KEK-TH-1769, RBRC 1095, DAMTP-2014-86

Domain wall QCD with physical quark masses

T. Blum,[1,2] P.A. Boyle,[3] N.H. Christ,[4] J. Frison,[3] N. Garron,[5,6] R.J. Hudspith,[7]
T. Izubuchi,[2,8] T. Janowski,[9] C. Jung,[8] A. Jüttner,[9] C. Kelly,[2] R.D. Kenway,[3]
C. Lehner,[8] M. Marinkovic,[9,10] R.D. Mawhinney,[4] G. McGlynn,[4] D.J. Murphy,[4]
S. Ohta (太田滋生),[11,12,2] A. Portelli,[9] C.T. Sachrajda,[9] and A. Soni[8]

(RBC and UKQCD Collaborations)

[1]*Physics Department, University of Connecticut, Storrs, CT 06269-3046, USA*

[2]*RIKEN-BNL Research Center, Brookhaven National Laboratory, Upton, NY 11973, USA*

[3]*SUPA, School of Physics, The University of Edinburgh, Edinburgh EH9 3JZ, UK*

[4]*Physics Department, Columbia University, New York, NY 10027, USA*

[5]*Department of Applied Mathematics & Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK*

[6]*On leave from School of Mathematics, Trinity College Dublin, College Green, Dublin 2, Ireland*

[7]*Department of Physics and Astronomy, York University, Toronto, Ontario, M3J 1P3, Canada*

[8]*Department of Physics, Brookhaven National Laboratory, Upton, NY 11973, USA*

[9]*School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK*

[10]*CERN, Physics Department, 1211 Geneva 23, Switzerland*

[11]*Institute of Particle and Nuclear Studies, KEK, Tsukuba, Ibaraki, 305-0801, Japan*

[12]*Department of Particle and Nuclear Physics, Sokendai Graduate University of Advanced Studies, Hayama, Kanagawa 240-0193, Japan*

PACS numbers: 11.15.Ha, 11.30.Rd, 12.15.Ff, 12.38.Gc 12.39.Fe

ABSTRACT

We present results for several light hadronic quantities (f_π, f_K, B_K, m_{ud}, m_s, $t_0^{1/2}$, w_0) obtained from simulations of 2+1 flavor domain wall lattice QCD with large physical volumes and nearly-physical pion masses at two lattice spacings. We perform a short, $\mathcal{O}(3)\%$, extrapolation in pion mass to the physical values by combining our new data in a simultaneous chiral/continuum 'global fit' with a number of other ensembles with heavier pion masses. We use the physical values of m_π, m_K and m_Ω to determine the two quark masses and the scale - all other quantities are outputs from our simulations. We obtain results with sub-percent statistical errors and negligible chiral and finite-volume systematics for these light hadronic quantities, including: $f_\pi = 130.2(9)$ MeV; $f_K = 155.5(8)$ MeV; the average up/down quark mass and strange quark mass in the $\overline{\mathrm{MS}}$ scheme at 3 GeV, $2.997(49)$ and $81.64(1.17)$ MeV respectively; and the neutral kaon mixing parameter, B_K, in the RGI scheme, $0.750(15)$ and the $\overline{\mathrm{MS}}$ scheme at 3 GeV, $0.530(11)$.

I. INTRODUCTION

The low energy details of the strong interactions, encapsulated theoretically in the Lagrangian of QCD, are responsible for producing mesons and hadrons from quarks, creating most of the mass of the visible universe, and determining a vacuum state which exhibits symmetry breaking. For many decades, the methods of numerical lattice QCD have been used to study these phenomena, both because of their intrinsic interest and because QCD effects are important for many precision tests of quark interactions in the Standard Model. Many theoretical and computational advances have been made during this time and, in this paper, we report on the first simulations of 2+1 flavor QCD (*i.e.* QCD including the fermion determinant for u, d and s quarks with $m_u = m_d$) with essentially physical quark masses using a lattice fermion formulation which accurately preserves the continuum global symmetries of QCD at finite lattice spacing: domain wall fermions (DWF). This isospin symmetric version of QCD requires three inputs to perform a simulation at a single lattice spacing: a bare coupling constant, a degenerate light quark mass ($m_u = m_d$), and a strange quark mass. We fix these using the physical values for m_π, m_K, and m_Ω. In particular, for a fixed bare coupling, adjusting $m_u = m_d$ and m_s until m_π/m_Ω and m_K/m_Ω take on their physical values leads to a determination of the lattice spacing, a, for this coupling. All other low energy quantities, such as f_π and f_K, are now predictions. By repeating this for different lattice spacings, physical predictions in the continuum limit ($a \to 0$) for other low energy QCD observables are obtained. In this work, we used results from our earlier simulations to estimate the input physical quark masses and then we make a modest correction in our results, using chiral perturbation theory and simple analytic ansatz, to adjust to the required quark mass values, a correction of less than 10% in the quark mass. These physical quark mass simulations would not have been possible without IBM Blue Gene/Q resources [1–4].

For the past decade, the RBC and UKQCD collaborations have been steadily approaching the physical quark mass point with a series of 2+1 flavor domain wall fermion simulations. Recently [5] we reported on a combined analysis of three of our domain wall fermion ensembles with the Shamir kernel, namely our $32^3 \times 64$ and $24^3 \times 64$ ensemble sets with the Iwasaki gauge action at $\beta = 2.25$ and $\beta = 2.13$ ($a^{-1} = 2.383(9)$ GeV and $1.785(5)$ GeV) and lightest unitary pion masses of $302(1)$ MeV and $337(2)$ MeV respectively, and our coarser $32^3 \times 64$ Iwasaki+DSDR ensemble set with $\beta = 1.75$ ($a^{-1} = 1.378(7)$ GeV) but substantially lighter pion masses of $143(1)$ MeV partially-quenched and $171(1)$ MeV unitary. We refer to these as our 32I, 24I and 32ID ensembles, respectively. (The lattice spacings and other results for these ensembles quoted here come from global fits that include the new, physical quark mass ensembles, as well as new observable measurements on these older ensembles. As such, central values have shifted from earlier published values, generally within the published errors. Also, the new errors are smaller, because of the increased data.) For the latter 32ID ensembles, the use of a coarser lattice represented a compromise between the need to simulate with a large physical volume in order to keep finite-volume errors under control in the presence of such light pions and the prohibitive cost of increasing the lattice size. The DSDR term was used to suppress the dislocations in the gauge field that dominate the residual chiral symmetry breaking in the domain wall formulation at strong coupling. The addition of this ensemble set resulted in a factor of two reduction in the chiral extrapolation systematic error over our earlier analysis of the Iwasaki ensembles alone (24I and 32I) [6], but the total errors on our physical predictions remained on the order of 2%. Now, combining algorithmic advances with the power of the latest generation of supercomputers, we are finally able to perform large volume simulations directly at the physical point without the need for such compromises.

In this paper we present an analysis of two 2+1 flavor domain wall ensembles simulated essentially at the physical point. The lattice sizes are $48^3 \times 96$ and $64^3 \times 128$ with physical volumes of $(5.476(12)\ \mathrm{fm})^3$ and $(5.354(16)\ \mathrm{fm})^3$ ($m_\pi L = 3.86$ and 3.78). Throughout this document we refer to these ensembles with the labels 48I and 64I respectively. We utilize the Möbius domain wall action tuned such that the Möbius and Shamir kernels are identical up to a numerical factor, which allows us to simulate with a smaller fifth dimension, and hence a lower cost, for the same physics. This is discussed in more detail in Section II. The values of L_s are 24 and 12 for the 48I and 64I ensembles respectively. For the 48I ensemble, L_s would have to be more than twice as large to achieve the same residual mass with the Shamir kernel. The corresponding residual masses, m_{res}, comprise $\sim 45\%$ of the physical light quark mass for the 48I ensemble, and $\sim 30\%$ for the 64I. We use the Iwasaki gauge action with $\beta = 2.13$ and 2.25, giving inverse lattice spacings of $a^{-1} = 1.730(4)$ GeV and 2.359(7) GeV, and the degenerate up/down quark masses were tuned to give (very nearly) physical pion masses of 139.2(4) MeV and 139.2(5) MeV.

We also introduce a third ensemble generated with Shamir domain wall fermions and the Iwasaki gauge action at $\beta = 2.37$, corresponding to an inverse lattice spacing of 3.148(17) GeV, with a lattice volume of $32^3 \times 64$ and with $L_s = 12$. The lightest unitary pion mass is 371(5) MeV. Although these masses are unphysically heavy, this ensemble provides a third lattice spacing for each of the measured quantities, allowing us to bound the $\mathscr{O}(a^4)$ errors on our final results. We label this ensemble 32Ifine.

We have taken full advantage of each of our expensive 48I and 64I gauge configurations by developing a measurement package that uses EigCG to produce DWF eigenvectors in order to deflate subsequent quark mass solves, and that uses the all-mode-averaging (AMA) technique of Ref. [7]. In AMA, quark propagators are generated on every timeslice of the lattice but with reduced precision, and then corrected with a small number of precise measurements. To reduce the fractional overhead of calculating eigenvectors and the large I/O demands of storing them, we share propagators between m_π, m_K, f_π, f_K, B_K, the K_{l3} form factor $f_+^{K\pi}(q^2=0)$ and the $K\to(\pi\pi)_{\mathrm{I}=2}$ amplitude. (The last two quantities are not reported here.) By putting so many measurements into a single job, the EigCG setup costs are only $\sim 20\%$ of the total time, and we find this approach speeds up the measurement of these quantities by between 5 and 25 times, depending on the observable. Here again the Blue Gene/Q has been invaluable, since it has a large enough memory to store the required eigenvectors and the reliability to run for sufficient time to use them in all of the above measurements. In Section III we present the results of these measurements.

As mentioned already, in order to correct for the minor differences between the simulated and physical pion masses, we perform a short chiral extrapolation. As these new 48I and 64I ensembles have essentially the same quark masses, we must include data with other quark masses in order to determine the mass dependences. We achieve this by combining the 64I and 48I ensembles with the aforementioned $32^3\times 64$ and $24^3\times 64$ Iwasaki gauge action ensemble sets (32I and 24I, respectively), and the $32^3\times 64$ Iwasaki+DSDR ensemble set (32ID), in a simultaneous chiral/continuum 'global fit'. We also include the new 32Ifine ensemble, to give us a third lattice spacing with the same action, to improve the continuum extrapolation. We note that these are the same kinds of fits we have used in our previous work with the 24I, 32I and 32ID ensembles - here we have the addition of very accurate data at physical quark masses. In addition, we also have added Wilson flow measurements of the scale on all of our ensembles to the global fits. While the Wilson flow scale in physical units is an output of our simulations, the relative values on the various ensembles provide additional accurate data that helps to constrain the lattice spacing determinations. In Section IV we discuss our fitting strategy in more detail and the fit results are presented in Section V.

