Localization properties of lattice fermions with plaquette and improved gauge actions

Maarten Golterman

Department of Physics and Astronomy, San Francisco State University, San Francisco, California 94132, USA

Yigal Shamir and Benjamin Svetitsky

School of Physics and Astronomy, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, 69978 Tel Aviv, Israel
(Received 30 March 2005; published 5 August 2005)

We determine the location λ_c of the mobility edge in the spectrum of the Hermitian Wilson operator in pure-gauge ensembles with plaquette, Iwasaki, and DBW2 gauge actions. The results allow mapping a portion of the (quenched) Aoki phase diagram. We use Green function techniques to study the localized and extended modes. Where $\lambda_c > 0$ we characterize the localized modes in terms of an average support length and an average localization length, the latter determined from the asymptotic decay rate of the mode density. We argue that, since the overlap operator is commonly constructed from the Wilson operator, its range is set by the value of λ_c^{-1} for the Wilson operator. It follows from our numerical results that overlap simulations carried out with a cutoff of 1 GeV, even with improved gauge actions, could be afflicted by unphysical degrees of freedom as light as 250 MeV.

I. INTRODUCTION

Domain-wall and overlap fermions reconcile chiral symmetry with the lattice, allowing for exact chiral symmetry at finite lattice spacing in the euclidean path-integral formulation [1–5]. While chiral symmetry can be achieved for a range of nonzero bare coupling g_0, problems arise if the bare coupling is too large. For domain-wall fermions, chiral symmetry cannot be maintained in the strong-coupling limit [6]. For overlap fermions, the built-in (modified) chiral symmetry is exact, but at strong coupling one loses either locality [7,8] or control over the number of species [4,9].

In any numerical simulation it is important to stay away from the dangerous regions of the phase diagram. The lattice Dirac operators of domain-wall fermions (DWF) and overlap fermions are both based on a Wilson operator with a negative, supercritical bare mass m_0.[1] Locality and chirality in these formulations are controlled by the spectral properties of this underlying Wilson operator.

The outstanding features of the supercritical Wilson operator are best illustrated in a theory with two dynamical flavors of Wilson fermions. Here the absence of a spectral gap in part of the phase diagram implies the existence of propagating, light degrees of freedom. Moreover, a non-zero spectral density for vanishing eigenvalue signals spontaneous symmetry breaking, as follows from the Banks-Casher relation [10]. For Wilson fermions there is no chiral symmetry to be broken, so the spontaneously broken symmetry is vectorial. As discovered by Aoki [11], the pions become massless if the bare Wilson-quark mass m_0 is lowered from positive values towards a critical value $m_0 = m_c(g_0) < 0$. For $m_0 < m_c(g_0)$ the curvature of

the effective potential for pions becomes negative at the origin; a pionic condensate forms which breaks spontaneously isospin and parity. This is the Aoki phase. Inside the Aoki phase the condensing pion is massive, while the other two pions are Goldstone bosons of the spontaneously broken isospin generators. See Fig. 1 for a schematic representation of the phase diagram.
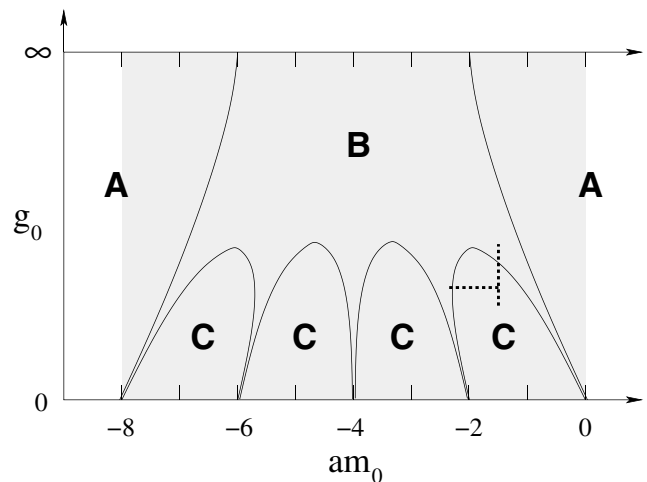

FIG. 1. A schematic representation of the Aoki phase diagram. The lightly shaded area is the supercritical region. Phase B is the (massless) Aoki phase, while phases A and C are massive. Depending on the action in the theory, each of the thin fingers where the Aoki phase touches the line $g_0 = 0$ could alternatively be replaced by a line of first-order phase transition [41]. Recent numerical results support the latter scenario in the case of dynamical Wilson fermions [42]. The phase diagram is symmetric under the replacement $am_0 \rightarrow -(8 + am_0)$, which can be undone by redefining the fields. Dotted lines indicate where we performed our numerical analysis.

---

[1] The supercritical region is $-8 < am_0 < 0$. Outside of this region the Wilson operator cannot have zero eigenvalues.

For DWF or overlap fermions one aims for a big gap, of order $1/a$, in the spectrum of the Wilson operator. Simply speaking, a bigger gap in the spectrum of the Wilson operator improves both the chiral symmetry of DWF (at fixed, finite extent of the fifth dimension) and the locality of overlap fermions. (As we will shortly see, in reality any such gap can be mostly, but not completely, devoid of eigenvalues.) Assuming that the phase diagram of the underlying Wilson operator remains qualitatively the same as in Fig. 1, this requires being inside one of the C phases. In practice, the rightmost C phase is used which, for $g_0 \to 0$, coincides with the interval $-2 < am_0 < 0$. In this interval, one lattice domain-wall (or overlap) field gives rise to one quark in the continuum limit.

When one studies the spectral properties of the Wilson operator as a kernel for DWF or overlap fermions, one is in fact considering a *quenched* Wilson-fermion theory, because the Boltzmann weight is derived from a different fermion operator. Usually, quenching means leaving out the fermion determinant altogether, but we will also consider the more general sense that the fermion determinant is that of DWF or the overlap operator. In this paper, we will study the spectrum of the Wilson operator for a variety of pure-gauge theories, but we will take our results as indicative of what might arise in a theory with dynamical fermions.

The Goldstone theorem connects spontaneous symmetry breaking with the appearance of massless poles in correlation functions. Quenching the Wilson fermions, however, opens the door to a new dynamical possibility. Considering again a two-flavor theory, let us suppose that all the correlation functions of the two (quenched) Wilson flavors decay exponentially. (As we will see, this is indeed the case well inside the C phases.) In the quenched theory, this does *not* preclude a nonzero pion condensate: The isospin symmetry can be broken spontaneously *without* creating a Goldstone boson. It was shown in [7,12] how this can be reconciled with the usual Ward-identity argument for the existence of a Goldstone pole.

There is now solid numerical and semianalytical evidence [13,14] for the existence of zero modes of the Wilson operator throughout practically the entire supercritical region. The Banks-Casher relation again leads to a nonzero pion condensate. In particular, the pion condensate is nonzero for parameter values in the C phase that yield good-quality DWF and overlap-fermion simulations. Despite the condensate, all the Wilson-fermion correlation functions are short ranged.

Reference [7] provides a theoretical explanation of this situation. The low-lying eigenvectors of the Hermitian Wilson operator may be either extended or (exponentially) localized. In the first case, the condensate must be accompanied by Goldstone bosons. In the second case, there is an alternative mechanism for saturating the relevant Ward identity, and all correlation functions can be (and, in fact,

are) short ranged. This gives the following physical picture for the quenched Wilson-fermion phase diagram.[2] In the entire supercritical region there is no gap in the spectrum of the Wilson operator, and the pion condensate is nonzero. For eigenvalues λ above a certain mobility edge λ_c, the eigenvectors of the Wilson operator are extended. If $\lambda_c > 0$, eigenvalues $|\lambda| < \lambda_c$ correspond to localized eigenmodes. In this case the pion condensate is nonzero, but there are still no long-range correlations. When $\lambda_c = 0$, on the other hand, the condensate arises from extended eigenmodes, and there are Goldstone pions.

The quenched Aoki phase is identified with the region where Goldstone pions exist; that is, it is defined by $\lambda_c = 0$. With this definition, the quenched phase diagram could be qualitatively similar to that depicted in Fig. 1. Early numerical evidence supporting this quenched phase structure may be found in Ref. [16]. The weak-coupling region may also be studied via an effective Lagrangian [17].

As far as DWF and overlap fermions are concerned, the requirement of a gap in the Wilson spectrum should be replaced by the requirement that $\lambda_c > 0$ [7]. In other words, one must work outside of the (quenched) Aoki phase, in one of the C phases. It is therefore important to map out the Aoki phase on any ensemble used for DWF or overlap-fermion numerical simulations. Furthermore, for practical reasons, one should not be too close to the Aoki phase. How close is "too close" depends on the underlying Boltzmann weight and on the construction of the fermion operator. We will discuss this very practical point at some length in our conclusions.

In this paper we study the spectral properties of the Wilson operator H_W via calculation of its resolvent $(H_W - z)^{-1}$ and correlation functions derived from it. The theoretical framework developed in Ref. [7] is directly applicable, and guides us in the numerical implementation. We measure the spectral density as well as properties that characterize the shape and size of the localized eigenmodes. The resolvent gives us these quantities much more economically than would the direct study of the eigenvalues and eigenvectors of H_W. The correlation functions address the Ward identities and the Banks-Casher relation directly.

The resolvent allows simultaneous treatment of localized and extended modes. In any volume, the eigenvalues corresponding to localized modes are random. When the resolvent is averaged over the gauge ensemble, the single-configuration spectral density, $V^{-1} \sum_n \delta(\lambda - \lambda_n)$, is smeared. Thus the ensemble-averaged spectral density is a continuous function of the eigenvalue for localized as well as extended modes. The essential physics of the localized modes lies not in their discreteness but in the compactness of their wave functions.

---

[2]See also Ref. [15], in particular, Fig. 2 therein.

Our measurements are carried out on pure-gauge ensembles (which is the usual meaning of quenching). We compare the spectral properties for three different pure-gauge actions: the standard plaquette action and the Iwasaki [18,19] and DBW2 [20,21] actions, two gauge actions motivated by renormalization-group considerations. These gauge actions have been used in quenched [22,23] and dynamical [24] DWF simulations, and in quenched overlap simulations [25,26].

This paper is organized as follows. In Sec. II we give basic definitions and derive the relevant Ward identities. In Sec. III we review the Banks-Casher relation as well as the localization alternative to Goldstone's theorem. A twisted-mass term [11,27] provides the "magnetic field" that determines the direction of the pion condensate. Careful study of the vanishing twisted-mass limit reveals that, if the low-lying eigenmodes are localized, the two-point function of the would-be Goldstone pions diverges linearly with the inverse twisted mass. This enables the relevant Ward identity to be saturated without a Goldstone pole.

We then turn to our numerical investigations, starting with the standard plaquette action for the gauge field. In Sec. IV we present results for the simplest quantity, the spectral density. In Sec. V we define the localization length and use it to determine the mobility edge λ_c for several points (g_0, m_0) in the phase diagram. Extrapolations of λ_c to zero allow us to map out a part of the boundary of the Aoki phase. We then proceed to detailed study of the localized modes. In Sec. VI we extend the investigation to the Iwasaki and DBW2 gauge actions. We conclude in Sec. VII with a discussion of the implications of our results for domain-wall and overlap fermions. The results of Sec. V yield several quantities that help locate regions of the phase diagram to be avoided in simulations.

A concise account of this work, not including the improved gauge actions, has already been given [28].

II. DEFINITIONS

A. Fermion action

The Wilson-Dirac operator is defined as

$$D(m_0) = \frac{1}{a}\begin{pmatrix} (W + am_0) & -C \\ C^\dagger & (W + am_0) \end{pmatrix}, \qquad (2.1)$$

where

$$C_{xy} = \frac{1}{2}\sum_\mu [\delta_{x+\hat{\mu},y}U_{x\mu} - \delta_{x-\hat{\mu},y}U_{y\mu}^\dagger]\sigma_\mu \qquad (2.2)$$

comes from the naive Dirac operator, and

$$W_{xy} = 4\delta_{xy} - \frac{1}{2}\sum_\mu [\delta_{x+\hat{\mu},y}U_{x\mu} + \delta_{x-\hat{\mu},y}U_{y\mu}^\dagger] \qquad (2.3)$$

is the Wilson operator that breaks chiral symmetry while preventing species doubling. Here $\sigma_\mu = (\boldsymbol{\sigma}, i)$, where σ_k are the three Pauli matrices; we are using a chiral basis for

the Dirac matrices, where γ_5 is diagonal. $U_{x\mu}$ is the SU(N) matrix representing the gauge field. We study the spectrum of the Hermitian Wilson-Dirac operator,

$$H_W(m_0) = D(m_0)\gamma_5. \qquad (2.4)$$

The corresponding eigenvalue equation (in a given gauge field) is

$$H_W(m_0)\Psi_n = \lambda_n \Psi_n, \qquad (2.5)$$

and we normalize the eigenvectors according to $a^4\sum_x |\Psi_n(x)|^2 = 1$.

Previous studies of the spectrum of H_W [13,29,30] were based on the calculation of individual eigenfunctions and eigenvalues. We find it more economical to calculate the Green function

$$G(z) = [H_W(m_0) - z]^{-1}, \qquad (2.6)$$

where $z = \lambda + im_1$, in order to extract information about the spectrum. $G(\lambda + im_1)$ is well defined in finite volume provided $m_1 \neq 0$. It has the spectral representation

$$G(x, y; \lambda + im_1) = \sum_n \frac{\Psi_n(x)\Psi_n^\dagger(y)}{\lambda_n - \lambda - im_1}. \qquad (2.7)$$

B. Two flavors and twisted mass

As mentioned in the Introduction, the spectral properties of H_W have profound effects on the realization of continuous symmetries when there is more than one flavor. Thus we will add an isospin index to the fermion field and consider the two-flavor theory defined by

$$S_F = \overline{\psi}(H_W - \lambda)\psi' \qquad = \overline{\psi}(D - \lambda\gamma_5)\psi, \qquad (2.8)$$

where $\psi' = \gamma_5\psi$. Spontaneous breaking of the flavor symmetry (and of parity) will be connected with the condensation of the "pion" field,

$$\pi_\pm(x) = i\overline{\psi}(x)\gamma_5\tau_\pm\psi(x), \qquad \pi_3(x) = i\overline{\psi}(x)\gamma_5\tau_3\psi(x), \qquad (2.9)$$

where $\tau_\pm = (\tau_1 \pm i\tau_2)/2$. The parameter λ has been introduced into Eq. (2.8) in order to shift the focus from zero to nonzero eigenvalues of H_W. In order to control the isospin orientation of the condensate, we add to the action a "magnetic field" in the guise of a twisted-mass term, giving finally

$$S_F = \overline{\psi}[H_W - (\lambda + i\tau_3 m_1)]\psi'$$
$$= \overline{\psi}[D - (\lambda + i\tau_3 m_1)\gamma_5]\psi. \qquad (2.10)$$

m_1 will be used as a regulator to avoid the singularities of $G(z)$ along the real axis.

