B273**, 663 (1986).
- [26] M. F. L. Golterman, Irreducible representations of the staggered fermion symmetry group, *Nucl. Phys.* **B278**, 417 (1986).
- [27] H. Kluberg-Stern, A. Morel, O. Napoly, and B. Petersson, Flavors of Lagrangian Susskind fermions, *Nucl. Phys.* **B220**, 447 (1983).
- [28] D. Versteegen, Symmetry properties of fermionic bilinears, *Nucl. Phys.* **B249**, 685 (1985).
- [29] W.-J. Lee, Perturbative matching of the staggered four fermion operators for e-prime/epsilon, *Phys. Rev. D* **64**, 054505 (2001).
- [30] W.-J. Lee and M. Klomfass, One spin trace formalism for B (K), *Phys. Rev. D* **51**, 6426 (1995).
- [31] W.-J. Lee, Perturbative improvement of staggered fermions using fat links, *Phys. Rev. D* **66**, 114504 (2002).
- [32] K. Orginos, D. Toussaint, and R. L. Sugar (MILC Collaboration), Variants of fattening and flavor symmetry restoration, *Phys. Rev. D* **60**, 054503 (1999).
- [33] G. P. Lepage, Flavor symmetry restoration and Symanzik improvement for staggered quarks, *Phys. Rev. D* **59**, 074502 (1999).
- [34] M. Luscher and P. Weisz, On-shell improved lattice gauge theories, *Commun. Math. Phys.* **97**, 59 (1985); Erratum, *Commun. Math. Phys.* **98**, 433 (1985).
- [35] M. Luscher and P. Weisz, Computation of the action for on-shell improved lattice gauge theories at weak coupling, *Phys. Lett.* **158B**, 250 (1985).
- [36] M. G. Alford, W. Dimm, G. P. Lepage, G. Hockney, and P. B. Mackenzie, Lattice QCD on small computers, *Phys. Lett. B* **361**, 87 (1995).
- [37] F. D. R. Bonnet, D. B. Leinweber, A. G. Williams, and J. M. Zanotti, Improved smoothing algorithms for lattice gauge theory, *Phys. Rev. D* **65**, 114510 (2002).
- [38] J. Kim, W. Lee, and S. R. Sharpe, One-loop matching of improved four-fermion staggered operators with an improved gluon action, *Phys. Rev. D* **83**, 094503 (2011).
- [39] J. Kim, W. Lee, and S. R. Sharpe, One-loop matching factors for staggered bilinear operators with improved gauge actions, *Phys. Rev. D* **81**, 114503 (2010).
- [40] P. de Forcrand, M. G. Pérez, and I.-O. Stamatescu, Topology of the SU(2) vacuum: A lattice study using improved cooling, *Nucl. Phys.* **B499**, 409 (1997).
- [41] P. de Forcrand, M. G. Pérez, and I.-O. Stamatescu, Improved cooling algorithm for gauge theories, *Nucl. Phys. B, Proc. Suppl.* **47**, 777 (1996).
- [42] K. Cichy *et al.*, Comparison of different lattice definitions of the topological charge, *Proc. Sci., LATTICE2014* (2014) 075 [[arXiv:1411.1205](https://arxiv.org/abs/1411.1205)].
- [43] A. Hasenfratz and C. Nieter, Instanton content of the SU(3) vacuum, *Phys. Lett. B* **439**, 366 (1998).
- [44] M. Falcioni, M. L. Paciello, G. Parisi, and B. Taglienti, Again on SU(3) glueball mass, *Nucl. Phys.* **B251**, 624 (1985).
- [45] G. C. Donald, C. T. H. Davies, E. Follana, and A. S. Kronfeld, Staggered fermions, zero modes, and flavor-singlet mesons, *Phys. Rev. D* **84**, 054504 (2011).
- [46] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, Cambridge, MA, 2016).
- [47] Popular and basic loss functions such as the mean squared error (MSE) and mean absolute error (MAE) are usually used for regression problems. On the contrary, the categorical cross-entropy loss function is most applicable to multi-class classification problems.
- [48] Popular optimization methods available in the market are stochastic gradient descent, AdaGrad, RMSprop, and Adam [49].
- [49] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).

- [50] Popular activation functions are the tanh, sigmoid, and ReLU. Here, we make use of ReLU for the hidden layers since it is the simplest and fastest. The softmax function is essential for the output layer of the multiclass classification.
- [51] F. Chollet *et al.*, Keras, <https://keras.io>.
- [52] H. Jeong, S. Lee, and W. Lee (to be published).
- [53] Y. Aoki *et al.*, Non-perturbative renormalization of quark bilinear operators and B(K) using domain wall fermions, *Phys. Rev. D* **78**, 054510 (2008).
- [54] C. Sturm, Y. Aoki, N.H. Christ, T. Izubuchi, C.T.C. Sachrajda, and A. Soni, Renormalization of quark bilinear operators in a momentum-subtraction scheme with a non-exceptional subtraction point, *Phys. Rev. D* **80**, 014501 (2009).
- [55] R. Lehoucq and D. C. Sorensen, Deflation techniques for an implicitly re-started arnoldi iteration, *SIAM J. Matrix Anal. Appl.* **17**, 789 (1996).
- [56] Y. Saad, Chebyshev acceleration techniques for solving nonsymmetric eigenvalue problems, *Math. Comput.* **42**, 567 (1984).
- [57] J. Cullum and R. Willoughby, A survey of Lanczos procedures for very large real symmetric eigenvalue problems, *J. Comput. Appl. Math.* **12–13**, 37 (1985).
- [58] J. Cullum and R. A. Willoughby, Computing eigenvalues of very large symmetric matrices an implementation of a Lanczos algorithm with no reorthogonalization, *J. Comput. Phys.* **44**, 329 (1981).
- [59] J. Cullum and R. Willoughby, Lanczos algorithms for large symmetric eigenvalue computations. Vol. I: Theory, in *Classics Applied Mathematics* (Siam, 2002), <https://doi.org/10.1137/1.9780898719192>.
- [60] M. A. Clark, C. Jung, and C. Lehner, Multi-grid lanczos, *EPJ Web Conf.* **175**, 14023 (2018).
- [61] Y.-C. Jang and C. Jung, Split grid and block lanczos algorithm for efficient eigenpair generation, *Proc. Sci., LATTICE2018* (2019) 309.
- [62] M. Luscher, Local coherence and deflation of the low quark modes in lattice QCD, *J. High Energy Phys.* **07** (2007) 081.
- [63] T. Bae, D. H. Adams, C. Jung, H.-J. Kim, J. Kim, K. Kim, W. Lee, and S. R. Sharpe, Taste symmetry breaking with HYP-smearred staggered fermions, *Phys. Rev. D* **77**, 094508 (2008).