Given the length of this paper and the many details discussed, we present a summary of our physical results in Table I as the last part of this introduction. These are continuum results for isospin symmetric 2+1 flavor QCD without electromagnetic effects. Our input values are m_π,

m_K, m_Ω, and the results in Table I are outputs from our simulations. For results quoted in the $\overline{\text{MS}}$ scheme, the first error is statistical and the second is the error from renormalization. For other quantities, the error is the statistical error. The other usual sources of error (finite volume, chiral extrapolation, continuum limit) have all been removed through our measurements and any error estimates we can generate for these possible systematic errors are dramatically smaller than the (already small) statistical error quoted. This is discussed at great length in Section V. The Conclusions section (Section VI) summarizes our results and gives comparisons of them with experiment and/or the results of other lattice simulations.

Quantity	Value
f_π	130.19 ± 0.89 MeV
f_K	155.51 ± 0.83 MeV
f_K/f_π	1.1945 ± 0.0045
$m_u = m_d(\overline{\text{MS}}, 3\text{ GeV})$	$2.997 \pm 0.036 \pm 0.033$ MeV
$m_s(\overline{\text{MS}}, 3\text{ GeV})$	$81.64 \pm 0.77 \pm 0.88$ MeV
$m_s/m_u = m_s/m_d$	27.34 ± 0.21
$t_0^{1/2}$	$0.7292 \pm 0.0041\ \text{GeV}^{-1}$
w_0	$0.8742 \pm 0.0046\ \text{GeV}^{-1}$
$B_K(\text{SMOM}(\not{q},\not{q}), 3\text{ GeV})$	0.5341 ± 0.0018
$B_K(\overline{\text{MS}}, 3\text{ GeV})$	$0.5293 \pm 0.0017 \pm 0.0106$
$\hat{B}_K$	$0.7499 \pm 0.0024 \pm 0.0150$
$L_4^{(2)}(\Lambda_{\chi PT} = 1\text{ GeV})$	-0.000171 ± 0.000064
$L_5^{(2)}(\Lambda_{\chi PT} = 1\text{ GeV})$	0.000513 ± 0.000078
$L_6^{(2)}(\Lambda_{\chi PT} = 1\text{ GeV})$	-0.000146 ± 0.000036
$L_8^{(2)}(\Lambda_{\chi PT} = 1\text{ GeV})$	0.000631 ± 0.000041

TABLE I. Summary of results from the simulations reported here. The first error is the statistical error, which for most quantities is much larger than any systematic error we can measure or estimate. The exception is for the quantities in $\overline{\text{MS}}$ and $\hat{B}_K$. For these quantities, the second error is the systematic error on the renormalization, which is dominated by the perturbative matching between the continuum RI-MOM scheme and the continuum $\overline{\text{MS}}$ scheme.

The layout of this document is as follows: In Section II we present the details of our new ensembles, including a more general discussion of the Möbius domain wall action. The associated simulated values of the pseudoscalar masses and decay constants, the Ω-baryon mass, the vector and axial current renormalization factors, the neutral kaon mixing parameter, B_K, and the Wilson flow scales, $t_0^{1/2}$ and w_0, are given in Section III. In Section IV we provide an overview of our global fitting procedure for those quantities, the results of which are given in Section V. Finally, we present our conclusions in Section VI.

II. SIMULATION DETAILS AND ENSEMBLE PROPERTIES

Substantial difficulties must be overcome in order to work with physical values of the light quark mass. Common to all fermion formulations are the challenges of increasing the physical spacetime volume to avoid the large finite-volume errors that would result from decreasing the pion mass at fixed volume. Similarly, the range of eigenvalues of the Dirac operator increases substantially, requiring many more iterations for the computation of its inverse and motivating the use of deflation and all-mode-averaging to reduce this computational cost. For domain wall fermions it is also necessary to decrease the size of the residual chiral symmetry breaking to reduce the size of the residual mass to a level below that of the physical light quark masses. While this could have been accomplished using the Shamir domain wall formulation [8, 9] used in previous RBC and UKQCD work, this would have required a doubling or tripling of the length of the fifth dimension, L_s, at substantial computational cost.

Instead, our new, physical ensembles have been generated with a modified domain wall fermion action that suppresses residual chiral symmetry breaking, resulting in values for the residual mass that lie below that of the physical light quark, but without the substantial increase in L_s that would have been required in the original domain wall framework.

We use the Möbius framework of Brower, Neff and Orginos [10–12]. Although the action has been changed, we remain within the subspace of the Möbius parametrization that preserves the $L_s \to \infty$ limit of domain wall fermions. The changes to the Symanzik effective action resulting from this change in fermion formulation can be made arbitrarily small and are of the same size as the observed level of residual chiral symmetry breaking. As discussed in Section II A, we are therefore able to combine our new ensembles in a continuum extrapolation with previous RBC and UKQCD ensembles.

A. Möbius fermion formalism

In this section and in Appendix A we describe the implementation of Möbius domain wall fermions, and provide a self-contained derivation of many of the properties of this formulation on which our calculation depends.

Of central importance is the degree to which the present results from the Möbius version of the domain wall formalism can be combined with those from our earlier Shamir calculations when taking a continuum limit. As reviewed below and in Appendix A, the Shamir and Möbius fermion formalisms result in very similar approximate sign functions, $\varepsilon(H_M)$, having the form given in Eq. (24) below. In fact, the only differences between the two functions $\varepsilon(H_M)$ corresponding to Shamir and Möbius fermions is the choice of L_s and an overall scale factor entering the definition of the kernel operator, H_M. Thus, in the limit $L_s \to \infty$ both theories agree with the same, chirally symmetric, overlap theory. The differences of both Shamir and Möbius fermions from that theory, and therefore from each other, vanish in this chiral limit. Note, this equivalence in the chiral limit holds for both the fermion determinant that is used to generate the gauge ensembles (shown below) and for the 4-D propagators (shown in Appendix A) which determine all of the Green's functions which appear in our measurements and define our lattice approximation to QCD.

Thus, we expect that all details of the four dimensional approximation to QCD defined by the Shamir and Möbius actions must agree in the limit $L_s \to \infty$ and, in our case of finite L_s, will show differences on the order of the residual chiral symmetry breaking, the most accessible effect of finite L_s. Since this constraint holds at finite lattice spacing, we conclude that the coefficients of the $O(a^2)$ corrections which appear in the four-dimensional, effective Symanzik Lagrangians for the Shamir and Möbius actions should agree at this same, sub-percent level, allowing a consistent continuum limit to be obtained from a combination of Shamir and Möbius results.

To understand this argument in greater detail, it is useful to connect the Shamir and Möbius theories in two steps. We might first discuss the relation between two Shamir theories: one with a smaller L_s and larger residual chiral symmetry breaking, and a second with a larger value of L_s and a value for $m_{\rm res}$ below the physical light quark mass. In the second step we can compare this large L_s Shamir theory with a corresponding Möbius theory that has the same approximate degree of residual chiral symmetry breaking. For example, when comparing our $\beta = 2.13$ Shamir and Möbius ensembles, we might begin with our $24^3 \times 64$, $L_s = 16$, 24I ensemble with $m_{\rm res}a = 0.003154(15)$ which is larger than the physical light quark mass. Next we consider a fictitious, $L_s = 48$ ensemble which should have a value of $m_{\rm res}$ very close to the 0.0006102(40) value of our 48I Möbius ensemble. In this comparison we would work with the same Shamir formalism and simply approach the chiral limit more closely by increasing L_s from 16 to 48. Clearly the $5\times$ reduction in the light quark mass will produce a significant change in the theory, which to a large degree should be equivalent to reducing the input quark mass in a theory with a large fixed value of L_s. Of course, there will be smaller changes as well. In addition to reducing the size of $m_{\rm res}$, we will also reduce the size of the dimension-five, $O(a)$ Sheikholeslami-Wohlert term (whose effects are expected to be at the $m_{\rm res}a^2 \le 0.1\%$ level even for the smaller value of L_s). There will be further small changes coming from approaching the $L_s \to \infty$ limit, for example the 3% change in the lattice spacing discussed in Appendix C.

The second comparison can be made between the fictitious $L_s = 48$ Shamir ensemble and our actual 48I Möbius ensemble with $L_s = 24$ and $b+c = 2$. Since the product of $L_s(b+c)$ is the same for these two examples, the approximate sign function will agree for eigenvalues of the kernel H_M which are close to zero. In fact, a study of the eigenvalues λ of H_M for the Shamir normalization shows that they lie in the range $0 \le \lambda \le 1.367(14)$ for $\beta = 2.13$. One can then examine the ratio of the two approximate sign functions, which determine the corresponding 4-D Dirac operators, over this entire eigenvalue range and show that the approximate Shamir and Möbius sign functions $\varepsilon(H_M)$ agree at the 0.1% level. Thus, in this second step we are comparing two extremely similar theories whose description of QCD is expected to differ in all aspects at the 0.1% level. We now turn to a detailed discussion of the Shamir and Möbius operators and their relation to the overlap theory.

Our conventions are as follows. The usual Wilson matrix is

$$D_W(M) = M + 4 - \frac{1}{2}D_{\rm hop}, \tag{1}$$

where

$$D_{\rm hop} = (1-\gamma_\mu)U_\mu(x)\delta_{x+\mu,y} + (1+\gamma_\mu)U_\mu^\dagger(y)\delta_{x-\mu,y}\,. \tag{2}$$

For our physical point ensembles we use a generalized form of the domain wall action [10–12],

$$S^5 = \bar{\psi} D^5_{GDW}\psi\,, \tag{3}$$

where

$$D^5_{GDW} = \begin{pmatrix} \tilde{D} & -P_- & 0 & \dots & 0 & mP_+ \\ -P_+ & \ddots & \ddots & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \ddots & \ddots & -P_- \\ mP_- & 0 & \dots & 0 & -P_+ & \tilde{D} \end{pmatrix}, \tag{4}$$

and we define

$$D_+ = (bD_W + 1) \quad ; \quad D_- = (1 - cD_W) \quad ; \quad \tilde{D} = (D_-)^{-1} D_+ \,. \tag{5}$$

This generalized set of actions reduces to the standard Shamir action in the limit $b = 1$, $c = 0$, and it can also be taken to give the polar approximation to the Neuberger overlap action as another limiting case [13, 14]. In all of our simulations we take the coefficients b and c as constant across the fifth dimension. This setup is well known to yield a tanh approximation to the overlap sign function. Coefficients that vary across the fifth dimension can also be used to introduce other rational approximations to the sign function, such as the Zolotarev approximation [15–17].

As in the Shamir domain wall fermion formulation we identify "physical", four-dimensional quark fields q and $\bar{q}$ whose Green's functions define our domain wall fermion approximation to continuum QCD. We choose to construct these as simple chiral projections of the five-dimensional fields ψ and $\bar{\psi}$ which appear in the action given in Eq. (3):

$$\begin{aligned} q_R &= P_+ \psi_{L_s} \quad q_L = P_- \psi_1 \,, \\ \bar{q}_R &= \bar{\psi}_{L_s} P_- \quad \bar{q}_L = \bar{\psi}_1 P_+ \,. \end{aligned} \tag{6}$$

While there is considerable freedom in this choice of the physical, four-dimensional quark fields, as is shown in Appendix A, this choice results in four-dimensional propagators which agree with those of the corresponding overlap theory up to a contact term in the $L_s \to \infty$ limit. This choice is also dictated by the requirement that we be able to combine results from the present, physical point calculation with earlier results using Shamir fermions in taking a continuum limit. With this choice both the Möbius and Shamir theories will yield 4-dimensional fermion propagators which differ only at the level of the residual chiral symmetry breaking. The choice of physical quark fields given in Eq. (6) has the added benefits that the corresponding four-dimensional propagators satisfy a simple γ^5 hermiticity relation and a hermitian, partially-conserved axial current can be easily defined.