C. SU(2) flavor symmetry and Ward identities

For $m_1 = 0$, the fermion action (2.10) has a (vector) SU(2) flavor symmetry. For $m_1 \neq 0$, the Ward identity of

the broken symmetry is obtained by performing a local flavor transformation,

$$\psi(x) \rightarrow \psi(x) + \alpha(x)\delta_+\psi(x), \qquad (2.11)$$

and similarly for $\bar{\psi}(x)$, where

$$\delta_+\psi(x) = i\tau_+\psi(x), \qquad \delta_+\bar{\psi}(x) = -i\bar{\psi}(x)\tau_+. \quad (2.12)$$

We find for any operator $\mathcal{O}$ that

$$\partial_\mu^*\langle J_\mu^+(x)\mathcal{O}(y)\rangle + 2m_1\langle\pi_+(x)\mathcal{O}(y)\rangle = \frac{i\delta_{xy}}{a^4}\langle\delta_+\mathcal{O}(y)\rangle. \qquad (2.13)$$

Here the backward lattice derivative is defined by $\partial_\mu^* f(x) = [f(x) - f(x - \hat{\mu})]/a$, and the vector current corresponding to Eq. (2.12) is

$$J_\mu^+(x) = \frac{1}{2}[\bar{\psi}(x)\tau_+(\gamma_\mu - 1)U_\mu(x)\psi(x + \hat{\mu})$$
$$+ \bar{\psi}(x + \hat{\mu})\tau_+(\gamma_\mu + 1)U_\mu^\dagger(x)\psi(x)]. \quad (2.14)$$

While the notation $\langle\ldots\rangle$ indicates an integration over both fermion and gauge fields, the Ward identity (2.13) in its various guises is in fact valid for each gauge configuration separately.

We define the pion two-point function

$$\Gamma(x, y) = \langle\pi_+(x)\pi_-(y)\rangle \qquad (2.15)$$

as well as

$$\Gamma_\mu(x, y) = \langle J_\mu^+(x)\pi_-(y)\rangle. \qquad (2.16)$$

[Note that this is the usual pion only when $\lambda = m_1 = 0$; see Eq. (2.10).] If we take $\mathcal{O}(y) = \pi_-(y)$ in Eq. (2.13), we find the Ward identity

$$\partial_\mu^*\Gamma_\mu(x, y) + 2m_1\Gamma(x, y) = \frac{\delta_{xy}}{a^4}\langle\pi_3(y)\rangle. \qquad (2.17)$$

Applying a Fourier transform, viz.

$$\tilde{\Gamma}(p) = \frac{a^8}{V}\sum_{xy}e^{ip(y-x)}\Gamma(x, y), \qquad (2.18)$$

$$\tilde{\Gamma}_\mu(p) = \frac{a^8}{V}\sum_{xy}e^{ip(y-x)}\Gamma_\mu(x, y), \qquad (2.19)$$

we derive the momentum-space Ward identity,

$$\frac{1}{a}\sum_\mu(1 - e^{-iap_\mu})\tilde{\Gamma}_\mu(p) + 2m_1\tilde{\Gamma}(p) = \langle\pi_3\rangle. \quad (2.20)$$

III. GOLDSTONE'S THEOREM AND LOCALIZATION

The Ward identity (2.20) is valid for arbitrary λ and m_1. In a quenched theory, however, despite the Goldstone theorem, $\langle\pi_3\rangle \neq 0$ does *not* necessarily imply the exis-

tence of a massless pole in $\tilde{\Gamma}_\mu(p)$ in the limit $m_1 \rightarrow 0$. Let us recall [7,12] how this comes about.

A. Condensate and Banks-Casher relation

The volume-averaged pion condensate

$$\langle\pi_3\rangle = (a^4/V)\sum_x\langle\pi_3(x)\rangle, \qquad (3.1)$$

in the two-flavor theory (2.10) can be expressed in terms of the Green function $G(\lambda \pm im_1)$. We will denote the expectation value in a given gauge field by $\langle\ldots\rangle_U$. Then

$$a^4\sum_x\langle\pi_3(x)\rangle_U = -ia^4\text{Tr}[G(\lambda + im_1) - G(\lambda - im_1)]$$

$$= 2\sum_n\frac{m_1}{(\lambda_n - \lambda)^2 + m_1^2}, \qquad (3.2)$$

where we have used the spectral representation (2.7). Averaging this over the gauge field gives the translation-invariant result

$$\langle\pi_3\rangle = 2\int d\lambda'\rho(\lambda')\frac{m_1}{(\lambda' - \lambda)^2 + m_1^2}, \qquad (3.3)$$

where $\rho(\lambda)$ is the eigenvalue density defined by

$$\rho(\lambda) = \frac{1}{V}\left\langle\sum_n\delta(\lambda - \lambda_n)\right\rangle. \qquad (3.4)$$

In the limit $m_1 \rightarrow 0$, we obtain

$$\langle\pi_3\rangle = 2\pi\rho(\lambda). \qquad (3.5)$$

This is a generalized Banks-Casher relation; the original relation [10] is Eq. (3.5) at $\lambda = 0$.

B. Localization as an alternative to Goldstone's theorem

Naively taking the limit $m_1 \rightarrow 0$ in Eq. (2.20) gives

$$\langle\pi_3\rangle \overset{?}{=} \frac{1}{a}\sum_\mu(1 - e^{-iap_\mu})\tilde{\Gamma}_\mu(p) \approx ip_\mu\tilde{\Gamma}_\mu(p), \quad (3.6)$$

where the last expression is the approximate form for $ap \ll 1$. As we shall see shortly, Eq. (3.6) is sometimes false, but it contains the Goldstone theorem: $\tilde{\Gamma}_\mu(p)$ must have a massless pole for any λ such that $\langle\pi_3\rangle \neq 0$. By the generalized Banks-Casher relation (3.5), this happens whenever $\rho(\lambda) \neq 0$. Apart from $\Gamma_\mu(x, y)$, one expects long-range power-law decay also for other correlation functions, including in particular $\Gamma(x, y)$.

In the physics of disordered systems [31] it is well known that the eigenmodes of a Hamiltonian in a random background divide into two classes: *extended* and *local-*

ized. In fact, the spectrum splits into bands, each band containing only eigenmodes of one type.[3] A point in the spectrum separating an extended band from a localized band is a *mobility edge*. $\Gamma_\mu(x, y)$ will exhibit a power-law decay when λ lies in an extended band, while for a localized band it will decay exponentially. We expect that the same basic separation applies to H_W as well.

If $\rho(\lambda)$ comes from localized eigenmodes and $\tilde{\Gamma}_\mu(p)$ has no pole at zero momentum, what has become of the Goldstone theorem? In other words, how is the Ward identity (2.20) satisfied? The way out of this conundrum is the following. In the limit $m_1 \to 0$, $\tilde{\Gamma}(p)$ *diverges* as $1/m_1$ for a range of values of p that includes the point $p = 0$. The limiting value of $m_1\tilde{\Gamma}(p)$ is finite. For $p \to 0$, we arrive at an alternative to Eq. (3.6),

$$\lim_{m_1 \to 0} 2m_1 \tilde{\Gamma}(0) = \lim_{m_1 \to 0} \langle \pi_3 \rangle. \tag{3.7}$$

C. Divergence of the pion two-point function

Let us consider further the $1/m_1$ divergence in $\tilde{\Gamma}(p)$. The (finite-volume) spectral representation of the charged-pion two-point function is

$$\Gamma(x, y) = \left\langle \sum_{n^\pm} \Psi^\dagger_{n^+}(x) \Psi_{n^-}(x) \right.$$
$$\times \frac{1}{\lambda_{n^-} - \lambda + im_1} \Psi^\dagger_{n^-}(y) \Psi_{n^+}(y)$$
$$\left. \times \frac{1}{\lambda_{n^+} - \lambda - im_1} \right\rangle, \tag{3.8}$$

where terms with the subscript $n^\pm$ are associated with the propagator for the corresponding quark flavor. As explained in Sec. 3 of Ref. [7], a $1/m_1$ divergence may arise only from the terms with $n^+ = n^-$, so that

$$\Gamma(x, y) = \frac{1}{m_1} \left\langle \sum_n |\Psi_n(x)|^2 |\Psi_n(y)|^2 \frac{m_1}{(\lambda_n - \lambda)^2 + m_1^2} \right\rangle$$
$$+ O(1). \tag{3.9}$$

In analogy with Eq. (3.4), we define the eigenmode-density correlation function,

$$\mathcal{R}_\lambda(x, y) = \left\langle \sum_n |\Psi_n(x)|^2 |\Psi_n(y)|^2 \delta(\lambda - \lambda_n) \right\rangle, \tag{3.10}$$

and its Fourier transform,

$$\tilde{\mathcal{R}}_\lambda(p) = \frac{1}{V} \left\langle \sum_n |H_n(p)|^2 \delta(\lambda - \lambda_n) \right\rangle, \tag{3.11}$$

where

$$H_n(p) = a^4 \sum_x |\Psi_n(x)|^2 e^{ipx}. \tag{3.12}$$

As $m_1 \to 0$, these may be interpreted as the contribution to $\Gamma(x, y)$, or to its Fourier transform, of eigenmodes with eigenvalue λ. Repeating the analysis leading from Eqs. (3.2) and (3.5) we find, for $m_1 \ll 1$,

$$\Gamma(x, y) = \frac{\pi \mathcal{R}_\lambda(x, y)}{m_1} + O(1). \tag{3.13}$$

Observe that $\mathcal{R}_\lambda(x, y)$ is the ensemble average (3.10) of a quantity that is strictly positive. Also, since the eigenmodes are normalized, we have that

$$\tilde{\mathcal{R}}_\lambda(0) = \frac{a^8}{V} \sum_{xy} \mathcal{R}_\lambda(x, y) = \rho(\lambda), \tag{3.14}$$

and so $\rho(\lambda) \neq 0$ means that $\mathcal{R}_\lambda(x, y)$ must be nonzero for at least one pair of values (x, y). Consequently, in any *finite* volume, $\rho(\lambda) \neq 0$ implies the existence of a $1/m_1$ divergence in the coordinate-space two-point function. [This must be true for at least one value of (x, y), but is expected to hold for practically every (x, y).] This result is valid for the generalized quenched theory defined by Eq. (2.10) for any Boltzmann weight; the only assumption we used is that the Boltzmann weight does not depend on λ and m_1.

[An *unquenched* theory with fermion action (2.10) would include $\det(H_W - \lambda - i\tau_3 m_1)$ in the Boltzmann weight. This determinant will suppress eigenvalues $\lambda_n \approx \lambda$, and the spectral density measured by Eq. (3.5) will be zero in any finite volume in the limit $m_1 \to 0$.]

We next consider the infinite-volume limit. The asymptotic behavior of an exponentially localized eigenmode is

$$|\Psi_n(x)|^2 \sim \exp\left(-\frac{|x - x_n^0|}{l_n}\right), \qquad |x - x_n^0| \gg l_n, \tag{3.15}$$

where l_n is the *localization length* and x_n^0 is the center of the localized eigenmode. Equation (3.15) is valid only at distances that are large compared to the size of the region containing most of the eigenmode's density. In principle, nothing forbids the occurrence of eigenmodes with a very short localization length $l_n \lesssim a$.

An extended eigenmode is one that does not satisfy Eq. (3.15) for any finite l_n. Evidently, truly extended eigenmodes exist only in infinite volume. In finite volume, the clear-cut identification of an eigenmode as localized demands that $|\Psi_n(x)|^2$ is exponentially small on most of the lattice. In later sections we will give a more quantitative criterion.

In infinite volume, we expect the Fourier transform of the eigenmode density (3.15) to have the following small-p behavior:

$$H_n(p) \approx \frac{e^{ipx_n^0}}{1 + p^2 l_n^2}. \tag{3.16}$$

The region $p^2 l_n^2 \ll 1$ will reflect only the exponentially

---

[3]There seems to be no rigorous proof of this fact except in one dimension [31].

decaying envelope (3.15) of the eigenmode density but not the short-distance fluctuations. The overall normalization is set by $H_n(0) = 1$ for a normalized eigenmode. Substituting the ansatz (3.16) into Eq. (3.11) and going to small p we find, in analogy with Eq. (3.13),

$$\tilde{\Gamma}(p) = \frac{\pi \rho(0)}{m_1}[1 + O(p^2 \bar{l}^2)] + O(1), \qquad (3.17)$$

where $\bar{l}$ is some average localization length for the eigenmodes with eigenvalue λ. [This extends Eq. (3.7) to nonzero p.] From Eq. (2.20) we conclude that

$$\tilde{\Gamma}_\mu(p) \sim ip_\mu \rho(0)\bar{l}^2. \qquad (3.18)$$

Thus there is no Goldstone pole when the eigenmodes with eigenvalue λ are exponentially localized.[4]

When the eigenmodes at the given λ are extended, the transition from Eq. (3.8) to Eq. (3.9) is not justified if the limit $m_1 \to 0$ is taken after the infinite-volume limit, because of interference effects between eigenmodes with infinitesimally close eigenvalues. In this case we expect Eq. (3.6) to be valid, indicating a Goldstone pole.