In practice, one solves for physical quark propagators using the linear system

$$D_- D^5_{GDW} \psi = D_- \eta \,. \tag{7}$$

To find the 4d effective action which corresponds to our choice of physical fields we must first perform some changes to the field basis as follows. We write

$$S^5 = \bar{\psi} D^5_{GDW} \psi = \bar{\chi} D^5_{\chi} \chi \,, \tag{8}$$

where, for now leaving a matrix Q_- undefined, $\chi = \mathscr{P}^{-1}\psi$, $\bar{\chi} = \bar{\psi}\gamma_5 Q_-$, $D^5_\chi = Q_-^{-1}\gamma_5 D^5_{GDW}\mathscr{P}$, and

$$\mathscr{P} = \begin{pmatrix} P_- & P_+ & 0 & \dots & 0 \\ 0 & \ddots & \ddots & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \ddots & P_+ \\ P_+ & 0 & \dots & 0 & P_- \end{pmatrix} . \tag{9}$$

Then with

$$\tilde{H} = \gamma_5 (D_-)^{-1} D_+ = \gamma_5 (H_-)^{-1} H_+ \,, \tag{10}$$

and $H_- = \gamma_5 D_-$, $H_+ = \gamma_5 D_+$ we may write

$$D^5_\chi = Q_-^{-1} \begin{bmatrix} \tilde{H} & P_- & 0 & \dots & 0 & mP_+ \\ -P_+ & \ddots & \ddots & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \ddots & \ddots & P_- \\ -mP_- & 0 & \dots & 0 & -P_+ & \tilde{H} \end{bmatrix} \mathscr{P} \,. \tag{11}$$

We may choose Q_- to place the matrix D^5_χ in a particularly convenient form as follows,

$$\begin{aligned} Q_- &= \tilde{H}P_- - P_+ = \gamma_5 [H_-]^{-1}[H_+ P_- - H_- P_+] \\ Q_+ &= \tilde{H}P_+ + P_- = \gamma_5 [H_-]^{-1}[H_+ P_+ - H_- P_-] \,, \end{aligned} \tag{12}$$

and introduce the so-called transfer matrix as

$$\begin{aligned} T^{-1} &= -(Q_-)^{-1} Q_+ \\ &= -[\gamma_5 \tfrac{(b+c)D_W}{2+(b-c)D_W} - 1]^{-1}[\gamma_5 \tfrac{(b+c)D_W}{2+(b-c)D_W} + 1] \\ &= -[H_M - 1]^{-1}[H_M + 1]. \end{aligned}$$

Here the Möbius kernel is

$$H^M = \gamma_5 \frac{(b+c)D_W}{2+(b-c)D_W}\,. \tag{13}$$

We find D_χ^5 takes the following form,

$$D_\chi^5 = \begin{bmatrix} P_- - mP_+ & -T^{-1} & 0 & \dots & \dots & 0 \\ 0 & 1 & -T^{-1} & 0 & \dots & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 & \vdots \\ \vdots & \dots & 0 & 1 & -T^{-1} & 0 \\ 0 & \dots & \dots & 0 & 1 & -T^{-1} \\ -T^{-1}(P_+ - mP_-) & 0 & \dots & \dots & 0 & 1 \end{bmatrix}, \tag{14}$$

for which we can perform a UDL decomposition around the top left block:

$$\begin{pmatrix} D & C \\ B & A \end{pmatrix} = \begin{pmatrix} 1 & CA^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} S_\chi & 0 \\ 0 & A \end{pmatrix} \begin{pmatrix} 1 & 0 \\ A^{-1}B & 1 \end{pmatrix}. \tag{15}$$

Here, the Schur complement is $S_\chi = D - CA^{-1}B$, where

$$A = \begin{pmatrix} 1 & -T^{-1} & 0 & \dots & \vdots \\ 0 & 1 & -T^{-1} & 0 & \vdots \\ 0 & 0 & 1 & -T^{-1} & 0 \\ 0 & \dots & 0 & 1 & -T^{-1} \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix} \qquad A^{-1} = \begin{pmatrix} 1 & T^{-1} & T^{-2} & \dots & T^{-(L_s-2)} \\ 0 & 1 & T^{-1} & \dots & T^{-(L_s-3)} \\ 0 & 0 & 1 & T^{-1} & \vdots \\ 0 & \dots & 0 & 1 & T^{-1} \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}, \tag{16}$$

$$D = P_- - mP_+, \tag{17}$$

$$C = (\, -T^{-1} \;\; 0 \; \dots \; \dots \; 0 \,), \tag{18}$$

$$B^T = (\, 0 \;\; \dots \; 0 \;\; -T^{-1}(P_+ - mP_-) \,), \tag{19}$$

$$CA^{-1}B = T^{-L_s}(P_+ - mP_-)\,. \tag{20}$$

Denoting the left and right factors as U and $L(m)$ respectively, we write this factorization as $D_\chi^5 = UD_S(m)L(m)$. The determinants of the U and $L(m)$ are unity, and the determinant of the product is simply

$$\det D_\chi^5 = \det A \det S_\chi = \det S_\chi\,, \tag{21}$$

where

$$S_\chi(m) = -(1+T^{-L_s})\gamma_5 \left[\frac{1+m}{2} + \frac{1-m}{2}\gamma_5 \frac{T^{-L_s}-1}{T^{-L_s}+1} \right]. \tag{22}$$

We can see that after the removal of the determinant of the Pauli Villars fields with $m = 1$ in our ensembles we are left with the determinant of an effective overlap operator, which is the following rational function of the kernel:

$$\det D_{PV}^{-1} D(m) = \det D_{ov} = \det\left(\frac{1+m}{2} + \frac{1-m}{2}\gamma_5 \frac{(1+H_M)^{L_s} - (1-H_M)^{L_s}}{(1+H_M)^{L_s} + (1-H_M)^{L_s}}\right). \tag{23}$$

We identify D_{ov} as an approximation to the overlap operator with approximate sign function

$$\varepsilon(H_M) = \frac{(1+H_M)^{L_s} - (1-H_M)^{L_s}}{(1+H_M)^{L_s} + (1-H_M)^{L_s}}, \tag{24}$$

with

$$\lim_{L_s\to\infty} \varepsilon(H_M) = \mathrm{sgn}(H_M). \tag{25}$$

Note that since $\mathrm{sgn}(H_M) = \mathrm{sgn}(\alpha H_M)$ for all positive α, changing the Möbius parameters $b+c$ while keeping $b-c = 1$ fixed leaves our kernel H_M proportional to the kernel for the Shamir formulation. This therefore changes only the approximation to the overlap sign function, but not the form of the $L_s \to \infty$ limit of the action.

In this way, our new simulations with the Möbius action will differ from those with Shamir domain wall fermions only through terms proportional to the residual chiral symmetry breaking. In particular the change of action is not fundamentally different from simulating with a different L_s.

Other, equivalent views of this approximation to the sign function are useful. Noting

$$-\tanh\frac{1}{2}\log z = \frac{1-z}{1+z}, \tag{26}$$

we see that since

$$T^{-1} = \frac{1+H_M}{1-H_M} \iff H_M = \frac{1-T}{1+T}, \tag{27}$$

we have

$$\frac{T^{-L_s}-1}{T^{-L_s}+1} = \tanh\left(-\frac{L_s}{2}\log|T|\right) = \tanh\left(L_s \tanh^{-1} H_M\right), \tag{28}$$

and for this reason our approximation to the sign function is often called the tanh approximation. For eigenvalues of H_M near zero, this tanh expression becomes a poor approximation to the sign function and it is for these small eigenvalues that the largest contributions to residual chiral symmetry breaking typically occur. For small eigenvalues λ of H_M, the tanh approximation is a steep, but not discontinuous, function at $\lambda = 0$. Examining Eq. (24) one can easily see that

$$\varepsilon(\alpha\lambda) \sim L_s \alpha\lambda, \tag{29}$$

which approaches the discontinuity of the sign function only as $L_s \to \infty$. The quality of the sign function approximation for small eigenvalues can be improved by either increasing L_s (at a linear cost) or by increasing the Möbius scale factor $\alpha = b + c$ while keeping $b - c = 1$ (close to cost-free), or both. One concludes that the scale factor $b+c$ should be increased to the maximum extent consistent with keeping the upper edge of the spectrum of H_M within the bounded region in which $\varepsilon(H_M)$ is a good approximation to the sign function. In the limit of large L_s a simulation with $(b+c) > 1$ will have the same degree of chiral symmetry breaking as a simulation in which that scale factor has been set to one but with L_s increased to $L_s(b+c)$.

In Appendix A we continue the above review of the relation between the DWF and overlap operators, demonstrating the equality of the Shamir and Möbius four-dimensional fermion propagators in the limit $L_s \to \infty$. We also introduce a practical construction of the conserved vector and axial currents for Möbius fermions, appropriate for our choice of physical fermion fields.

B. Simulation parameters and ensemble generation

We generated three domain wall ensembles with the Iwasaki gauge action. The 48I and 64I ensembles were generated with Möbius domain wall fermions and with (near-)physical pion masses, and the 32Ifine ensemble was generated with Shamir DWF and with a heavier mass but finer lattice spacing. The results from previous fits to our older ensembles were used to choose the input light and strange quark masses to the simulations. The input parameters are listed in Table II. As discussed above, the Möbius parameters for the 48I and 64I ensembles are chosen with $b - c = 1$ such that the Shamir and Möbius kernels are identical. The values of $\alpha = b + c$, which to a first approximation gives the ratio of fifth-dimensional extents between the Möbius and the equivalent Shamir actions, are listed in the table.

We use an exact hybrid Monte Carlo algorithm for our ensemble generation, with five intermediate Hasenbusch masses, (0.005, 0.017, 0.07, 0.18, 0.45), for the two flavor part of the algorithm of both the 48I and 64I ensembles, and three intermediate masses, (0.005, 0.2, 0.6), for the 32Ifine. A rational approximation was used for the strange quark determinant. The integrator layout and parameters are given in Tables III and IV.

Each trajectory of the 48I ensemble required 3.5 hours on 2 racks of Blue Gene/Q (BG/Q) (2×1024 nodes), and those of the 64I required 0.67 hours on 8 racks of BG/Q. We generated 2200 and 2850 trajectories for the 48I and 64I ensembles respectively. The first 1100 trajectories of the

	48I	64I	32Ifine
Size	$48^3 \times 96 \times 24$	$64^3 \times 128 \times 12$	$32^3 \times 64 \times 12$
β	2.13	2.25	2.37
am_l	0.00078	0.000678	0.0047
am_h	0.0362	0.02661	0.0186
α	2.0	2.0	1.0
a^{-1}(GeV)	1.730(4)	2.359(7)	3.148(17)
L (fm)	5.476(12)	5.354(16)	2.006(11)
$m_\pi L$	3.863(6)	3.778(8)	3.773(42)
$\langle P \rangle$	0.5871119(25)	0.6153342(21)	0.6388238(37)
$\langle \bar{\psi}\psi \rangle$	0.0006385(12)	0.0002928(9)	0.0006707(15)
$\langle \bar{\psi}\gamma^5\psi \rangle$	-0.0000043(31)	-0.0000000(34)	-0.0000013(26)

TABLE II. Input parameters and relevant quantities for the three new Iwasaki ensembles. Here L is the spatial lattice extent in lattice units, and $\alpha = b + c$ is Möbius scaling factor (recall the 32Ifine is a Shamir DWF ensemble, and therefore has $\alpha = 1.0$). The last three entries are the average plaquette, chiral condensate, and pseudoscalar condensate respectively. The lattice spacings are determined in Section V of this document.

	48I	64I	32Ifine
Steps per traj.	15	9	6
$\Delta\tau$	0.067	0.111	0.167
Metropolis acc.	84%	87%	82%
CG iters per traj.	$\sim 5.9 \times 10^5$	$\sim 6.1 \times 10^5$	$\sim 8.4 \times 10^4$

TABLE III. The number of steps per HMC trajectory, the MD time-step $\Delta\tau$, the Metropolis acceptance and the total number of CG iterations for the three new ensembles.

64I ensemble were generated with $L_s = 10$ and produced a pion mass of about 170 MeV, due to the residual mass being larger than anticipated. Changing to $L_s = 12$ reduced the residual mass, allowing us to simulate at essentially the physical pion mass. The 32Ifine ensemble required 5 minutes on 1 rack of BG/Q, and we generated 6940 trajectories for this ensemble.

Level (i)	S_i	Integrator	n_i	Step size (48I,64I,32Ifine)
1	$\sum S_Q + \sum S_R$	FGI QPQPQ	1	1/15, 1/9, 1/6
2	S_G	FGI QPQPQ	4	-

TABLE IV. The integrator layout for our three ensembles. Here $\sum S_Q$ and $\sum S_R$ are the sum of the quotient and rational quotient actions used for the light and strange quarks respectively. The sums are over the intermediate mass listed in the text. S_G is the gauge action, FGI QPQPQ is a particular form of the force gradient integrator [18], and n_i are the number of steps comprising a single update of the corresponding action. The coarsest time-steps are at level 1, and the step sizes are chosen such that the total trajectory length is 1 MD time unit. More detail regarding the notation and integrators can be found in Appendix A of Ref. [5].

C. Ensemble properties

In Figure 1 we plot the Monte Carlo evolution of the topological charge, plaquette, and the light quark scalar and pseudoscalar condensates, after thermalization. In addition we plot the time histories of the Clover-form energy density evaluated at the Wilson flow times w_0^2 and t_0 in Figure 2. We measured the topological charge by cooling the gauge fields with 60 rounds of APE [19] smearing (smearing coefficient 0.45), and then measured the field-theoretic topological charge density using the 5Li discretization of Ref. [20], which eliminates the $\mathcal{O}(a^2)$ and $\mathcal{O}(a^4)$ terms at tree level. In Figure 3 we plot histograms of the topological charge distributions.

In Figure 4 we plot the integrated autocorrelation time for the same observables on the 32Ifine, 48I, and 64I ensembles as a function of the cutoff in Molecular Dynamics (MD) time separation, Δ_{cutoff}:

$$\tau_{\text{int}}(\Delta_{\text{cutoff}}) = 1 + 2 \sum_{\Delta=1}^{\Delta_{\text{cutoff}}} C(\Delta) \,, \tag{30}$$

where

$$C(\Delta) = \left\langle \frac{\left(Y(t) - \overline{Y}\right)\left(Y(t+\Delta) - \overline{Y}\right)}{\sigma^2} \right\rangle_t \tag{31}$$

is the autocorrelation function associated with the observable $Y(t)$. The mean and variance of $Y(t)$ are denoted $\overline{Y}$ and σ^2, and Δ is the lag measured in MD time units. The error on the integrated autocorrelation time is estimated using a method discussed in our earlier paper [5]: for each fixed Δ in Eq. (31) we bin the set of measurements $\left(Y(t) - \overline{Y}\right)\left(Y(t+\Delta) - \overline{Y}\right)$ over neighboring configurations and estimate the error on the mean $\langle \cdots \rangle_t$ by bootstrap resampling. We then increase the

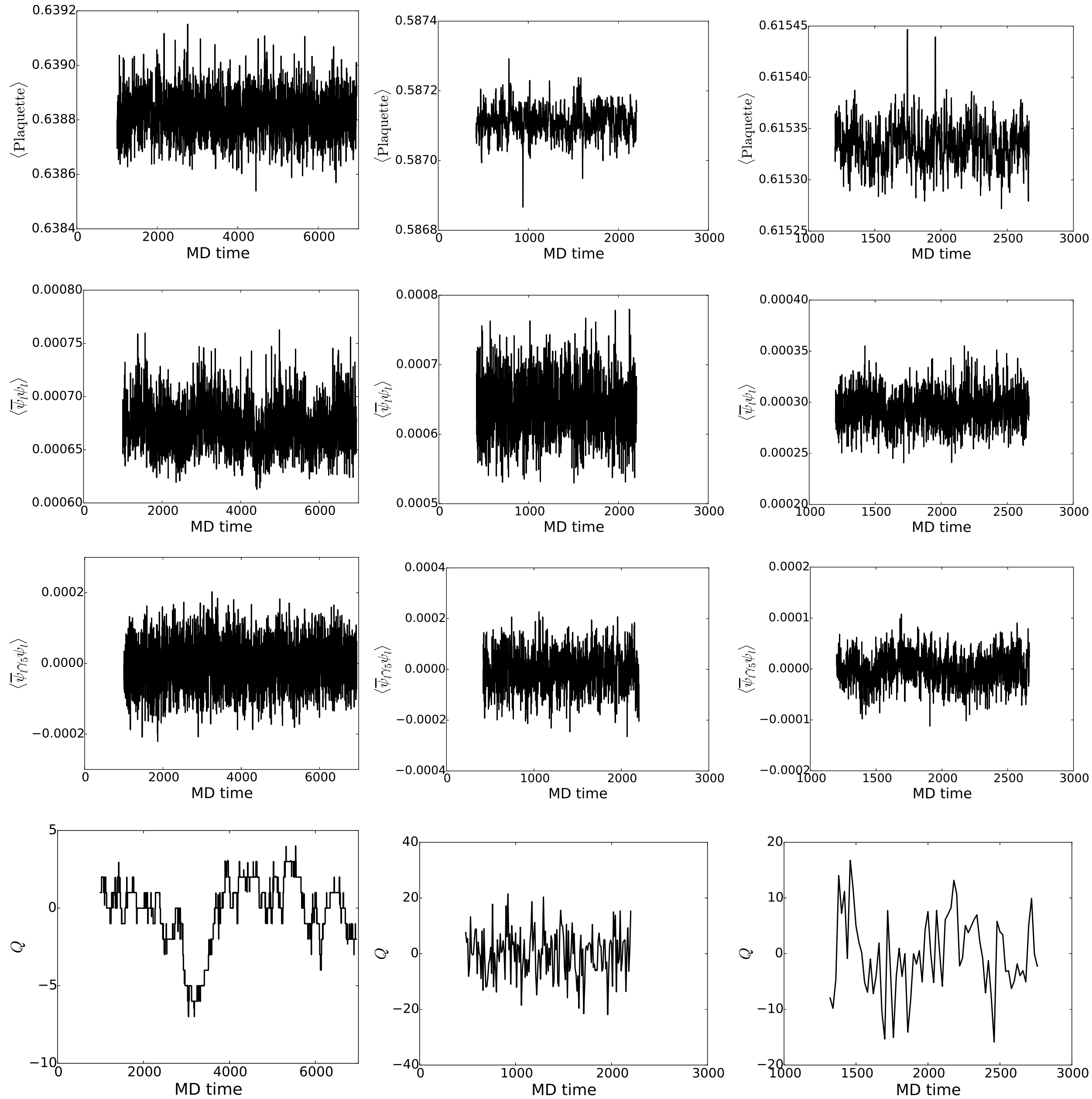


FIG. 1. Monte Carlo evolution of the average plaquette (first row), light quark chiral condensate (second row), light quark pseudoscalar condensate (third row), and topological charge (fourth row) after thermalization on the 32Ifine (left column), 48I (middle column) and 64I (right column) ensembles.

bin size until the error bars stop growing, which we found to correspond to bin sizes of 960, 100, and 200 MD time units on the 32Ifine, 48I, and 64I ensemble, respectively. The error on τ_{int} is then computed from the bootstrap sum in Eq. (30).

In Table V we tabulate estimates of the autocorrelation lengths for each of the various quantities

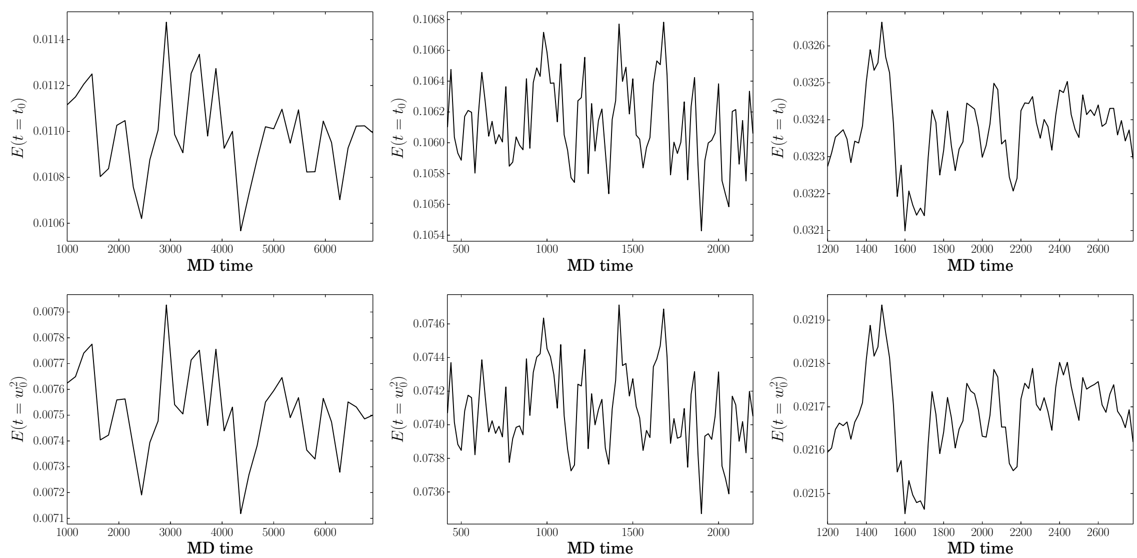


FIG. 2. Time history plots for the energy density evaluated at the Wilson flow times t_0 (top line) and w_0^2 (bottom line) on the 32Ifine (left column), 48I (middle column), 64I (right column).