IV. WILSON PLAQUETTE ACTION: SPECTRAL DENSITY

We now turn to our numerical investigation. We begin with quenched ensembles generated with the Wilson plaquette action for the SU(3) gauge theory,

$$S = \frac{\beta}{3} \sum_{\substack{x \\ \mu < \nu}} \text{ReTr}(1 - U_{x\mu\nu}). \qquad (4.1)$$

We began calculating at $\beta = 6.0$, which is usually taken to correspond to the lattice scale $a^{-1} \simeq 2$ GeV, and at[5] $m_0 = -1.5$, between the Aoki "fingers" that point to the $\beta = \infty$ ($g_0 = 0$) axis at $m_0 = 0$ and -2. Then we moved downward in β towards the Aoki phase, calculating at $\beta = 5.85$, 5.7 (where $a^{-1} \simeq 1$ GeV), 5.6, 5.5, and 5.4. We will show that the Aoki phase is entered just below $\beta = 5.6$ (see Fig. 1).

Returning to $\beta = 5.7$, we moved "sideways" by changing m_0 to -2.0 and -2.4. The latter turns out to be very close to or in the second Aoki finger. We put these choices of (β, m_0) into the context of other work in Sec. VII.

For each value of β we generated 120 uncorrelated gauge configurations (except where otherwise noted) on a lattice of 16^4 sites, using the MILC pure-gauge overrelaxation code. For a given β, results for all values of m_0, λ, and m_1 were calculated on the same ensemble; thus correlations had to be taken into account in all fits and statistical analysis.

---

[4]See also Sec. IV of Ref. [7].

[5]For the remainder of the paper we rescale $am_0 \to m_0$, giving the bare Wilson mass in lattice units.

We shall illustrate our methods by discussing in detail the analysis for $\beta = 5.85$ and $m_0 = -1.5$. Results for other values of (β, m_0) will be summarized in the tables.

A. Spectral density from Banks-Casher relation

From the Banks-Casher relation (3.2) and (3.5) we have

$$2\pi\rho(\lambda) = \lim_{m_1 \to 0} \langle \pi_3 \rangle = \frac{2}{V} \lim_{m_1 \to 0} \langle \text{Im Tr}G(\lambda + im_1) \rangle. \quad (4.2)$$

Thus the volume- and ensemble-averaged Green function, extrapolated to $m_1 = 0$, gives $\rho(\lambda)$ directly. Of course m_1 must be kept nonzero for actual calculation in order for G to be bounded.

The spectral sum (3.2) shows how to do the extrapolation. For any gauge configuration $\mathcal{U}$ we consider the sum

$$\sum_n \frac{m_1}{(\lambda_n - \lambda)^2 + m_1^2}. \qquad (4.3)$$

The summand tends to a δ function as $m_1 \to 0$, but before the limit is taken it has a finite width equal to m_1. The given configuration will make a contribution of $O(1/m_1)$ if H_W has an eigenmode whose eigenvalue λ_n satisfies $|\lambda_n - \lambda| \lesssim m_1$; these contributions, summed over configurations $\mathcal{U}$, will add up to a finite limit as $m_1 \to 0$. On the other hand, all the eigenmodes that are far from λ, with $a|\lambda_n - \lambda| = O(1)$, will make a contribution of $O(m_1)$. This indicates a linear extrapolation,

$$\frac{\langle \pi_3 \rangle}{2\pi} = c_0 + c_1 m_1, \qquad (4.4)$$

where $c_{0,1}$ will depend on λ. Then c_0 is an estimate for $\rho(\lambda)$.

We calculated TrG using a single random source per gauge configuration. Averages were obtained for up to seven values of m_1: $0.01, 0.02, \ldots, 0.07$. The upper graph in Fig. 2 shows the linear extrapolation for two values of λ. For $\lambda = 0.3$ the fit works well. For $\lambda = 0$, on the other hand, the extrapolated intercept is very small and the precision attained is inadequate.

The problem is that, when $\rho(\lambda)$ is too small, there are few eigenmodes (for the volume we use) within the broadened δ function in Eq. (4.3). Then the sum is dominated by the $O(m_1)$ contributions of the more distant eigenmodes, with large fluctuations. Making m_1 even smaller will suppress these contributions, but it will be self-defeating because even fewer modes will lie within the broadened δ function, increasing the fluctuations in their contribution as well.

B. Improved estimator

A solution lies in using an improved estimator for the spectral density, one that suppresses the contribution of distant eigenmodes and thus approaches the $m_1 \to 0$ limit faster than linearly. We introduce the (dimensionless) differential operator $\mathcal{J}$ defined by
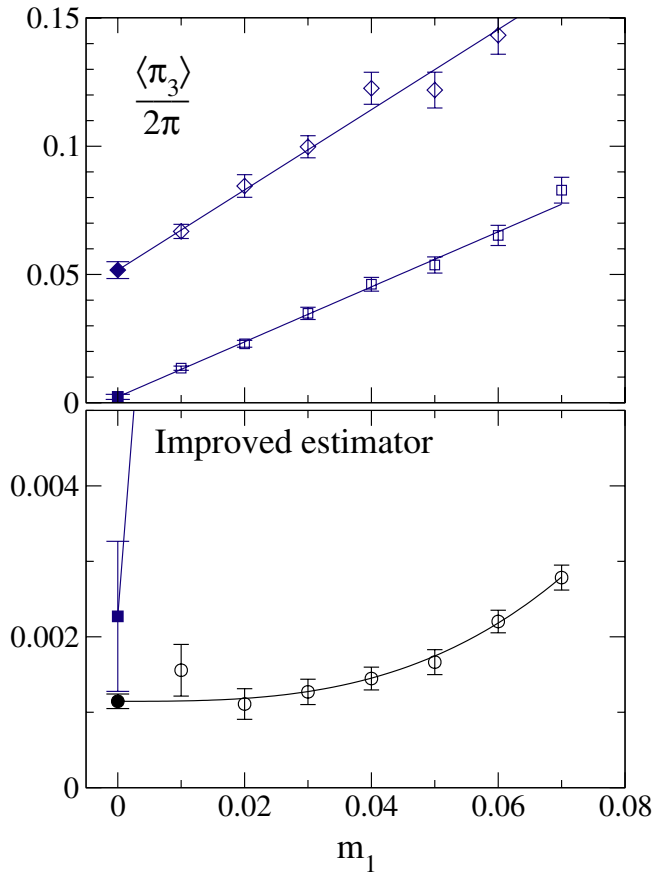
FIG. 2 (color online). Top graph: the pion condensate $(2\pi)^{-1}\langle\pi_3\rangle$ as a function of m_1 (open symbols), and its linear extrapolation to yield the spectral density $\rho(\lambda)$ at $m_1 = 0$, for $\lambda = 0.3$ (upper curve) and $\lambda = 0$ (lower curve). Bottom graph: cubic extrapolation of improved estimator for $\lambda = 0$ (circles), compared to the result of the linear fit above (square). All data are for Wilson action at $\beta = 5.85$; $m_0 = -1.5$.

$$J f = -m_1^2 \frac{\partial}{\partial m_1}\left(\frac{1}{m_1} f\right) \qquad (4.5)$$

for any $f(m_1)$. It is designed to remove the linear term in $f(m_1)$. Applying it to the pion condensate, we have

$$J \langle\pi_3\rangle = \frac{2}{V}\langle \mathrm{Im\,Tr}G - m_1 \mathrm{Re\,Tr}G^2 \rangle, \qquad (4.6)$$

with the spectral representation

$$J \langle\pi_3\rangle = \frac{4}{V}\left\langle \sum_n \frac{m_1^3}{[(\lambda_n - \lambda)^2 + m_1^2]^2} \right\rangle. \qquad (4.7)$$

The $m_1 \to 0$ limit again yields $2\pi\rho(\lambda)$ since $\int dx\, m_1^3/(x^2 + m_1^2)^2 = \pi/2$.

The operator J indeed removes the leading, linear contribution of the eigenmodes with $a|\lambda_n - \lambda| = O(1)$. We can see from Eq. (4.7) that the contribution of these eigenmodes is now proportional to m_1^3. In view of this, we attempt a *cubic* extrapolation,

$$\frac{J\langle\pi_3\rangle}{2\pi} = c_0 + c_3 m_1^3. \qquad (4.8)$$

The lower graph in Fig. 2 shows the new extrapolation, again at $\lambda = 0$. Linear and quadratic terms are in fact unnecessary to attain an excellent fit to the data points; this is the final justification of the model (4.8). Most important, the precision in the extrapolation is improved by a factor of 10.

Using the improved estimator (4.6) requires a second inversion for each random source. The number of conjugate-gradient (CG) iterations needed for the second inversion can reach twice the number for the first inversion. The cost of the improved estimator is thus up to (roughly) 3 times the original cost. Were we to invest the additional computer time instead in increased statistics using the simple estimator (4.2), the anticipated reduction in the statistical error would not come anywhere close to the factor of 10 achieved with the improved estimator.

C. Spectral density from the two-point function

Since the two-point function $\Gamma(x, y)$ is the eigenmode-density correlator [see Eq. (3.9)], it is a rich source of information about spectral properties. The first application we discuss is an alternative method to calculate the spectral density $\rho(\lambda)$.

We project $\Gamma(x, y)$ to zero spatial momentum and calculate the time-correlation function,

$$\Gamma(t) = \frac{a^6}{V_3}\sum_{\mathbf{x},\mathbf{y}}\Gamma(0, \mathbf{x}; t, \mathbf{y}), \qquad (4.9)$$

where V_3 is the three volume. This calculation requires a random source on time slice 0, as well as a random sink for each time slice t; again we use one set of random sources per gauge configuration. If we sum $\Gamma(t)$ over all t, we can use translation invariance and Eq. (3.9) to show that

$$\lim_{m_1\to 0} am_1 \sum_t \Gamma(t) = \pi\rho(\lambda). \qquad (4.10)$$

This corresponds to setting $p = 0$ in the finite-volume Ward identity (2.20). [Note the similarity to Eq. (3.14).] We verified Eq. (4.10) for two values of the bare coupling ($\beta = 6.0$ and 5.7) by calculating both the condensate [Eq. (4.2)] and the two-point function. In all subsequent measurements we did not calculate the condensate separately since it gives only the spectral density, while $\Gamma(t)$ yields this and much more.

The extrapolation of $m_1\Gamma(t)$ to $m_1 = 0$ suffers from the same fluctuation problems as that of $\langle\pi_3\rangle$ when the spectral density is small. The solution again lies in an improved estimator, but now for the eigenmode-density correlation function $\mathcal{R}(x, y; \lambda)$ [see Eq. (3.10)]. We define the "improved" two-point function,

$$\Gamma_{\mathrm{imp}}(x, y) = \mathrm{Tr}\langle \mathcal{G}(x, y)\mathcal{G}(y, x) \rangle, \qquad (4.11)$$

where

$$\mathcal{G}(x, y) = a^4 \sum_z G(x, z; \lambda + im_1)G(z, y; \lambda - im_1)$$

$$= \sum_n \frac{\Psi_n(x)\Psi_n^\dagger(y)}{(\lambda_n - \lambda)^2 + m_1^2}, \qquad (4.12)$$

and project it to zero spatial momentum,

$$\Gamma_{\text{imp}}(t) = \frac{a^6}{V_3} \sum_{\mathbf{x},\mathbf{y}} \Gamma_{\text{imp}}(0, \mathbf{x}; t, \mathbf{y}). \qquad (4.13)$$

Instead of Eq. (3.9) one has

$$m_1^3 \Gamma_{\text{imp}}(x, y) = \left\langle \sum_n |\Psi_n(x)|^2 |\Psi_n(y)|^2 \frac{m_1^3}{[(\lambda_n - \lambda)^2 + m_1^2]^2} \right\rangle$$
$$+ O(m_1^2), \qquad (4.14)$$

whence

$$\lim_{m_1 \to 0} m_1^3 \Gamma_{\text{imp}}(x, y) = \frac{\pi}{2} \mathcal{R}_\lambda(x, y). \qquad (4.15)$$

Equation (3.14) then yields the spectral density [cf. Eq. (4.10)],

$$\lim_{m_1 \to 0} a m_1^3 \sum_t \Gamma_{\text{imp}}(t) = \frac{\pi}{2} \rho(\lambda). \qquad (4.16)$$

Results obtained using $\Gamma(t)$ were extrapolated to $m_1 = 0$ linearly in m_1, while those obtained using $\Gamma_{\text{imp}}(t)$ were extrapolated[6] linearly in m_1^3.

We also tried to improve the estimator for the eigenmode density by calculating $J\Gamma(x, y)$ [cf. Eq. (4.5)]. This has the advantage that the application of J to all terms in the Ward identity (2.20) generates a new identity. As it turns out, while some reduction in the statistical error was achieved this way, the improvement was not significant. The reason is that the application of J to $\Gamma(x, y)$ would be effective in suppressing the contribution of eigenmodes with $a|\lambda_{n^\pm} - \lambda| \approx 1$ only if the approximation Eq. (3.9) is valid. As it turns out, for the m_1 values we used, Eq. (3.9) is not a good approximation to Eq. (3.8) when the spectral density is small. The correlator $\Gamma_{\text{imp}}(x, y)$ does a better job in suppressing the contribution of eigenmodes with $a|\lambda_{n^\pm} - \lambda| \approx 1$.

D. Numerical results

We present our results for the spectral density for the Wilson plaquette action in Tables I, II, III, IV, V, VI, VII, and VIII and in Figs. 3 and 4.

At each value of β and m_0, we performed our calculations for a range of values of λ starting from zero, increas-

---

[6]The form of Eq. (4.14) suggests an m_1^2 term as well, but we found it to be unnecessary for a good fit.

TABLE I. Spectral properties for $\beta = 6.0$, $m_0 = -1.5$. The mobility edge is at $\lambda_c \simeq 0.41$, marked by the space in the table, below which $\lambda_\ell = \infty$. The quantities l_ℓ, $\mathcal{D}$, l_s, and R, which characterize the localized modes, are defined in Sec. V.

λ	ρ	l_ℓ	$\mathcal{D}$	l_s	R
0.0[a,b]	0.000 11(2)	0.61(1)	0.000 038(8)	2.9(8)	31
0.1[a,b]	0.000 24(2)	0.61(1)	0.000 10(1)	2.4(3)	25
0.2[a]	0.0016(1)	0.70(1)	0.000 53(6)	3.0(4)	16
0.3[a]	0.0214(9)	1.16(4)	0.0058(4)	3.7(3)	8
0.4[a]	0.128(3)	3.8(6)	0.025(2)	5.1(4)	5.3
0.5	0.267(7)	∞	0.033(5)	8(1)	4.4
0.6	0.390(9)	∞	0.042(9)	9(2)	4

[a]Results from improved estimator.
[b]Used 1200 gauge configurations for measurements with $m_1 = 0.01, 0.02, 0.03$.

TABLE II. Spectral properties for $\beta = 5.85$, $m_0 = -1.5$. The mobility edge is at $\lambda_c \simeq 0.32$.

λ	ρ	l_ℓ	$\mathcal{D}$	l_s	R
0.0[a]	0.0011(1)	0.64(1)	0.000 36(5)	3.1(5)	17
0.1[a]	0.0019(1)	0.71(1)	0.000 54(7)	3.5(5)	15
0.2[a]	0.0088(4)	0.92(2)	0.0024(2)	3.7(3)	10
0.3[a]	0.056(1)	2.2(2)	0.0132(8)	4.2(3)	6.5
0.4	0.168(7)	∞	0.036(6)	4.7(8)	4.9
0.5	0.27(1)	∞	0.023(7)	12(4)	4.4
0.6	0.39(2)	∞	0.055(9)	7(1)	4

[a]Results from improved estimator.

TABLE III. Spectral properties for $\beta = 5.7$, $m_0 = -1.5$. The mobility edge is at $\lambda_c \simeq 0.25$.

λ	ρ	l_ℓ	$\mathcal{D}$	l_s	R
0.0[a]	0.0089(3)	1.04(3)	0.0023(2)	3.9(4)	10
0.1[a]	0.0126(7)	1.29(5)	0.0029(3)	4.3(5)	9.4
0.2[a]	0.039(1)	2.3(1)	0.0097(5)	4.0(2)	7.1
0.3[a]	0.106(2)	∞	0.026(1)	4.1(2)	5.5
0.4[a]	0.199(6)	∞	0.036(4)	5.5(6)	4.7
0.5[a]	0.308(8)	∞	0.052(4)	5.9(5)	4.2

[a]Results from improved estimator.

TABLE IV. Spectral properties for $\beta = 5.6$, $m_0 = -1.5$. The mobility edge is at $\lambda_c = 0.14(2)$. (We give the statistical error in λ_c here because it is larger than in other cases. All suffer from systematic error in the extrapolation.)

λ	ρ	l_ℓ	$\mathcal{D}$	l_s	R
0.0	0.027(2)	2.1(1)	. . .	. . .	7.8
0.1	0.034(3)	4(1)	. . .	. . .	7.4
0.2	0.072(4)	∞	0.012(3)	6(2)	6.1
0.3	0.135(7)	∞	0.016(5)	8(3)	5.2
0.4	0.219(8)	∞	0.017(6)	13(5)	4.6

TABLE V. Spectral properties for $\beta = 5.5$, $m_0 = -1.5$. The mobility edge is at zero.

λ	ρ	l_ℓ	$\mathcal{D}$	l_s	R
0.0[a]	0.057(1)	∞	0.0112(6)	5.1(3)	6.5
0.1[a]	0.073(1)	∞	0.013(1)	5.6(4)	6.1
0.2	0.116(5)	∞	0.015(4)	8(2)	5.4
0.3	0.172(7)	∞	0.019(4)	9(2)	4.9
0.4	0.25(1)	∞	0.029(6)	9(2)	4.5

[a]Results from improved estimator.

TABLE VI. Spectral properties for $\beta = 5.4$, $m_0 = -1.5$. The mobility edge is at zero.

λ	ρ	l_ℓ	$\mathcal{D}$	l_s	R
0.0	0.066(3)	∞	...	...	6.2
0.1	0.089(5)	∞	0.008(3)	11(4)	5.8
0.2	0.148(6)	∞	0.022(4)	7(1)	5.1
0.3	0.194(7)	∞	0.018(4)	11(3)	4.8
0.4	0.27(1)	∞	0.028(6)	10(2)	4.4

TABLE VII. Spectral properties for $\beta = 5.7$, $m_0 = -2.0$. The mobility edge is at $\lambda_c \simeq 0.21$.

λ	ρ	l_ℓ	$\mathcal{D}$	l_s	R
0.0	0.014(2)	1.14(4)	0.005(2)	3(1)	9.1
0.1	0.022(2)	1.37(5)	0.005(2)	4(2)	8.2
0.2	0.13(2)	3.8(5)	0.03(1)	4(2)	5.3
0.3	0.24(1)	∞	0.028(4)	9(1)	4.5
0.4	0.41(2)	∞	0.05(1)	8(2)	4.0

TABLE VIII. Spectral properties for $\beta = 5.7$, $m_0 = -2.4$. The mobility edge is very close to zero.

λ	ρ	l_ℓ	$\mathcal{D}$	l_s	R
0.0	0.043(2)	4(1)	0.006(2)	7(2)	6.9
0.1	0.112(5)	∞	0.014(3)	8(2)	5.5
0.2	0.234(8)	∞	0.026(6)	9(2)	4.5
0.3	0.40(1)	∞	0.052(6)	8(1)	4.0
0.4	0.54(1)	∞	0.05(1)	11(2)	3.7

ing in steps of 0.1. In the ranges studied (the maximal λ was 0.4 to 0.6) the spectral density typically increases by 1 to 3 orders of magnitude. For the ensemble generated by the plaquette action, $\rho(\lambda)$ shows no remarkable behavior as λ passes λ_c, the mobility edge (to be determined below). In all cases, however, the mobility edge is encountered for $\rho(\lambda_c) \approx 0.1$. There is a steady rise in the spectral density as we decrease β (Fig. 3), which bespeaks an increased disorder in the gauge field.
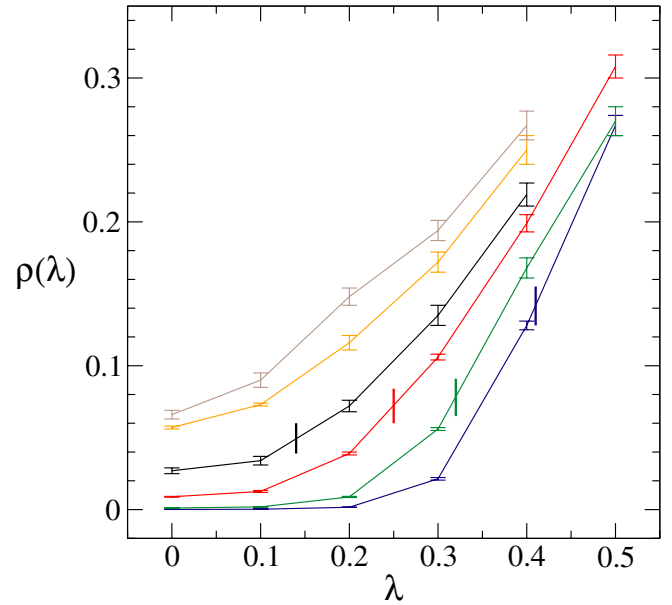

FIG. 3 (color online). Eigenvalue density $\rho(\lambda)$ for (bottom to top) $\beta = 6.0, 5.85, 5.7, 5.6, 5.5, 5.4$. Mobility edges $\lambda_c(\beta)$ (when $\lambda_c > 0$) are indicated by vertical bars. Wilson action; $m_0 = -1.5$.

A final note on improvement: For the larger values of λ it is possible to achieve acceptable precision without using the improved estimator. Since the number of CG iterations required grows rapidly with the spectral density, we limited the use of improved estimators to those cases where results obtained without improvement were poor.
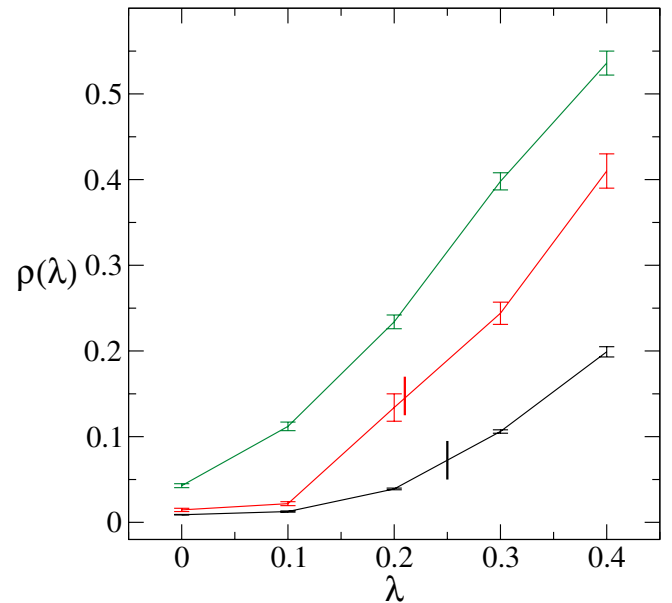

FIG. 4 (color online). Eigenvalue density $\rho(\lambda)$ for $\beta = 5.7$ and (bottom to top) $m_0 = -1.5, -2.0, -2.4$. Mobility edges are indicated by vertical bars. For $m_0 = -2.4$ the mobility edge is probably very close to $\lambda = 0$. Wilson action.

V. WILSON PLAQUETTE ACTION: LOCALIZATION PROPERTIES

A central goal of the work presented here is the determination of the mobility edge λ_c at various places in the (β, m_0) phase diagram. Equally interesting is the shape and density of the localized eigenmodes below λ_c (when $\lambda_c > 0$).

The two-point function $\Gamma(x, y)$ contains complete information about the eigenmodes of H_W, and it is our primary tool both in determining λ_c and in studying the localized modes. When dealing with localized eigenmodes, we aim for parameter values where the single spectral sum (3.9) provides a good approximation to the exact expression (3.8). When m_1 is small enough, $\Gamma(x, y)$ then reduces to the eigenmode-density correlator $\mathcal{R}_\lambda(x, y)$ [Eq. (3.10)]. Our analysis is based on this feature. The approximation (3.9) ceases to hold when the eigenmodes become too dense or too extended, that is, when they interfere in the double sum. This occurs when λ is above or too close to the mobility edge. We will develop a criterion to establish the consistency of our analysis of the localized eigenmodes.

We begin by giving a precise definition of the localization length $l_\ell(\lambda)$. Its divergence marks the mobility edge $\lambda_c(\beta, m_0)$. By definition, $\lambda_c = 0$ marks the Aoki phase. We then turn to study other properties of the localized eigenmodes outside the Aoki phase, where $\lambda_c > 0$.

A. The localization length

We have loosely defined the localization length l_n of an individual eigenmode in Eq. (3.15). We can use l_n, nebulous as it is, to motivate the definition of an *average* localization length that can be obtained from the large-t behavior of $\Gamma(t)$. Once we reach this definition, a precise definition of l_n will be superfluous.

We begin by introducing the restricted spectral density $\rho_l(\lambda)$ that contains the contributions of eigenmodes with localization length $l_n \leq l$ only. It is given by

$$\rho_l(\lambda) = \frac{1}{V}\left\langle \sum_n \delta(\lambda_n - \lambda)\theta(l - l_n) \right\rangle, \quad (5.1)$$

and its derivative, the differential spectral density, is given by

$$\rho_l'(\lambda) = \frac{d\rho_l(\lambda)}{dl} = \frac{1}{V}\left\langle \sum_n \delta(\lambda_n - \lambda)\delta(l_n - l) \right\rangle. \quad (5.2)$$

We also introduce a probability distribution $\mathcal{P}_\lambda(l)$ for the localization length l of eigenmodes with given eigenvalue λ by writing $\rho_l'(\lambda) = \rho(\lambda)\mathcal{P}_\lambda(l)$. This distribution is normalized because, in finite volume,

$$\int_0^L dl\rho_l'(\lambda) = \rho(\lambda) \quad \text{and hence} \quad \int_0^L dl\mathcal{P}_\lambda(l) = 1,$$
$$(5.3)$$

where L is the (largest) linear size of the lattice. Below, this upper limit will be implicit.

[In infinite volume there can be truly extended eigenmodes. Since we expect the eigenmodes at a given λ to be either all extended or all localized, we have correspondingly either $\mathcal{P}_\lambda(l) = 0$ for all finite l, or $\int_0^\infty dl\mathcal{P}_\lambda(l) = 1$.]