Ensemble	$\langle P\rangle$	E_{t_0}	$E_{w_0^2}$	Q	Q^2	$\langle\bar{\psi}\gamma^5\psi\rangle$	$\langle\bar{\psi}\psi\rangle$
32Ifine	2.9(7)	29(77)	51(66)	340(120)	240(140)	2.6(8)	24(4)
48I	4.1(1.0)	10(26)	10(24)	1.1(1.6)	0.2(5)	1.9(3)	1.4(3)
64I	4.7(1.7)	38(24)	30(22)	19(7)	5(9)	6(8)	2.0(4)

TABLE V. Estimated integrated autocorrelation times for various quantities on the 32Ifine, 48I and 64I ensembles.

included in the above figures. We can estimate τ_{int} from the upper bound on the error for the slowest mode, which corresponds to the energy densities on the 64I and 48I ensembles, and the topological charge on the 32Ifine. This suggests $\tau_{\text{int}} \sim 35$ MDTU for the 48I ensemble, $\tau_{\text{int}} \sim 50$ MDTU for the 64I ensemble and $\tau_{\text{int}} \sim 460$ MDTU for the 32Ifine ensemble.

For all quantities considered, we observe that the chosen bin sizes are sufficient to account for the autocorrelations suggested by Figure 4. We also observe a significant decrease in the rate of tunneling between configurations with different topological charge as the lattice spacing becomes finer, as evidenced by the long autocorrelation time on the 32Ifine ensemble.

After generating our ensembles we discovered that there are spurious correlations between U(1) random numbers generated by the Columbia Physics System (CPS) random number generator (RNG) with a new seed. Fortunately, as discussed in Appendix G, we determined that the corre-

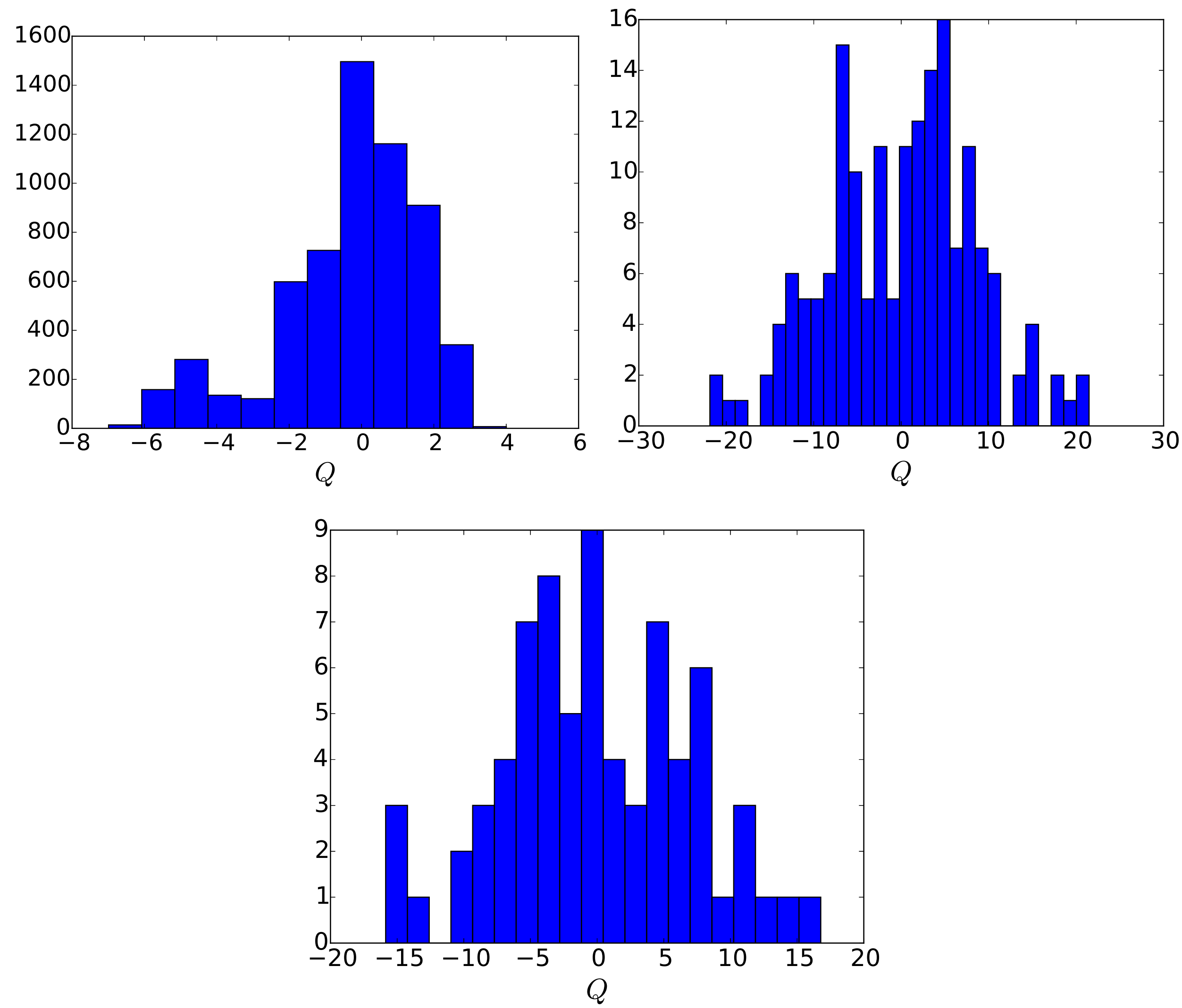


FIG. 3. Topological charge distributions for the 32Ifine (top left), 48I (top right), and 64I (bottom) ensembles.

lation present in the freshly-seeded RNG state was lost during thermalization, and consequently that this had no measurable effect on our thermalized gauge configurations or measurements.

III. SIMULATION MEASUREMENT RESULTS

In this section we present the results of fitting to a number of observables on the 48I and 64I ensembles. On the 48I ensemble we used data from 80 configurations in the range 420–2000 with a separation of 20 MD time units. The 64I measurements were performed on 40 configurations in the range 1200–2760 and separated by 40 MD time units. The data on both ensembles were

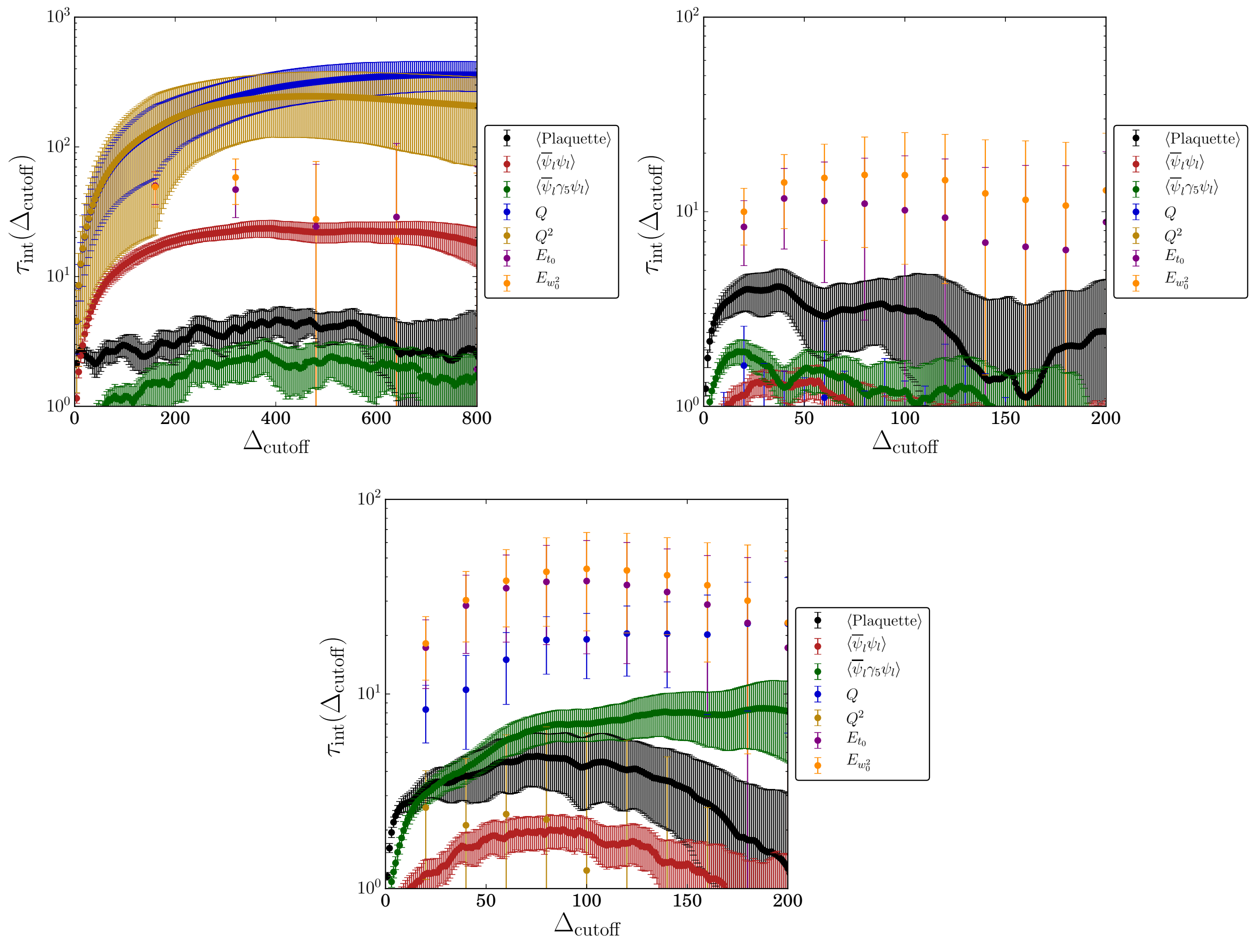


FIG. 4. The integrated autocorrelation time as a function of the cutoff MD time separation, $\Delta_{\rm cutoff}$, for the average plaquette; light quark scalar and pseudoscalar densities; topological charge Q and its square; and the Clover-form energy densities evaluated at Wilson flow times t_0 and w_0^2, E_{t_0} and $E_{w_0^2}$ respectively. These are plotted for the 32Ifine ensemble (top left), 48I (top right), and 64I (bottom) ensembles. The data has been binned over 960, 100, and 200 MD time units on the 32Ifine, 48I, and 64I ensemble, respectively.

binned over 5 successive configurations, corresponding to 100 MD time units and 200 MD time units respectively. On the 64I ensemble, we measured the cheaper Wilson flow scales every 20 configurations (as opposed to every 40 for the other measurements) in the range 1200–2780 and binned over 10 successive configurations. We also present similar results computed on 36 configurations of the 32Ifine ensemble in the range 1000–6600, measuring every 160 MD time units and using a bin size of 6 configurations (960 MD time units).