The decay rate of individual localized eigenmodes can be related to the decay rate of $\Gamma(t)$ for $m_1 \to 0$ [7]. Using Eq. (3.15) in Eq. (3.9) gives

$$m_1\Gamma(x, y) \sim \left\langle \sum_n \exp\left(-\frac{|x - x_n^0| + |y - x_n^0|}{l_n}\right) \right.$$
$$\left. \times \frac{m_1}{(\lambda_n - \lambda)^2 + m_1^2} \right\rangle. \quad (5.4)$$

Averaging over the position of x_n^0 gives

$$m_1\Gamma(x, y) \sim \frac{1}{V}\left\langle \sum_n \exp\left(-\frac{|x - y|}{l_n}\right)\frac{m_1}{(\lambda_n - \lambda)^2 + m_1^2} \right\rangle, \quad (5.5)$$

since the average is dominated by locations x_n^0 near the straight line connecting x and y. Hence

$$\lim_{m_1 \to 0} m_1\Gamma(x, y) \sim \pi\rho(\lambda)\int dl\mathcal{P}_\lambda(l)\exp\left(-\frac{|x - y|}{l}\right), \quad (5.6)$$

where we have used Eq. (5.2). Equation (4.9) then gives

$$\lim_{m_1 \to 0} m_1\Gamma(t) \sim \pi\rho(\lambda)\int dl\mathcal{P}_\lambda(l)\exp\left(-\frac{t}{l}\right). \quad (5.7)$$

Thus $\Gamma(t)$ decays exponentially, as a weighted average of exponentials with all localization lengths. Motivated by this result and ignoring power corrections, we finally define the average localization length as

$$l_\ell(\lambda) = \frac{1}{\mu(\lambda)}, \quad (5.8)$$

where $\mu(\lambda)$ is the extrapolation to $m_1 = 0$ of the decay rate M of the two-point function,

$$\Gamma(t) \sim \exp[-M(\lambda, m_1)t], \quad t \gg a. \quad (5.9)$$

Equation (5.7) actually suggests a slightly different definition of $\mu(\lambda)$, based on the extrapolation of the correlation function as a whole,

$$\lim_{m_1 \to 0} m_1\Gamma(t) \sim \exp[-\mu(\lambda)t], \quad t \gg a. \quad (5.10)$$

In practice, the extraction of $\mu(\lambda)$ using Eq. (5.10) turns out to be extremely noisy, but compatible with the results obtained using Eq. (5.9).[7] We therefore used the definition based on Eq. (5.9) throughout our calculations.

---

[7]Some further details are provided in Ref. [28].

We extrapolate to $m_1 = 0$ by fitting

$$M^2(\lambda, m_1) = \mu^2(\lambda) + \alpha(\lambda)m_1. \qquad (5.11)$$

See below for an explanation of this choice. We illustrate this extrapolation in Fig. 5. The tables list the results for $l_\ell = \mu^{-1}$, for comparison with other characteristic lengths.[8]

The derivation leading to Eq. (5.7) relies on Eq. (3.9), where interference effects present in Eq. (3.8) have been dropped. This is certainly true when the exponentially localized eigenmodes are *isolated*. A precise definition of what this means will be given later. For now it suffices to say that, intuitively, the eigenmodes with eigenvalue near λ are isolated if every such eigenmode is supported in a different part of the lattice.

As λ increases, the spectral density grows and so does the localization length of individual eigenmodes. Eventually, interference sets in, meaning that the approximation of Eq. (3.9) is not valid for the m_1 values we use. Interference will cause $\Gamma(t)$ to decay faster than the decay rate of individual eigenmodes. The definition $l_\ell = \mu^{-1}$ will then systematically underestimate the true (average) localization length, as extracted perhaps by calculating individual eigenmodes and matching them to Eq. (3.15). If we are to use Eq. (5.9) to calculate the localization length, we must verify that the eigenmodes are indeed isolated. We return to this issue below.
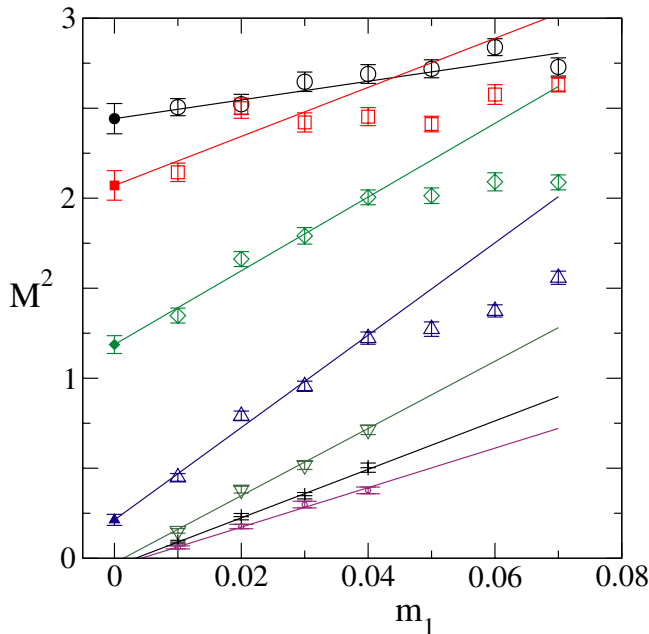
B. The mobility edge

As we increase λ, we reach a critical value λ_c where $\mu(\lambda_c) = 0$ (see Fig. 5). What is the physical significance of λ_c? According to the preceding discussion, $\mu^{-1}(\lambda)$ provides a reasonable average value for the localization lengths of individual eigenmodes if there is no interference, while it underestimates the average when interference sets in. Either way, $\mu(\lambda) = 0$ implies that the average localization length of individual eigenmodes is infinite; the eigenmodes have become *extended*. The point λ_c therefore marks the *mobility edge*.

Would a different procedure yield a different value for the mobility edge? One might consider, for instance, an alternative determination based on the numerical calculation of individual eigenmodes. We believe that such ambiguity can only be an artifact of finite-size effects. Such effects are perforce significant when $\mu(\lambda) \lesssim 1/L$, since the notion of extended eigenmodes is truly meaningful only in infinite volume. Thus some disagreement between different determinations of the mobility edge is to be expected in any finite volume. Nonetheless, we expect that all methods will converge to the same value in the infinite-volume limit [31]. We are only interested here in obtaining an overall picture, which does not justify the resources that would be needed for, say, a finite-size scaling analysis.

We estimate λ_c simply by a linear extrapolation of $\mu(\lambda)$ to zero, using the last two nonzero values we have found.
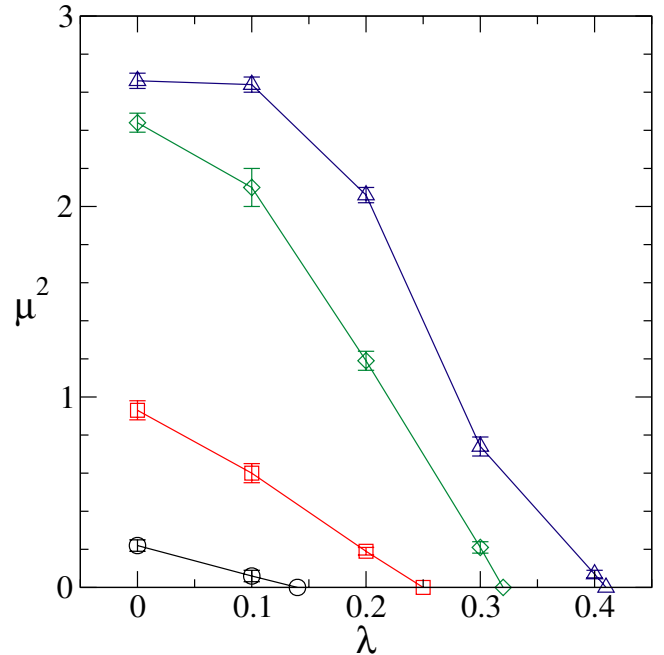

FIG. 5 (color online). Squared decay rate M^2 of the two-point function $\Gamma(t)$ vs m_1 (open symbols) and its linear extrapolation to $m_1 = 0$. Top to bottom: $\lambda = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6$. The mobility edge is at $\lambda \simeq 0.32$. All data are for Wilson action at $\beta = 5.85$; $m_0 = -1.5$.

FIG. 6 (color online). The inverse squared localization length μ^2 vs λ. Top to bottom: $\beta = 6.0, 5.85, 5.7, 5.6$. Intercepts at the λ axis are extrapolated from the last two points in each set, and provide estimates for the mobility edge λ_c. Data are for Wilson action; $m_0 = -1.5$.
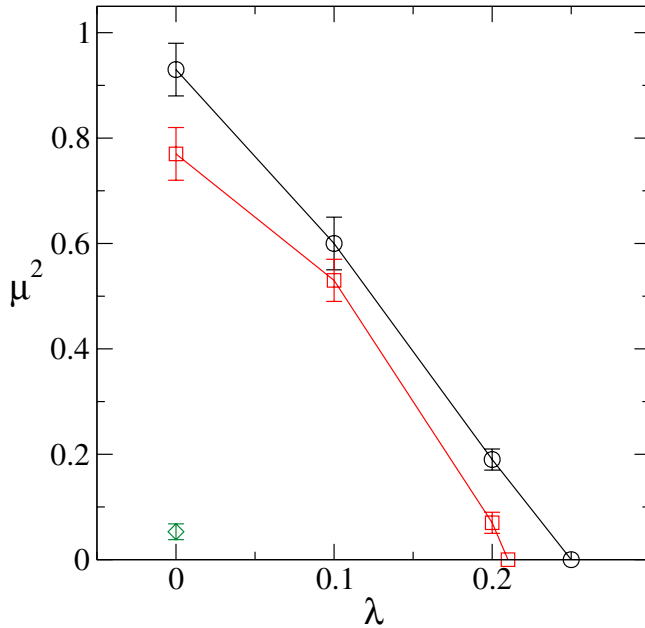
FIG. 7 (color online). Same as Fig. 6, at $\beta = 5.7$. Top to bottom: $m_0 = -1.5, -2.0, -2.4$. For $m_0 = -2.4$ the mobility edge is probably very close to $\lambda = 0$. Wilson action.

TABLE IX. Mobility edge $\lambda_c(\beta, m_0)$. Where no error is shown, the statistical error is less than one in the last digit.

β	m_0	λ_c
∞	-1.5	$1/2$
6.0	-1.5	0.41
5.85	-1.5	0.32
5.7	-1.5	0.25
5.6	-1.5	0.14(2)
5.5	-1.5	0.0
5.4	-1.5	0.0
5.7	-2.0	0.21
5.7	-2.4	≈ 0

the largest value of λ that is below the mobility edge λ_c. (This uncertainly does not occur for lower values of λ.) In fact, the calculated l_ℓ^{-1} at this value of λ is likely to be too high; its true value could even be zero. In other words, this value of λ could lie *above* the mobility edge. This means that the last measured point on each curve in Figs. 6 and 7 is unreliable.

C. Entering the Aoki phase

Our estimates for the mobility edge $\lambda_c(\beta, m_0)$ are collected in Table IX. Consider first the $m_0 = -1.5$ results. For reference, we include the free-theory limit ($\beta = \infty$) [7], where λ_c coincides with the gap of the free H_W. At $\beta = 6.0$, λ_c is still close to its free-field value. As we move towards strong coupling, the curve $\lambda_c(\beta)$ steepens before reaching zero somewhere between $\beta = 5.6$ and $\beta = 5.5$. As discussed in Sec. III, where $\lambda_c = 0$ the two-point function $\Gamma(t)$ has a zero-momentum pole even at $\lambda = 0$. The Goldstone theorem is valid, the Ward identity (3.6) is satisfied, and the (quenched) Wilson-fermion theory possesses a Goldstone boson associated with the spontaneous breakdown of its SU(2) flavor symmetry. This is the Aoki phase, and we put its boundary between $\beta = 5.6$ and 5.5.

Table IX also shows results at the other two m_0 values for $\beta = 5.7$. We find only a small change between $m_0 = -1.5$ and $m_0 = -2.0$, while the $m_0 = -2.4$ result suggests that here, too, one is near or within the Aoki phase.

The implications of these results for DWF and overlap fermions are discussed in Sec. VII.

D. Participation number and the support length

Now we turn to a more detailed description of the localized eigenmodes, for which $|\lambda| < \lambda_c$. We begin by defining a measure of the size of the support of a localized eigenmode. In the sequel we use this measure to sharpen the notion of isolated localized eigenmodes, and verify the consistency of our approach.

We define [29,31] the *participation number* P of a normalized eigenmode $\Psi(x)$ in d dimensions by

The justification for this procedure is that, for λ near λ_c, we expect that $\mu(\lambda)$ will exhibit the characteristic behavior of a (continuous) phase transition. See Figs. 6 and 7.

Above the mobility edge we expect the $1/m_1$ divergence of $\Gamma(x, y)$ to disappear in the infinite-volume limit; hence Eq. (3.6) is valid, and $\tilde{\Gamma}_\mu(p)$ should have a Goldstone pole. We expect that the physics of these Goldstone poles to be governed by some effective chiral Lagrangian. Borrowing from the familiar physics of ordinary Goldstone bosons, we expect $M^2 \propto m_1$ in the presence of the symmetry-breaking field m_1. This is the motivation for the term linear in m_1 in the fit function (5.11). Below the mobility edge M^2 does not vanish for $m_1 \to 0$, and therefore it is immaterial whether we extrapolate M or M^2 to $m_1 = 0$. Put together, these considerations suggest that Eq. (5.11) is an appropriate form both below and above the mobility edge.

Above the mobility edge, the linear extrapolation often gives a small negative result for μ^2. We have neglected to include possible logarithms which, in the context of a chiral Lagrangian, would occur at next-to-leading order in m_1. These logarithms, together with finite-size corrections, should move the extrapolated value to zero.

Our linear extrapolation to find the zero of $\mu(\lambda) = l_\ell^{-1}(\lambda)$ suffers from an uncertainty that stems from the interference to which we alluded above. On further analysis (see Sec. V E below) we will conclude that, for all (β, m_0) where $\lambda_c > 0$, our estimate of l_ℓ is unreliable for

---

[8]We calculated the localization length from $\Gamma(t)$, without improvement. Where the tables indicate improvement, it applies to ρ and $\mathcal{D}$ only.

$$P^{-1} = a^d \sum_x (|\Psi(x)|^2)^2. \qquad (5.12)$$

The physical meaning is easy to see. Suppose that the magnitude of the eigenmode density is $|\Psi(x)|^2 \approx 1/l_s^d$ over a region whose linear size is l_s. Then $P^{1/d} \approx l_s$ is a measure of the linear size of the support of $\Psi(x)$. We similarly define a generalized participation number by

$$P_{d,d'}^{-1} = a^{d+d'} \sum_{x_{d'+1},...,x_d} \left(\sum_{x_1,...,x_{d'}} |\Psi(x_\mu)|^2 \right)^2. \qquad (5.13)$$

[Equation (5.12) corresponds to $d' = 0$.] In the case considered we would have

$$P_{d,d'}^{-1} \approx l_s^{d-d'} (l_s^{d'-d})^2 = l_s^{d'-d}. \qquad (5.14)$$

In the special case $d' = d - 1$, we have $P_{d,d-1} \approx l_s$.