With the bin sizes given above, the number of binned samples on the 48I, 64I and 32Ifine ensembles are 16, 8 and 6 respectively. We emphasize however that each measurement on the 64I

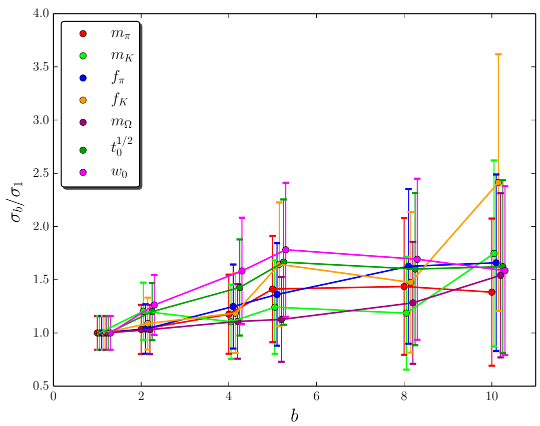


FIG. 5. The dependence of the error for the simulated data on the 64I ensemble. The vertical axis plots the ratio σ_b/σ_1 for bin size b along the horizontal axis, where σ is the statistical error and the subscript indicates the bin size for which that error was computed. The upper and lower bounds were obtained by varying σ_b by $1/\sqrt{N}$, where N is the number of samples.

ensemble is obtained from an average over 128 timeslices, and those on the 48I and 32Ifine over 96 and 64 timeslices, respectively. Nevertheless, the numbers of binned samples on the 64I and 32Ifine ensembles are considerably smaller than those typically encountered in lattice simulations and we therefore provide evidence that our use of this small number of large bins does to not lead to an inaccurate assignment of errors.

First, based on the integrated autocorrelation times determined in the previous section, the expected effective time separation between uncorrelated measurements is ~ 100 MDTU on the 64I ensemble, half of the actual bin size chosen. (Recall this is estimated as $2 \times \tau_{\text{int}}$). Our choice is therefore quite conservative. For the 48I and 32Ifine ensembles the time separation between uncorrelated measurements is ~ 70 and ~ 920 MDTU, respectively, which are comparable to our bin sizes of 100 and 960. However, these estimates are obtained from the energy densities and topological charge respectively, and the latter may be misleadingly large for the following reason. In a study by the ALPHA collaboration [21] the authors point out that for an HMC algorithm which is invariant under parity, such as ours, the correlations seen in parity-even observables, which we study, will correspond to modes in the HMC evolution which are determined by parity-even quantities such as Q^2. We have included this quantity also in Figure 4 and Table V, for which we observe substantially smaller autocorrelation lengths, suggesting that our 48I and 32Ifine bin sizes are also quite conservative.

Of the 32Ifine and 64I data sets, the latter is the most important to our analysis. In Figure 5 we plot the error on the 64I simulated data as a function of increasing bin size, where we estimate the error on the error as $\sim 1/\sqrt{N}$ were N is the number of binned samples. In Figure 19 of Section V we show a similar plot but for the physical predictions of our global fits, again as a function of the 64I bin size. From these figures we observe no statistically significant dependence on the 64I bin size, suggesting that we are not underestimating our errors by making our choice of 100 MDTU bins for this ensemble.

The ability to generate physical mass ensembles forced us to seek dramatic improvements in our measurement strategy, since the statistical error for kaon observables increases with decreasing light quark mass (holding the strange quark mass fixed). For an example of this behavior, consider the kaon two-point function,

$$C(t) = \sum_{\vec{y},\vec{x}} [\bar{s}u](\vec{y},t)[\bar{u}s](\vec{x},0)\,, \tag{32}$$

which in the limit of large t goes as

$$\langle C(t)\rangle = Ae^{-m_K t} + \dots\,, \tag{33}$$

where $\langle..\rangle$ is the average over the gauge field ensemble. The standard deviation on this quantity, i.e. its statistical error, goes as $\sqrt{\langle C^2(t)\rangle}$, which contains two strange quark propagators and two light quark propagators. This quantity can also be represented as a linear combination of exponentially decaying terms:

$$\langle C^2(t)\rangle = Be^{-(m_{s\bar{s}}+m_\pi)t} + \dots\,, \tag{34}$$

where $m_{s\bar{s}}$ is the mass of the $s\bar{s}$ state. The signal-to-noise ratio goes as $\exp\left(-[m_K - (m_{s\bar{s}} + m_\pi)/2]t\right)$ in the large time limit, and therefore decays faster with lighter pions.

The first component of our measurement strategy involves maximally reusing propagators for all of our measurements, which include m_π, m_K, m_Ω, f_π, f_K, B_K, $f_+^{K\pi}(0)$ and $K \to (\pi\pi)_{I=2}$. (Note that the latter two quantities are not reported on in this document.) Reusing propagators requires choosing a common source for our propagators that remains satisfactory across the entire range of measurements. Also, since we measure both two- and three-point functions, we need to be able to control the spatial momentum of the sources in order to project out unwanted momenta. We performed numerous studies of Coulomb gauge fixed wall sources and Coulomb gauge fixed box sources for many of these observables. (The box sources were generically chosen so that an integer multiple of their linear dimension would fit in the lattice volume, allowing us to obtain zero momentum projections by using all possible box sources.) While the box sources showed faster projection onto the desired ground state, the statistical errors on the wall sources were much smaller, such that the errors on the measured quantities per unit of computer time were essentially the same. From these studies, we chose to use the simple Coulomb gauge fixed wall sources.

In previous work on the $\eta - \eta'$ mass, which involves disconnected quark diagrams, we found that translating n-point functions over all possible temporal source locations reduced the error essentially as the square-root of the number of translations [22]. The calculation of such a large number of quark propagators on a single configuration can be accomplished much more quickly by a deflation algorithm. The EigCG algorithm [23] was used for $K \to (\pi\pi)_{I=0,2}$ measurements at unphysical kinematics in Ref. [24], and was adopted for this calculation. Measurements were again performed for all temporal translations of the n-point functions, and a factor of 7 speed-up was achieved. The major drawback of EigCG is the considerable memory footprint. However, BG/Q partitions have large memory and therefore this issue can be managed. In practice we found that only a fraction of the vectors generated by the EigCG method were good representations of the true eigenmodes, and in future we may be able to reduce the CG time further by pre-calculating exact low-modes using the implicitly restarted Lanczos algorithm with Chebyshev acceleration [25].

An alternative approach to generating a large number of quark propagators is to use inexact deflation. [26]. This approach had not been optimally formulated for the domain wall operator when the measurements on our new ensembles were begun. However, a new formulation of inexact deflation appropriate to DWF, known as HDCG [27], has since been developed, and has been shown to be more efficient than EigCG; this technique is now being used for our valence measurements.

The final component of our measurement package is the use of the all-mode averaging (AMA) [7] method to further reduce the cost of translating the propagator sources along the temporal direction. AMA is a generalization of low-mode averaging, in which one constructs an approximate propagator using exact low eigenmodes and a polynomial approximation to the high modes obtained by applying deflated conjugate gradient (CG) to a source vector on each temporal slice and averaging over the solutions. The stopping condition on the deflated conjugate gradient can generally be relaxed, reducing the iteration count. The remaining bias in the observable is corrected using a small number of exact solves obtained using the low modes and a precise deflated CG solve from a single timeslice for the high-mode contribution. The benefit of this procedure is that the CG solves used for the polynomial approximation can be performed very cheaply using inexact 'sloppy' stopping conditions of 10^{-4} or 10^{-5} as many of the low modes are already projected out exactly. The net result of combining the sloppy translated solution with the (typically small) bias correction is an exact result calculated many times more cheaply than if we were to perform precise deflated solves on every timeslice.

In order to avoid any bias due to the even-odd decomposed Dirac operator used in the CG, we calculate the eigenvectors using EigCG on a volume source spanning the entire four-dimensional volume, and the temporal slices where we perform the exact solves are chosen randomly for each configuration. We calculate low modes in single precision using EigCG in order to reduce the memory footprint, and also perform the sloppy solves in single precision. For the exact solves, we achieve double precision accuracy through multiple restarts of single precision solves, restarting the solve by correcting the defect as calculated in double precision. For the zero-momentum strange quark propagators required, we do a standard, accurate CG solve for sources on every timeslice. On the 48I, we performed our measurements using single rack BG/Q partitions (1024 nodes), calculating 600 low modes with EigCG (filling the memory) and running continuously for 5.5 days. (Note that this timing includes non-zero momentum light quark solves for measurements of f_+^K and additional light quark solves for $K \to (\pi\pi)_{I=2}$, which are not reported in this document.) For the 64I ensemble, the measurements were performed on between 8 and 32 rack BG/Q partitions at the ALCF and 1500 low modes were calculated by EigCG. On a 32 rack partition, the latter took 5.3 hours and the solver sustained 1 PFlops. (The EigCG setup time is efficiently amortized in these calculations by using the EigCG eigenvectors to deflate a large number of solves.)

The Coulomb gauge-fixing matrices for the 64I ensemble were not computed on the BG/Q and were instead determined separately (and more quickly) on a cluster, using the timeslice-by-timeslice Coulomb gauge FASD algorithm [28].

We simultaneously fit the residual mass, pseudoscalar masses and decay constants, axial and vector current renormalization coefficients (Z_A and Z_V, respectively), and kaon bag parameter (B_{lh}). A separate fit was performed for the Ω-baryon mass. The values for these observables obtained on each lattice, as well as the statistical errors computed by jackknife resampling, are summarized in Table VI. The corresponding fit ranges are summarized in Tables VII and VIII. In the following sections we discuss the fit procedures and plot effective masses and amplitudes for each observable.

	32Ifine	48I	64I
m_{ll}	0.11790(131)	0.08049(13)	0.05903(13)
m_{lh}	0.17720(118)	0.28853(14)	0.21531(17)
f_{ll}	0.04846(32)	0.07580(8)	0.05550(10)
f_{lh}	0.05358(22)	0.09040(9)	0.06653(10)
Z_A	0.77779(29)	0.71191(5)	0.74341(5)
Z_V	0.77700(8)	0.71076(25)	0.74293(14)
B_{lh}	0.5437(85)	0.5841(6)	0.5620(6)
m_{hhh}	0.5522(29)	0.9702(10)	0.7181(7)
m'_{hhh}	0.811(49)	1.273(10)	0.937(7)
m_{res}	0.0006296(58)	0.0006102(40)	0.0003116(23)
w_0	2.664(16)	1.50125(94)	2.0495(15)
$t_0^{1/2}$	2.2860(63)	1.29659(28)	1.74496(62)
m_{ll}/m_{hhh}	0.2135(26)	0.08296(17)	0.08220(20)
m_{lh}/m_{hhh}	0.3209(25)	0.29740(32)	0.29983(37)

TABLE VI. Summary of fit results in lattice units.