We are thus motivated to define $l_s(n) \equiv P_{4,3}(n)$ to be the linear size of the support of the nth eigenmode. Substituting in Eq. (3.9) and using translation invariance together with Eqs. (4.9) and (5.14) we have

$$m_1 \Gamma(t = 0) = \frac{1}{V} \left\langle \sum_n \frac{1}{l_s(n)} \frac{m_1}{(\lambda_n - \lambda)^2 + m_1^2} \right\rangle + O(m_1). \qquad (5.15)$$

Defining

$$\mathcal{D}(\lambda) = \pi^{-1} \lim_{m_1 \to 0} m_1 \Gamma(t = 0), \qquad (5.16)$$

we have

$$\mathcal{D}(\lambda) = \rho(\lambda) \int dl \mathcal{P}_\lambda^s(l) \frac{1}{l}, \qquad (5.17)$$

where $\mathcal{P}_\lambda^s(l)$ is the probability distribution for $l_s(n)$. On the basis of Eq. (5.17) we define the average *support length* as

$$l_s(\lambda) = \frac{\rho(\lambda)}{\mathcal{D}(\lambda)}. \qquad (5.18)$$

While we expect the average localization length l_ℓ and the average support length l_s to be quantities of similar magnitude, there is no reason why they should be the same.

We may also obtain $\mathcal{D}(\lambda)$ using the improved estimator (4.13). Again, results obtained using $\Gamma(t)$ were extrapolated to zero linearly in m_1, while those obtained using $\Gamma_{\mathrm{imp}}(t)$ were extrapolated linearly in m_1^3.

We list our results for l_s in the tables. Observe that below the mobility edge l_s turns out to be always larger than l_ℓ. The physical significance of these results is discussed below.

E. Separation distance of localized eigenmodes

We now return to the notion of *isolated* localized eigenmodes mentioned above. The use of Eq. (3.9) depends on neglecting interference among the modes appearing in Eq. (3.8), which is only justified when m_1 is sufficiently

small. We must check whether our smallest values of m_1 are indeed sufficiently small.

Spectral sums as in Eq. (3.9) show that m_1 is the resolution with which we detect eigenmodes near λ. Let us compare the average support and localization lengths l_s and l_ℓ to the mean distance between eigenmodes detected at this resolution. Our smallest value of m_1 is 0.01. The number of eigenmodes with eigenvalues near λ that we detect for a typical gauge configuration is thus $N(\lambda) \simeq 0.01 V \rho(\lambda)$. Hence we define

$$R(\lambda) \equiv [V/N(\lambda)]^{1/4} \simeq [0.01 \rho(\lambda)]^{-1/4} \qquad (5.19)$$

as a measure of the average separation between eigenmodes detected at this resolution. Results for $R(\lambda)$ are shown in the tables. [We do not quote errors but they can be easily worked out from the $\rho(\lambda)$ data.] If $l_s \ll R$ the eigenmodes are isolated, and correlation functions reflect properties of individual localized eigenmodes, with no interference.[9] Thus, our method for the extraction of the average localization and support lengths is valid.

We can also directly estimate the overlap among any two eigenmodes to see whether they interfere. This overlap satisfies the inequality

$$|\langle \Psi_1 | \Psi_2 \rangle|^2 \leq \left(a^4 \sum_x |\Psi_1(x)| |\Psi_2(x)| \right)^2. \qquad (5.20)$$

We assume $\lambda_1 \approx \lambda_2 \approx \lambda$. The "edge" of the support of Ψ_2 closest to the center of Ψ_1 is at an average distance of around $R - l_s/2$. Using Eq. (3.15) this gives

$$|\langle \Psi_1 | \Psi_2 \rangle|^2 \lesssim \exp[-(R - l_s/2)/l_\ell]. \qquad (5.21)$$

A review of the tables shows the following.[10] Leaving out the last value of λ just below the mobility edge we find that for all other $\lambda < \lambda_c$ one has $l_s < R/2$ and $l_\ell < R/5$. Hence $|\langle \Psi_1 | \Psi_2 \rangle|^2 \lesssim \exp[-(3R/4)/(R/5)] = \exp(-15/4) \simeq 0.024$. This demonstrates that, except for the last value of λ just below λ_c, all the quantities we measured indeed reflect properties of individual localized eigenmodes, with no interference.

Above λ_c, as well as (in most cases) for the λ value just below λ_c, we have $R \approx l_s \approx l_\ell$. This means that interference is not a negligible effect, and the values of l_s and l_ℓ no longer reflect the properties of individual eigenmodes.

F. The $1/m_1$ divergence of the two-point function: $p \neq 0$

Finally, we return to the Ward identity Eq. (2.20). As can be seen from Eq. (3.17), the localization alternative to Goldstone's theorem requires that the $1/m_1$ divergence

---

[9]Here it is significant that l_ℓ turns out to be smaller than l_s.
[10]This applies also to results obtained with improved gauge actions, to be discussed in Sec. VI below. We ignore those cases where data's for l_s are not available.

persist for a range of momenta, and that its coefficient depend smoothly on p.

In order to confirm this, we calculated the Fourier transform of $\Gamma(t)$,

$$\tilde{\Gamma}(\omega_n) = \sum_t \cos(\omega_n t)\Gamma(t) \qquad (5.22)$$

where $\omega_n = 2\pi n/16$, and extrapolated $m_1\tilde{\Gamma}(\omega_n)$ linearly to $m_1 = 0$. [The justification for the linear extrapolation is the same as for Eq. (4.4). We included data for $m_1 = 0.005$ in the fit.] The calculation, which provides a consistency check, was done for $\beta = 5.7$, $m_0 = -1.5$, $\lambda = 0$, where the higher spectral density (compared to $\beta = 6.0$) makes it easier to obtain good statistics; the mobility edge here is still far from zero. We plot the results in Fig. 8. The ω dependence of the $1/m_1$ divergence is indeed smooth, as indicated in Eq. (3.17) for small ω.

For comparison, we repeated the calculation for $\lambda = 0.5$, which lies above λ_c. The extrapolation of $m_1\tilde{\Gamma}(\omega_n = 0)$ to $m_1 = 0$ is again straightforward, as it must be since this gives $\rho(\lambda)$ according to Eq. (4.10). Doing the same with $\tilde{\Gamma}(\omega_n \neq 0)$, however, leads to a huge χ^2, showing that Eq. (3.17) is inapplicable. This highlights the qualitative difference between $|\lambda| < \lambda_c$ and $|\lambda| > \lambda_c$.

VI. IMPROVED GAUGE ACTIONS

As an alternative to the Wilson plaquette action, we also studied pure-gauge ensembles generated by the Iwasaki [18] and DBW2 [20] actions, which have been used in DWF [22–24] and overlap [25,26] simulations. We set the Wilson mass to $m_0 = -1.8$ (see Sec. VII for explanation of
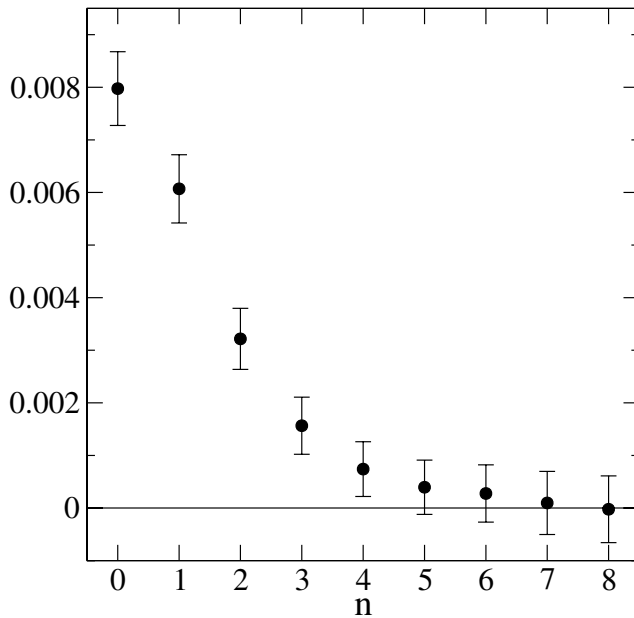

FIG. 8.   Coefficient of the $1/m_1$ divergence in $\tilde{\Gamma}(\omega_n)$ for $\beta = 5.7$, $m_0 = -1.5$, $\lambda = 0.0$.

TABLE X.   Spectral properties for Iwasaki gauge action, $\beta = 2.6$, $m_0 = -1.8$. The mobility edge is at $\lambda_c \simeq 0.38$. Number of configurations ranges from 600 at $\lambda = 0.0$ to 100 at $\lambda = 0.5$.

	l_ℓ	$\mathcal{D}$	l_s	R
0.0^a	$7(3) \times 10^{-7}$	$0.72(1)$	$\dots$	$\dots$	110
0.1^a	$1.8(5) \times 10^{-5}$	$0.73(1)$	$4(1) \times 10^{-6}$	$4(2)$	50
0.2^a	$2.3(9) \times 10^{-4}$	$0.80(1)$	$1.3(7) \times 10^{-4}$	$2(1)$	25
0.3^a	$9(2) \times 10^{-4}$	$1.20(2)$	$2.7(8) \times 10^{-4}$	$3(1)$	18
0.4^a	$0.111(2)$	∞	$\dots$	$\dots$	5.5
0.5^a	$0.277(5)$	∞	$0.040(4)$	$7(7)$	4.3

[a]Results from improved estimator.

TABLE XI.   Spectral properties for Iwasaki gauge action, $\beta = 2.2$, $m_0 = -1.8$. The mobility edge is at $\lambda_c \simeq 0.27$. Measurements on 200 gauge configurations.

λ	ρ	l_ℓ	$\mathcal{D}$	l_s	R
0.0^a	$0.0063(2)$	$0.80(1)$	$0.0021(1)$	$3.0(2)$	11
0.1^a	$0.0095(3)$	$0.91(2)$	$0.0032(2)$	$3.0(2)$	10
0.2	$0.034(2)$	$1.40(5)$	$0.010(2)$	$3.4(7)$	7.4
0.3	$0.120(5)$	∞	$0.025(4)$	$5(1)$	5.4
0.4	$0.27(1)$	∞	$0.017(7)$	$16(10)$	4.4

[a]Results from improved estimator.

TABLE XII.   Spectral properties for DBW2 gauge action, $\beta = 1.04$, $m_0 = -1.8$. The mobility edge is at $\lambda_c \simeq 0.39$. Between 200 and 400 gauge configurations were used.

λ	ρ	l_ℓ	$\mathcal{D}$	l_s	R
0.0^a	$\dots$	$0.74(1)$	$\dots$	$\dots$	$\dots$
0.1^a	$2(1) \times 10^{-7}$	$0.75(1)$	$\dots$	$\dots$	150
0.2^a	$6(2) \times 10^{-6}$	$0.82(1)$	$\dots$	$\dots$	65
$0.3^{a,b}$	$3.4(7) \times 10^{-4}$	$1.2(2)$	$\dots$	$\dots$	23
0.4	$0.058(2)$	∞	$\dots$	$\dots$	6

[a]Results from improved estimator.
[b]Extrapolation with Eq. (4.8) failed. We added terms linear and quadratic in m_1.

TABLE XIII.   Spectral properties for DBW2 gauge action, $\beta = 0.79$, $m_0 = -1.8$. The mobility edge is at $\lambda_c \simeq 0.32$. Measurements on 200 gauge configurations.

λ	ρ	l_ℓ	$\mathcal{D}$	l_s	R
0.0^a	$0.0013(1)$	$0.70(1)$	$0.00041(5)$	$3.2(5)$	17
0.1^a	$0.0023(2)$	$0.73(1)$	$0.0009(1)$	$2.6(4)$	14
$0.2^{a,b}$	$0.0064(4)$	$0.91(1)$	$0.0024(2)$	$2.7(3)$	11
0.3^a	$0.058(1)$	$2.1(1)$	$0.012(1)$	$4.8(4)$	6.4
0.4	$0.22(1)$	∞	$0.030(5)$	$7(1)$	4.6

[a]Results from improved estimator.
[b]Extrapolation with Eq. (4.8) failed. We added an m_1^2 term.

this choice). For each action, we studied two values of the lattice spacing, $a^{-1} \simeq 2$ GeV and $a^{-1} \simeq 1$ GeV. The corresponding bare couplings are $\beta = 2.6$ and 2.2, respectively, for the Iwasaki action [19]. For the DBW2 action they are $\beta = 1.04$ [23] and 0.79 respectively; the latter was determined by interpolation using β-values from Ref. [21].

We present our results in Tables X, XI, XII, and XIII and Figs. 9 and 10. A comparison of the (inverse squared) localization length for the three gauge actions at 2 GeV is shown in Fig. 9. The Iwasaki and DBW2 localization lengths are almost the same, an indication of the good scaling properties of the two actions. The mobility edge is almost the same for all three actions. Our results at 1 GeV are compared in Fig. 10. There is a gradual decrease in the localization length from Wilson to Iwasaki to DBW2. Differences in the value of the mobility edge, while bigger than in the 2 GeV case, remain small.

As can be seen from Tables X, XI, XII, and XIII, the main difference among the three actions is a dramatic reduction in the low-λ spectral density as we go from the Wilson action to the Iwasaki and then to the DBW2 gauge actions. What distinguishes the three actions is the coefficient of the 1×2 rectangle, which is zero for the standard plaquette action and in a common parametrization [22,23] is $c = -0.331$ for the Iwasaki action and $c = -1.4069$ for the DBW2 action. We see that the decrease in the low-λ spectral density is correlated with a more negative rectangle coefficient c. The rectangle term in the action sup-
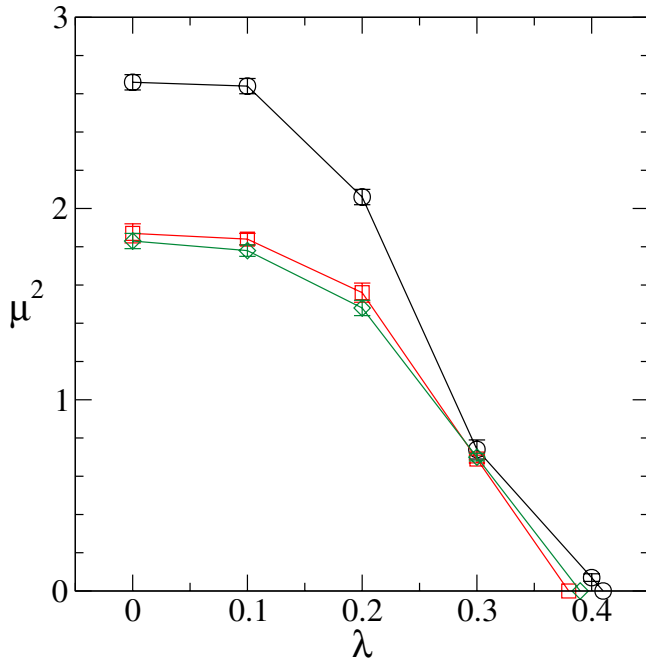

FIG. 10 (color online). As in Fig. 9, but at $a^{-1} \simeq 1$ GeV: Wilson action at $\beta = 5.7$, with $m_0 = -1.5$ (circles); Iwasaki action at $\beta = 2.2$, with $m_0 = -1.8$ (squares); DBW2 action at $\beta = 0.79$, with $m_0 = -1.8$ (diamonds).
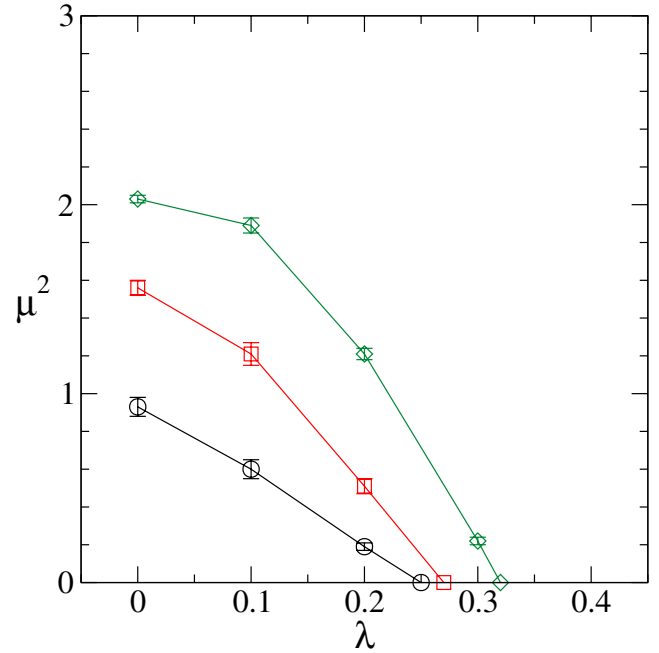
presses the small-size dislocations [20,23] responsible for the existence of the near-zero modes.

At (or just above) the mobility edge the spectral density remains of the same order as for the plaquette action: $\rho(\lambda_c) \approx 0.1$. This means that, unlike the plaquette action, the Iwasaki and DBW2 spectral densities rise steeply as λ approaches the mobility edge from below. At 2 GeV, a comparison of the spectral-density values just above and below the mobility edge shows a jump by 2 orders of magnitude for both improved actions.

VII. DISCUSSION

Our numerical results bear, first, on the phase structure of the supercritical Wilson operator in various ensembles; and, second, on the overlap and DWF theories based on the Wilson operator. We examine each of these in turn.

A. The Aoki phase diagram

The Aoki phase diagram has dictated our list of (β, m_0) values for calculation, and our results in turn yield information about the diagram. We can now discuss in more detail the choice of (β, m_0) values we have made.

There is considerable evidence (see e.g. Ref. [13]) that, for plaquette actions with $\beta = 5.7$ to 6.0, the low-λ spectral density of the Wilson operator varies slowly over the range $-2.0 < m_0 < -1.5$. In view of this, we have chosen to carry out most of our calculations with the plaquette action for the single value $m_0 = -1.5$. We have tested this

FIG. 9 (color online). The inverse squared localization length μ^2 vs λ, as in Fig. 6. Comparison of three gauge actions at $a^{-1} \simeq 2$ GeV: Wilson action at $\beta = 6.0$, with $m_0 = -1.5$ (circles); Iwasaki action at $\beta = 2.6$, with $m_0 = -1.8$ (squares); DBW2 action at $\beta = 1.04$, with $m_0 = -1.8$ (diamonds).

insensitivity to m_0 by considering as well $\beta = 5.7$, $m_0 = -2.0$. We see in Fig. 7 that the localization lengths l_ℓ for all λ are almost unchanged as m_0 changes from -1.5 to -2.0; hence, the mobility edge is unchanged as well. The rough equality extends to the support length and to the spectral density for λ sufficiently below the mobility edge, which means in practice $\lambda \lesssim 0.1$ (see Tables III and VII as well as Fig. 4).

Old results [16] on the Aoki phase diagram indicate that the point ($\beta = 6.0$, $m_0 = -2.4$) lies in the second "finger" of the Aoki phase. Our results show that the point $(5.7, -2.4)$ is indeed at a boundary of this finger or inside it.

Looking ahead to DWF, our results indicate that, for $\beta \gtrsim 6.0$, one can use any value of m_0 in the range $[-2.0, -1.5]$. (The same applies to the overlap construction.) Mean-field arguments suggest that the optimal m_0 (related to the "domain-wall height" M via $M = -m_0$) for DWF simulations is roughly $m_0 \approx -1 + m_c(g_0)$ [32]. The commonly used value at $a^{-1} \simeq 2$ GeV is $M = -m_0 = 1.8$. This is consistent with the mean-field estimate and lies comfortably in the middle of the range of m_0 we considered. In view of our evidence for the insensitivity of the key spectral properties over this range, we have chosen the value $m_0 = -1.8$ also for the improved-action calculations.

Our scan of β values for $m_0 = -1.5$ places the boundary of the Aoki phase of the plaquette action between $\beta = 5.5$ and $\beta = 5.6$ (cf. [16]).

B. Implications for overlap fermions

The overlap operator is constructed explicitly in terms of the Wilson operator H_W,

$$D_{\text{ov}} = 1 - \gamma_5 \frac{H_W}{|H_W|}. \qquad (7.1)$$

Our knowledge of the spectrum of H_W is thus of immediate import to overlap calculations; in particular, it is critical to understanding the locality of the operator. The long-distance tail of D_{ov} has two components [7,8], and both must be kept under control. One component comes from the near-zero modes; its range is the localization length $l_\ell(\bar{\lambda})$, where $\bar{\lambda}$ will be defined shortly. The other component comes from modes near the mobility edge, and its range is the inverse of λ_c.

We denote by l_{ov} the range of the overlap operator: For $|x - y| \gg a$ one has $|D_{\text{ov}}(x, y)| \sim \exp(-|x - y|/l_{\text{ov}})$. In Ref. [7] we discussed the different roles of the extended and the localized modes in determining l_{ov}, and arrived at an estimate for the contribution of the localized spec-

trum.[11] Here we revisit the issue in the light of our numerical findings.

In the treatment of the localized eigenmodes the essential assumption was the absence of interference between different modes. This should be true, in particular, for the low-lying modes that make up the *mutually isolated* subspace. These are all the localized eigenmodes in the interval $[-\bar{\lambda}, \bar{\lambda}]$ where $0 < \bar{\lambda} < \lambda_c$ is chosen such that for any two eigenvalues $\lambda_m, \lambda_n \in [-\bar{\lambda}, \bar{\lambda}]$, the corresponding localized eigenmodes are isolated (on average) in the sense of Sec. V E. The contribution of the mutually isolated subspace to the ensemble-averaged overlap operator is then estimated to be[12]

$$\langle |D_{\text{ov}}(x, y)| \rangle_{|\lambda| \leq \bar{\lambda}} \approx \int_{-\bar{\lambda}}^{\bar{\lambda}} d\lambda \rho(\lambda) \exp\left(-\frac{|x - y|}{2l_\ell(\lambda)}\right). \quad (7.2)$$

Since both $\rho(\lambda)$ and $l_\ell(\lambda)$ are monotonically increasing, the integral is dominated by its upper limit, viz.

$$\langle |D_{\text{ov}}(x, y)| \rangle_{|\lambda| \leq \bar{\lambda}} \approx \bar{\lambda} \rho(\bar{\lambda}) \exp\left(-\frac{|x - y|}{2l_\ell(\bar{\lambda})}\right). \qquad (7.3)$$

This is the first of the two pieces determining l_{ov}.

We may arrive at a numerical estimate as follows. Consider for definiteness our results for the Wilson gauge action at $\beta = 6.0$, shown in Table I. A conservative guess for the edge of the mutually isolated spectrum is $\bar{\lambda} \approx 0.2$. Since $\rho(0.2)$ is of the order of 10^{-3}, we find that the integrated spectral density $\int_{-\bar{\lambda}}^{\bar{\lambda}} d\lambda \rho(\lambda)$ is of the order of 10^{-4}. The total number of eigenvalues in the interval $[-\bar{\lambda}, \bar{\lambda}]$ is given by the volume V times this integral, and so the mean space-time separation $\bar{R}$ between any two eigenmodes in this spectral interval[13] is $\bar{R} \simeq 10$. Using similar considerations to Sec. V E, this corresponds to $|\langle \Psi_n | \Psi_m \rangle|^2 = O(0.01)$, which is certainly a small overlap. All the modes in this spectral interval are, therefore, mutually isolated.

Evaluating Eq. (7.3) for $\bar{\lambda} = 0.2$ gives

$$\langle |D_{\text{ov}}(x, y)| \rangle_{|\lambda| \leq \bar{\lambda}} \approx 10^{-4} \exp\left(-\frac{|x - y|}{1.4}\right). \qquad (7.4)$$

This semiquantitative estimate highlights two points concerning $\bar{\lambda}$. First, the small prefactor signifies a large value of $\bar{R}$. The latter must be appreciably larger than the support length l_s in order for the modes to be mutually isolated, and $\bar{\lambda}$ must be chosen to make it so. Moreover, pushing $\bar{\lambda}$ to a somewhat higher value would also increase $l_\ell(\bar{\lambda})$ and thus modify the coefficient in the exponent. Equation (7.4) still

---

[11]This estimate complements the discussion of Ref. [8]. The analytical techniques of Ref. [8] are inapplicable to the localized spectrum because it becomes dense in the infinite-volume limit.

[12]The introduction of $\bar{\lambda}$ improves on the discussion of Ref. [7], where the rather arbitrary cutoff $\lambda_c/2$ was imposed on the integral [see, in particular, Eq. (6.3) therein].

[13]Note that $\bar{R}$ measures the average separation between *all* modes in the interval $[-\bar{\lambda}, \bar{\lambda}]$, whereas $R(\lambda)$ defined in Sec. V E measures the average separation between modes in an $O(m_1)$ neighborhood around λ.

provides a rough idea of the contribution of the mutually isolated localized modes.

The modes with $|\lambda_n| \geq \bar{\lambda}$ are of two types: the localized eigenmodes in the interval $\bar{\lambda} \leq |\lambda_n| \leq \lambda_c$, and the extended modes above λ_c. The rise in the spectral density near and above λ_c leads us to expect that the resulting contribution to l_{ov} will be controlled by λ_c itself. For these modes the analysis of Ref. [8] applies, and a comparison of our results to theirs supports this conjecture. For $\beta = 6.0$ and $m_0 = -1 - s = -1.4$ and -1.6, Table I in Ref. [8] reports $l_{\mathrm{ov}}^{-1} \simeq 0.49$ and 0.45, respectively. This can be compared to our result for the mobility edge at $m_0 = -1.5$, namely, $\lambda_c \simeq 0.41$. Given the differences in the details, the close agreement may be a numerical coincidence. Still, this suggests that the contribution of the rest of the spectrum may indeed be dominated by the extended modes just above the mobility edge, and that

$$\langle|D_{\mathrm{ov}}(x, y)|\rangle_{|\lambda| \geq \bar{\lambda}} \approx C \exp(-\lambda_c|x - y|), \qquad (7.5)$$

where $C = O(1)$. This is the second contribution to the tail of D_{ov}.

As explained above, the prefactor in Eq. (7.4) is bound to be small. Comparing Eqs. (7.4) and (7.5) suggests that l_{ov} is dominated by the mobility edge and not by the near-zero modes. This is consistent with the results of Ref. [33]. There it is found that changing the gauge action from Wilson to Iwasaki or to DBW2 at quenched $a^{-1} \simeq 2$ GeV has little effect on l_{ov}. Indeed, while the spectral density of the localized modes is a sensitive function of the gauge action, the mobility edge itself varies little among the three gauge actions (see Tables X, XI, XII, and XIII and Fig. 9). Of course, if $2l_\ell(\bar{\lambda}) > \lambda_c^{-1}$ the small contribution of the near-zero modes will dominate the asymptotic tail of D_{ov} at large distances. For all cases we study here, however, the opposite inequality holds.

The overlap kernel's decay rate l_{ov}^{-1} can be interpreted as the mass of unphysical degrees of freedom. In practice, these degrees of freedom can be uncomfortably light. Assuming $l_{\mathrm{ov}} \approx \lambda_c^{-1}$, our data allow us to estimate their mass for the pure-gauge ensembles. When the cutoff is $a^{-1} \simeq 2$ GeV, we find $l_{\mathrm{ov}}^{-1} \simeq 0.4 \times 2\mathrm{GeV} = 800$ MeV for all three gauge actions.[14] This is presumably a high enough scale to qualify as part of the discretization errors. For $a^{-1} \simeq 1$ GeV, on the other hand, we obtain $l_{\mathrm{ov}}^{-1} \simeq 0.25 \times 1$ GeV $= 250$ MeV for the Wilson action, 270 MeV for Iwasaki, and 320 MeV for DBW2. This is an alarmingly low scale for unphysical degrees of freedom.[15]

---

[14]This rough equality among the actions is in agreement with the above mentioned results of Ref. [33].