Correlator	32Ifine	48I	64I
$PP^{WL}(ll)$	10:31	15:48	12:60
$PP^{WW}(ll)$	10:31	10:35	10:61
$AP^{WL}(ll)$	10:31	10:46	10:60
$PP^{WL}(lh)$	10:31	14:40	17:49
$PP^{WW}(lh)$	10:31	14:33	14:45
$AP^{WL}(lh)$	10:31	12:40	20:49
Z_A	11:52	6:89	10:117
Ω	6:20	5:17	5:19
m_{res}	6:57	9:86	10:117

TABLE VII. Summary of fit ranges $t_{\text{min}}/a \leq t/a \leq t_{\text{max}}/a$ used for each two-point correlator and ensemble.

Ensemble	Quantity	$\lvert t_{\text{source}} - t_{\text{sink}} \rvert / a$	t_{skip}/a
32Ifine	Z_V	16:4:32	8
	B_{lh}	52:4:56	10
48I	Z_V	12:4:24	6
	B_{lh}	20:4:40	10
64I	Z_V	15:5:40	6
	B_{lh}	25:5:40	10

TABLE VIII. Summary of fit ranges used for each three-point correlator and ensemble. We simultaneously fit to all source-sink separations in the given range, where the operator insertion is evaluated at times which are at least t_{skip}/a time slices away from the sources and sinks.

A. Residual mass

For domain wall fermions, the leading effect of having a finite fifth dimension is an additive renormalization to the bare quark masses known as the residual mass, m_{res}. We extract the residual mass from the ratio

$$\mathscr{C}_{m_{\text{res}}}(t) = \frac{\langle 0 | \sum_{\vec{x}} j_{5q}^{a}(\vec{x},t) | \pi \rangle}{\langle 0 | \sum_{\vec{x}} j_{5}^{a}(\vec{x},t) | \pi \rangle}, \tag{35}$$

where j_{5q}^{a} is the pseudoscalar density evaluated at the midpoint of the fifth dimension, and j_5^a is the physical pseudoscalar density constructed from the surface fields (cf. Ref. [29], Eqs. (8) and (9)). In Figure 6 we plot the effective residual mass, as well as the fit, on each ensemble.

B. Pseudoscalar Masses

The masses of the pion and kaon at the simulated quark masses, denoted m_{ll} and m_{lh} respectively, were extracted by fitting to two-point functions of the form

$$\mathscr{C}^{s_1 s_2}_{\mathscr{O}_1 \mathscr{O}_2}(t) = \langle 0 | \mathscr{O}_1^{s_1}(t) \mathscr{O}_2^{s_2}(0) | 0 \rangle, . \tag{36}$$

Here the subscripts indicate the interpolating operators and the superscripts denote the operator smearing used for the sink and source, respectively. In the following we have used Coulomb gauge-fixed wall (W) sources, and both local (L) and Coulomb gauge-fixed wall sinks. We extract the pseudoscalar meson masses by fitting three correlators simultaneously: PP^{LW}, PP^{WW}, and AP^{LW}, where P is the pseudoscalar operator and A is the temporal component of the axial current.

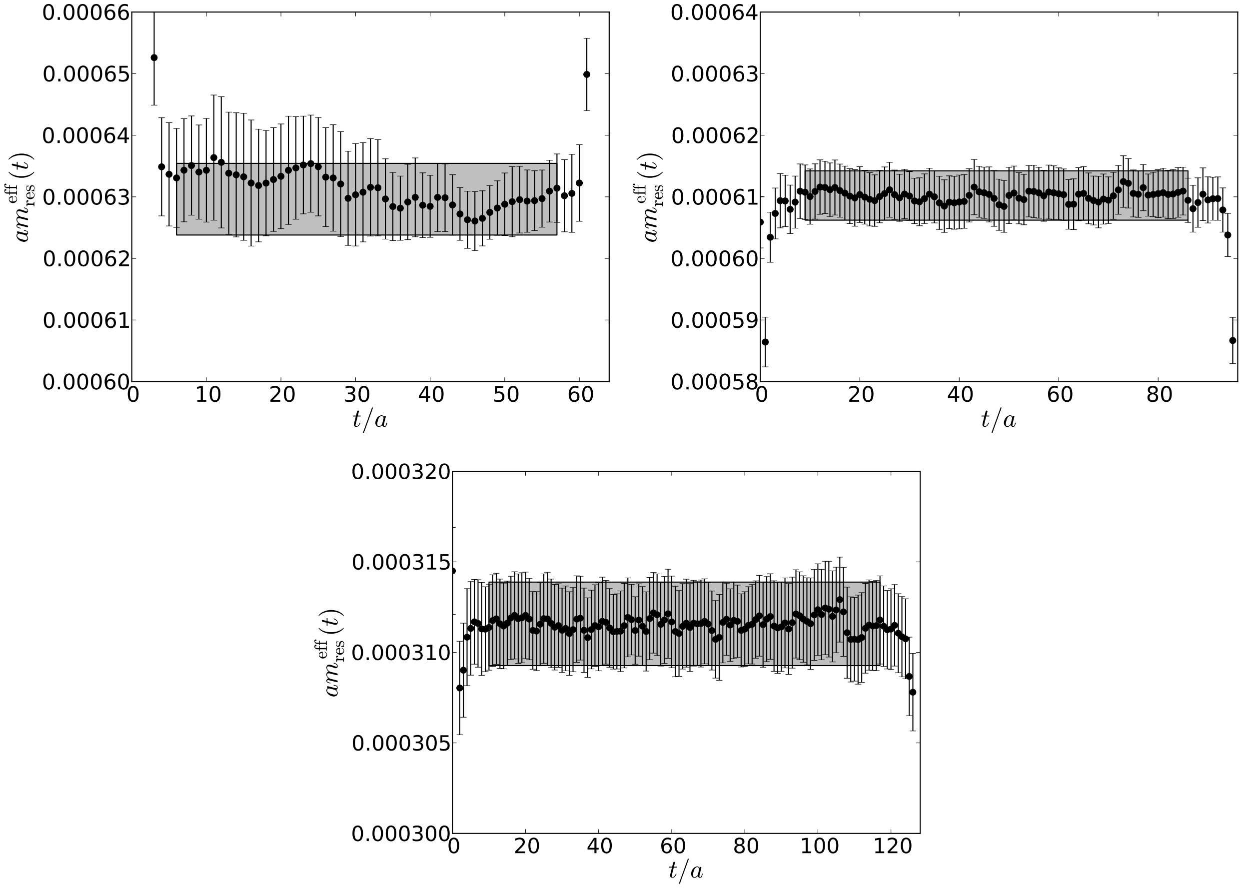


FIG. 6. Effective $m_{\rm res}$ on the 32Ifine (top left), 48I (top right), and 64I (bottom) ensembles.

These are fit to the following analytic form for the ground state of a Euclidean two-point correlation function:

$$\mathscr{C}_{\mathscr{O}_1\mathscr{O}_2}^{s_1 s_2}(t) = \frac{\langle 0|\mathscr{O}_1^{s_1}|X\rangle\langle X|\mathscr{O}_2^{s_2}|0\rangle}{2m_X V}\left(e^{-m_X t} \pm e^{-m_X(N_t - t)}\right), \tag{37}$$

where the + (-) sign corresponds to the PP (AP) correlators, and X denotes the physical state to which the operators couple. In the following sections we use

$$\mathscr{N}_{\mathscr{O}_1\mathscr{O}_2}^{s_1 s_2} \equiv \frac{\langle 0|\mathscr{O}_1^{s_1}|X\rangle\langle X|\mathscr{O}_2^{s_2}|0\rangle}{2m_X V} \tag{38}$$

to denote the amplitude for a given correlator. The effective mass plots associated with these correlators, as well as the fitted masses, are shown in Figures 7, 8, 9, and 10.

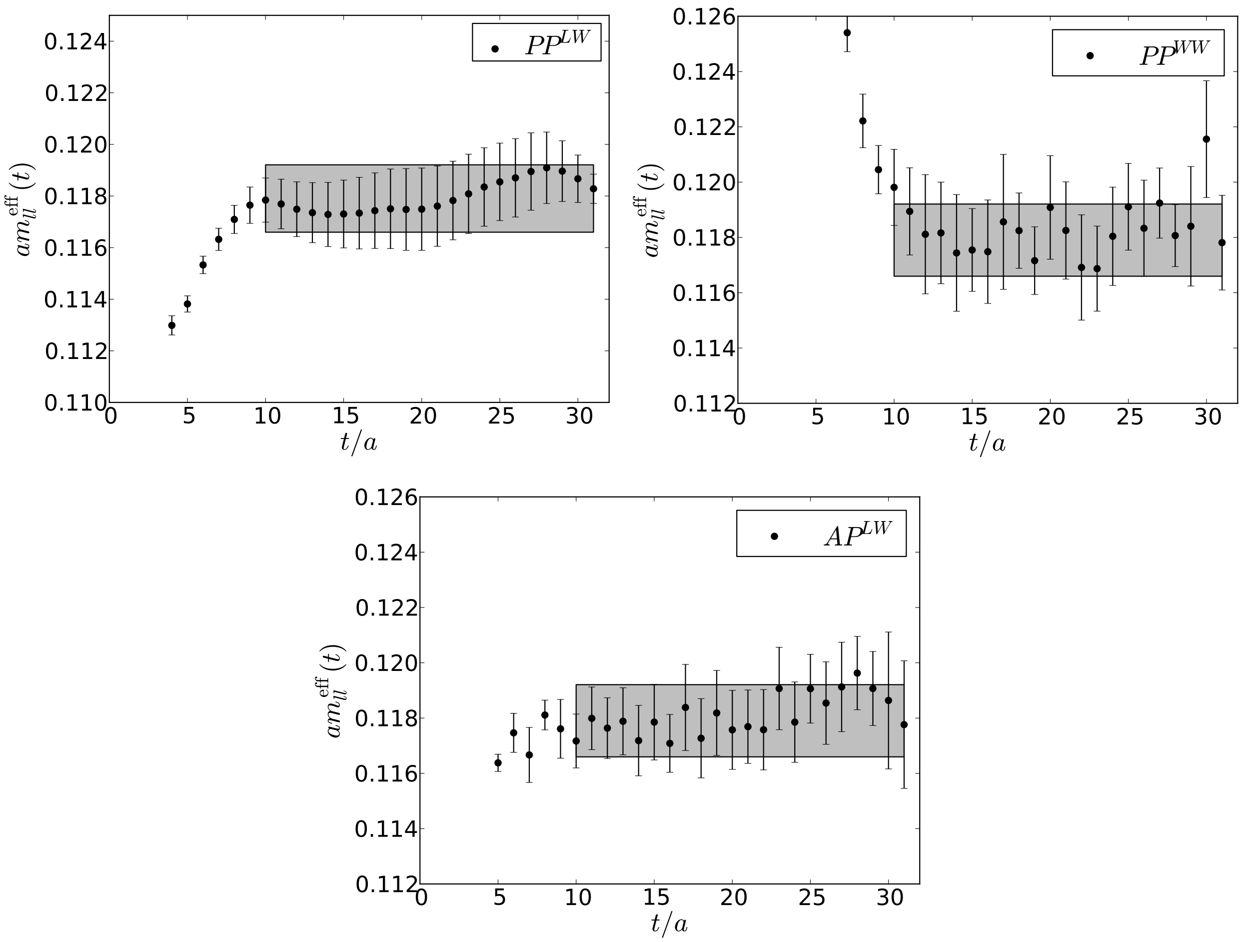


FIG. 7. Effective m_{ll} on the 32Ifine ensemble. We fit a common value of the mass to all three correlators.