[15]Note, in particular, the 1 GeV overlap simulations of Ref. [25].

C. Implications for domain-wall fermions

Our discussion of DWF will be less detailed for two reasons. First, the theoretical background has already been discussed in detail [7,34–36]. Also, the relevant "Hamiltonian" for DWF is not H_W itself but the logarithm of the fifth-dimension transfer matrix, a different (though closely related) operator.

DWF achieve exact chirality when the number of sites L_5 in the fifth dimension tends to infinity. DWF simulations are performed at finite L_5, typically in the range of 10–20. The main question is what is the size of chiral symmetry violations due to the finiteness of L_5. A quantitative measure of these violations is provided by the residual mass m_{res} [22,23,36], which is the small additive correction to the quark mass determined from the partially conserved axial-vector current (PCAC) relation [2]. (Alternatively, m_{res} can be determined from the extrapolation of the pion mass to the chiral limit.) Here, too, there are two terms that can be ascribed to extended and localized modes,

$$m_{\mathrm{res}} \approx C_1 \exp(-\tilde{\lambda}_c L_5) + \frac{C_2}{L_5}, \qquad (7.6)$$

where $C_1 = O(1)$ comes from the extended modes near the mobility edge and

$$C_2 \approx L_5 \int_{-1/L_5}^{1/L_5} d\lambda \tilde{\rho}(\lambda) \approx \tilde{\rho}(1/L_5) \qquad (7.7)$$

comes from the low-lying, localized modes. Because of the rapidly growing spectral density, the localized modes' contribution is dominated by modes with $|\lambda_n| \approx 1/L_5$; we ignore a power-law correction to the extended modes' contribution.[16] The tildes indicate spectral quantities of the new "Hamiltonian,"

$$\tilde{H} = -\log(T^2)/(2a_5), \qquad (7.8)$$

where $T(a_5)$ is the transfer matrix for hopping in the fifth direction and a_5 is the corresponding lattice spacing (conventional DWF have $a_5 = 1$).

The "Hamiltonians" H_W and $\tilde{H}$ share identical zero modes [2]. Thus we expect that for the near-zero modes, the spectral density (and other properties) of $\tilde{H}$ is fairly close to that of H_W. Replacing $\tilde{\rho}(\lambda)$ by $\rho(\lambda)$ in Eq. (7.7) should yield a reasonably good approximation. On the other hand, further away from $\lambda = 0$ the spectra of $\tilde{H}$ and H_W need not be equal. In particular the mobility edges

---

[16]Equation (7.7), derived from Appendix C.2 of Ref. [34], provides a slightly better estimate of the contribution of the localized modes than that given in Ref. [35]. Equation (7.6) corrects the discussion of the near-zero modes' contribution to m_{res} given in Ref. [7]. In particular, Eq. (6.12) therein as well as the first term on the right-hand side of Eq. (7.1) there are erroneous. The limits of integration in Eq. (7.7) here come of demanding that $\exp(-\lambda L_5) = O(1)$ for modes included in the integral.

could be quite different. Nonetheless, again because of the identity of the zero modes, the mobility edges of $\tilde{H}$ and H_W reach zero simultaneously. Thus the same Aoki phase defines the forbidden region for both overlap and DWF.

As noted, the numerical results of this paper are relevant for determining the near-zero modes' contribution to m_{res}, but separate calculations would be needed to determine $\tilde{\lambda}_c$, the mobility edge of $\tilde{H}$, which governs the extended modes' contribution to m_{res}. One can alternatively determine these quantities by calculating the residual mass as a function of L_5 [22,23,36] and fitting to Eq. (7.6).

Equation (7.6) shows that, unlike overlap fermions, the physics of finite-L_5 DWF simulation and, in particular, the value of m_{res}, are sensitive functions of the near-zero modes of $\tilde{H}$ (or of H_W). In quenched DWF simulations, using the Iwasaki or DBW2 actions allows reaching negligibly small values of m_{res} [22,23]. For overlap simulations, on the other hand, these gauge actions are advantageous for a technical reason: overlap simulations require an exact treatment (within numerical precision) of all the Wilson eigenvalues in a certain interval $[-\delta, \delta]$, and reducing the number of modes in this interval speeds up the simulation. We note that there exist versions of DWF, in particular, the so-called Möbius fermions [37], where the near-zero modes' contribution to m_{res} decreases much faster with L_5 (see also [38–40]).

D. SUMMARY

The continuum limit of lattice QCD with either DWF or overlap fermions can be taken while letting $m_0 \to -1$. Since all correlation lengths associated with the Wilson operator itself remain finite in lattice units, there is little doubt that the continuum limit is correct, and that no unphysical excitations can survive it. Issues addressed in this paper have to do with MC simulations at finite lattice spacing.

Using Green function techniques, we have determined the mobility edge of the Hermitian Wilson operator for a number of pure-gauge ensembles with plaquette, Iwasaki and DBW2 gauge actions. Our results allow mapping a portion of the (quenched) Aoki phase diagram. Where the mobility edge is nonzero, we have also characterized the localized spectrum in terms of an average support length and an average localization length, the latter determined from the asymptotic decay rate of the mode density.

Our results are of direct relevance to the overlap operator. For the near-zero modes, or, more precisely, for the mutually isolated subspace of the localized modes, we found that the localization length is consistently smaller than the support length. We have also found that twice the localization length is not bigger than the inverse mobility edge. Together, these findings imply that the near-zero modes play little role in setting the range of the overlap operator.

We argue that the range of the overlap operator is set by, and is roughly equal to, the inverse mobility edge. Our results for the mobility edge suggest that it is fairly safe to perform (quenched) overlap simulations at $a^{-1} = 2$ GeV. The same is not true when the cutoff is 1 GeV; for all three gauge actions we find that the standard overlap operator is likely to contain unphysical quarklike degrees of freedom as light as 250–300 MeV. This casts serious doubt on the validity of 1 GeV overlap simulations. In general, the determination of the range of the overlap operator is an important test that must be carried out for any new overlap simulation.

Closely related to the mobility edge is the mass of the lowest pseudoscalar excitation of the Wilson-fermion action—Wilson's original would-be pion. It, too, is a nonphysical excitation where the overlap operator is concerned. Where we demand that couplings be chosen such that the mobility edge is far from zero, the same can be said of the mass of the Wilson pion.[17]

Obtaining the corresponding information for DWF will require the study of a different "Hamiltonian," the logarithm of the fifth-dimension transfer matrix, $\tilde{H}$. In particular, it is important to study the range of the effective four-dimensional operator D_{eff} obtained by integrating out the five-dimensional bulk modes and the pseudofermions [38], and to determine how the range of D_{eff} is affected by $\tilde{H}$ as a function of L_5 and a_5. The near-zero spectrum of $\tilde{H}$ is similar to that of the Wilson operator, and we are thus able to confirm the picture that the near-zero modes (of either H_W or $\tilde{H}$) make a major contribution to the residual mass.

Last, we note that numerical simulations with dynamical DWF [24] (or dynamical overlap fermions) require the largest available computers. Not only production runs, but also exploratory runs are very expensive. The process of closing in on an optimal set of simulation parameters can greatly benefit from the (low-cost) determination of the quantities studied here—the mobility edge and the characteristics of the localized modes.

---

[17]If it turns out that the A and C regions in Fig. 1 are separated by first-order transitions instead of fingers, the Wilson pion may still have a very small mass near the transitions [17,41].

[1] D. B. Kaplan, Phys. Lett. B **288**, 342 (1992).

[2] Y. Shamir, Nucl. Phys. **B406**, 90 (1993); V. Furman and Y. Shamir, Nucl. Phys. **B439**, 54 (1995).

[3] R. Narayanan and H. Neuberger, Phys. Lett. B **302**, 62 (1993); Nucl. Phys. **B412**, 574 (1994); **B443**, 305 (1995).

[4] H. Neuberger, Phys. Lett. B **417**, 141 (1998).

[5] P. H. Ginsparg and K. G. Wilson, Phys. Rev. D **25**, 2649 (1982); H. Neuberger, Phys. Lett. B **427**, 353 (1998); M. Lüscher, Phys. Lett. B **428**, 342 (1998).

[6] R. C. Brower and B. Svetitsky, Phys. Rev. D **61**, 114511 (2000); F. Berruto, R. C. Brower, and B. Svetitsky, Phys. Rev. D **64**, 114504 (2001).

[7] M. Golterman and Y. Shamir, Phys. Rev. D **68**, 074501 (2003).

[8] P. Hernández, K. Jansen, and M. Lüscher, Nucl. Phys. **B552**, 363 (1999).

[9] M. Golterman and Y. Shamir, J. High Energy Phys. 09 (2000) 006.

[10] T. Banks and A. Casher, Nucl. Phys. **B169**, 103 (1980).

[11] S. Aoki, Phys. Rev. D **30**, 2653 (1984); **33**, 2399 (1986); **34**, 3170 (1986); Phys. Rev. Lett. **57**, 3136 (1986); Nucl. Phys. Proc. Suppl. **60A**, 206 (1998).

[12] A. J. McKane and M. Stone, Ann. Phys. (N.Y.) **131**, 36 (1981).

[13] R. G. Edwards, U. M. Heller, and R. Narayanan, Nucl. Phys. **B522**, 285 (1998); **B535**, 403 (1998); Phys. Rev. D **60**, 034502 (1999).

[14] F. Berruto, R. Narayanan, and H. Neuberger, Phys. Lett. B **489**, 243 (2000).

[15] M. Golterman and Y. Shamir, Nucl. Phys. Proc. Suppl. **129**, 149 (2004).

[16] S. Aoki and A. Gocksch, Phys. Lett. B **231**, 449 (1989); Phys. Rev. D **45**, 3845 (1992); S. Aoki, T. Kaneda, and A. Ukawa, *ibid.* **56**, 1808 (1997).

[17] M. Golterman, S. Sharpe, and R. Singleton, Phys. Rev. D **71**, 094503 (2005).

[18] Y. Iwasaki, UTHEP-117, 1983; UTHEP-118, 1983; Y. Iwasaki and T. Yoshie, Phys. Lett. B **143**, 449 (1984).

[19] Y. Iwasaki, K. Kanaya, T. Kaneko, and T. Yoshie, Phys. Rev. D **56**, 151 (1997).

[20] P. de Forcrand *et al.* (QCD-TARO Collaboration), Nucl. Phys. **B577**, 263 (2000).

[21] A. Borici and R. Rosenfelder, Nucl. Phys. Proc. Suppl. **63**, 925 (1998).

[22] A. Ali Khan *et al.* (CP-PACS Collaboration), Phys. Rev. D **63**, 114504 (2001); Nucl. Phys. Proc. Suppl. **94**, 725 (2001).

[23] K. Orginos *et al.* (RBC Collaboration), Nucl. Phys. Proc. Suppl. **106**, 721 (2002); Y. Aoki *et al.* (RBC Collaboration), Phys. Rev. D **69**, 074504 (2004).

[24] Y. Aoki *et al.* (RBC Collaboration), hep-lat/0411006.

[25] Y. Chen *et al.*, Phys. Rev. D **70**, 034502 (2004).

[26] D. Galletly *et al.* (QCDSF-UKQCD Collaboration), Nucl. Phys. Proc. Suppl. **129**, 453 (2004).

[27] R. Frezzotti, P. A. Grassi, S. Sint, and P. Weisz, J. High Energy Phys. 08 (2001) 058.

[28] M. Golterman, Y. Shamir, and B. Svetitsky, Phys. Rev. D **71**, 071502 (2005).

[29] K. Jansen, C. Liu, H. Simma, and D. Smith, Nucl. Phys. Proc. Suppl. **53**, 262 (1997).

[30] S. Aoki and Y. Taniguchi, Phys. Rev. D **65**, 074502 (2002); S. Aoki, Nucl. Phys. Proc. Suppl. **A109**, 70 (2002); S. Aoki *et al.* (CP-PACS Collaboration), Nucl. Phys. Proc. Suppl. **106**, 718 (2002).

[31] D. J. Thouless, Phys. Rep. **13**, 93 (1974).

[32] T. Blum, A. Soni, and M. Wingate, Phys. Rev. D **60**, 114507 (1999); S. Aoki, T. Izubuchi, Y. Kuramashi, and Y. Taniguchi, Phys. Rev. D **62**, 094502 (2000).

[33] T. DeGrand *et al.* (MILC collaboration), Phys. Rev. D **63**, 034503 (2001); T. G. Kovacs, *ibid.* **67**, 094501 (2003); T. DeGrand, A. Hasenfratz, and T. G. Kovacs, *ibid.* **67**, 054501 (2003).

[34] Y. Shamir, Phys. Rev. D **62**, 054513 (2000).

[35] M. Golterman and Y. Shamir, Phys. Rev. D **71**, 034502 (2005).

[36] T. Blum *et al.* (RBC Collaboration), Phys. Rev. D **69**, 074502 (2004).

[37] R. C. Brower, H. Neff, and K. Orginos, hep-lat/0409118.

[38] H. Neuberger, Phys. Rev. D **57**, 5417 (1998); Y. Kikukawa and T. Noguchi, hep-lat/9902022.

[39] A. Borici, Nucl. Phys. Proc. Suppl. **83**, 771 (2000); hep-lat/9912040.

[40] K. Jansen and K. Nagai, J. High Energy Phys. 12 (2003) 038.

[41] S. R. Sharpe and R. J. Singleton, Phys. Rev. D **58**, 074501 (1998).

[42] E. M. Ilgenfritz *et al.*, Phys. Rev. D **69**, 074511 (2004); F. Farchioni *et al.*, Eur. Phys. J. C **39**, 421 (2005).

[43] Available from http://www.physics.utah.edu/~detar/milc/