C. Pseudoscalar Decay Constants and Axial Current Renormalization

The pseudoscalar decay constants, f_π and f_K, are defined in terms of the coupling of the pseudoscalar meson fields to the local four-dimensional axial current A_μ^a:

$$\begin{cases} \langle 0|A_\mu^a(x)|\pi^b(p)\rangle = -i\delta^{ab} f_\pi p_\mu e^{ip\cdot x} \\ \langle 0|A_\mu^a(x)|K^b(p)\rangle = -i\delta^{ab} f_K p_\mu e^{ip\cdot x} \end{cases}, \tag{39}$$

where

$$A_\mu^a(x) = \overline{q}(x)\gamma_\mu\gamma_5\lambda^a q(x) \tag{40}$$

is formed from the surface fields $q(x)$. In order to match this operator to the physically normalized Symanzik-improved axial operator A_μ^{Sa}, we must derive the appropriate renormalization factor, Z_A. In the domain wall fermion formalism it is also possible to define a five-dimensional current $\mathscr{A}_\mu^a$

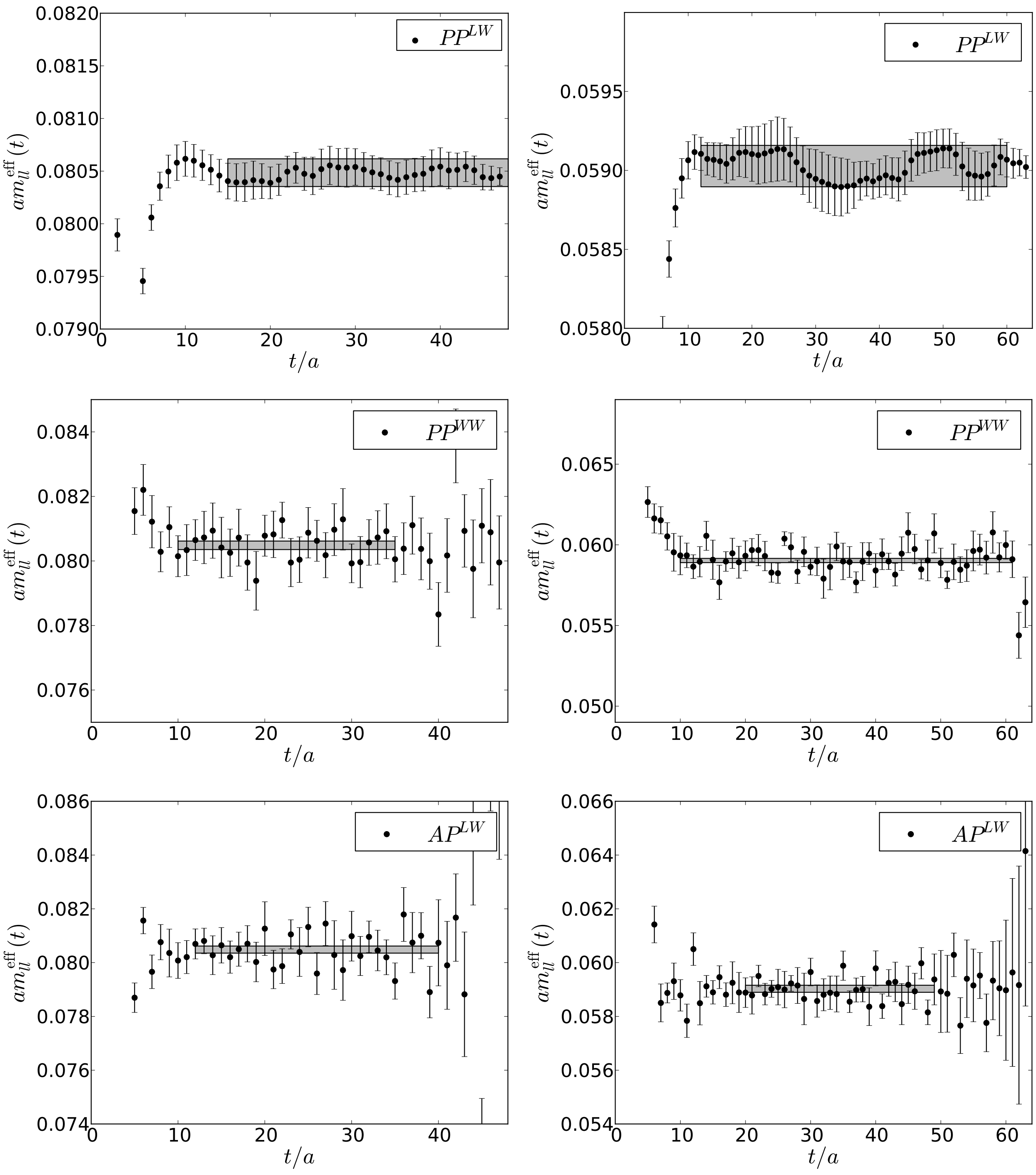


FIG. 8. Effective m_{ll} on the 48I (left column) and 64I (right column) ensembles. We fit a common value of the mass to all three correlators on a given ensemble.

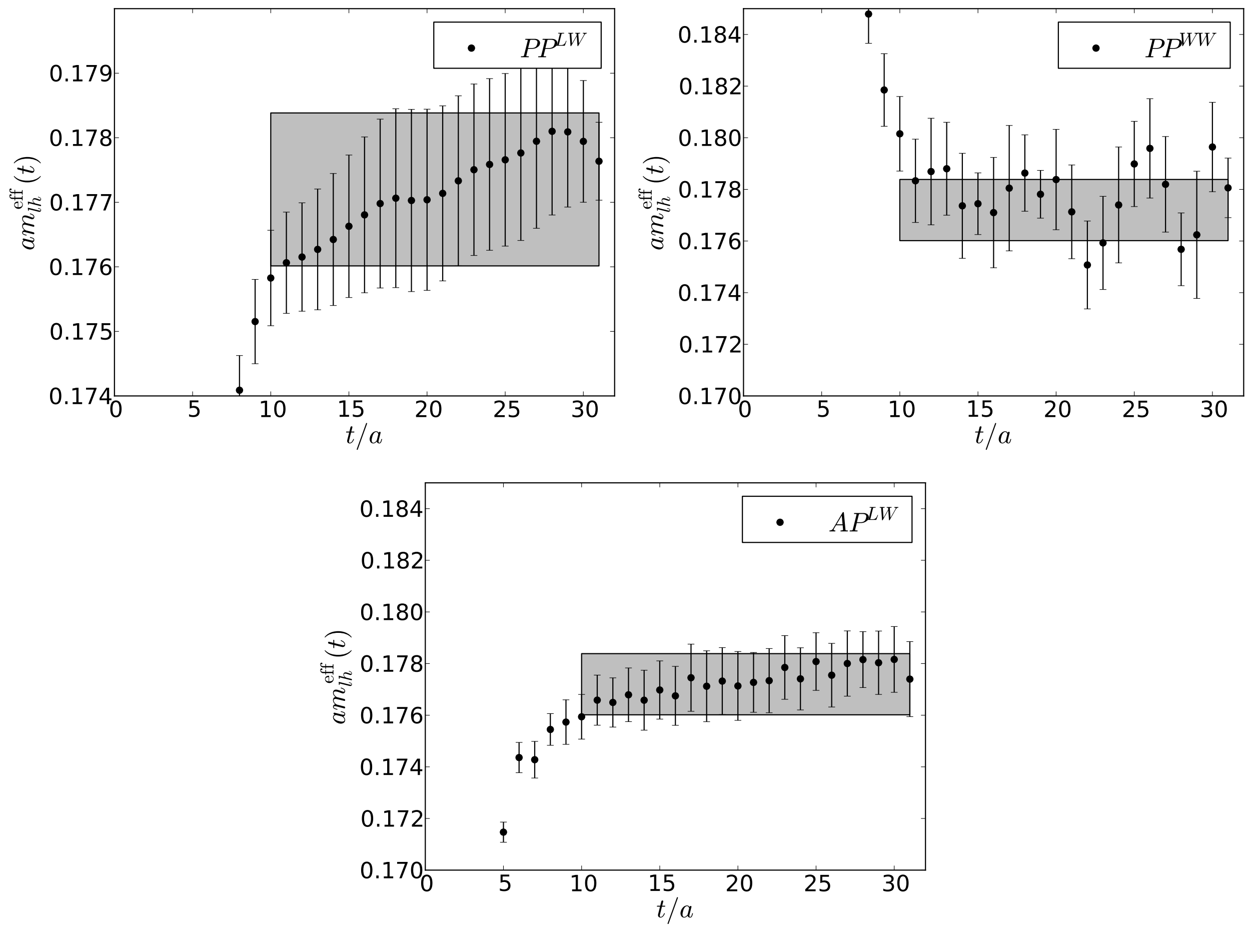


FIG. 9. Effective m_{lh} on the 32Ifine ensemble. We fit a common value of the mass to all three correlators.

which satisfies the discretized partially-conserved axial current (PCAC) relation,

$$\Delta_\mu^- \langle \pi(x) | \mathscr{A}_\mu^a(y) \rangle = \langle \pi(x) | 2m j_5^a(y) + 2 j_{5q}^a(y) \rangle, \tag{41}$$

where Δ_μ^- is the backwards discretized derivative. The factor relating this to the Symanzik current is denoted $Z_{\mathscr{A}}$.

In the past, we took advantage of the fact that $Z_{\mathscr{A}} = 1 + \mathscr{O}(m_{\rm res})$ to approximate Z_A as $Z_A/Z_{\mathscr{A}}$, which can be computed directly via the following ratio:

$$Z_A \approx \frac{Z_A}{Z_{\mathscr{A}}} = \frac{\langle 0 | \sum_{\vec{x}} \mathscr{A}_\mu^a(\vec{x},t) | \pi \rangle}{\langle 0 | \sum_{\vec{x}} A_\mu^a(\vec{x},t) | \pi \rangle}. \tag{42}$$

The 5-D current $\mathscr{A}_\mu^a(x)$ is properly defined as the current carried by the link *between* x and $x+\mu$, whereas the 4-D current $A_\mu^a(x)$ is defined on the lattice site x. The correlation functions $C(t+\frac{1}{2}) = \sum_{\vec{x}} \langle \mathscr{A}_0^a(\vec{x},t) \pi^a(\vec{0},0) \rangle$ and $L(t) = \sum_{\vec{x}} \langle A_0^a(\vec{x},t) \pi^a(\vec{0},0) \rangle$, that one would use to compute the above ratio, are therefore not defined at the same temporal coordinate. By taking appropriate

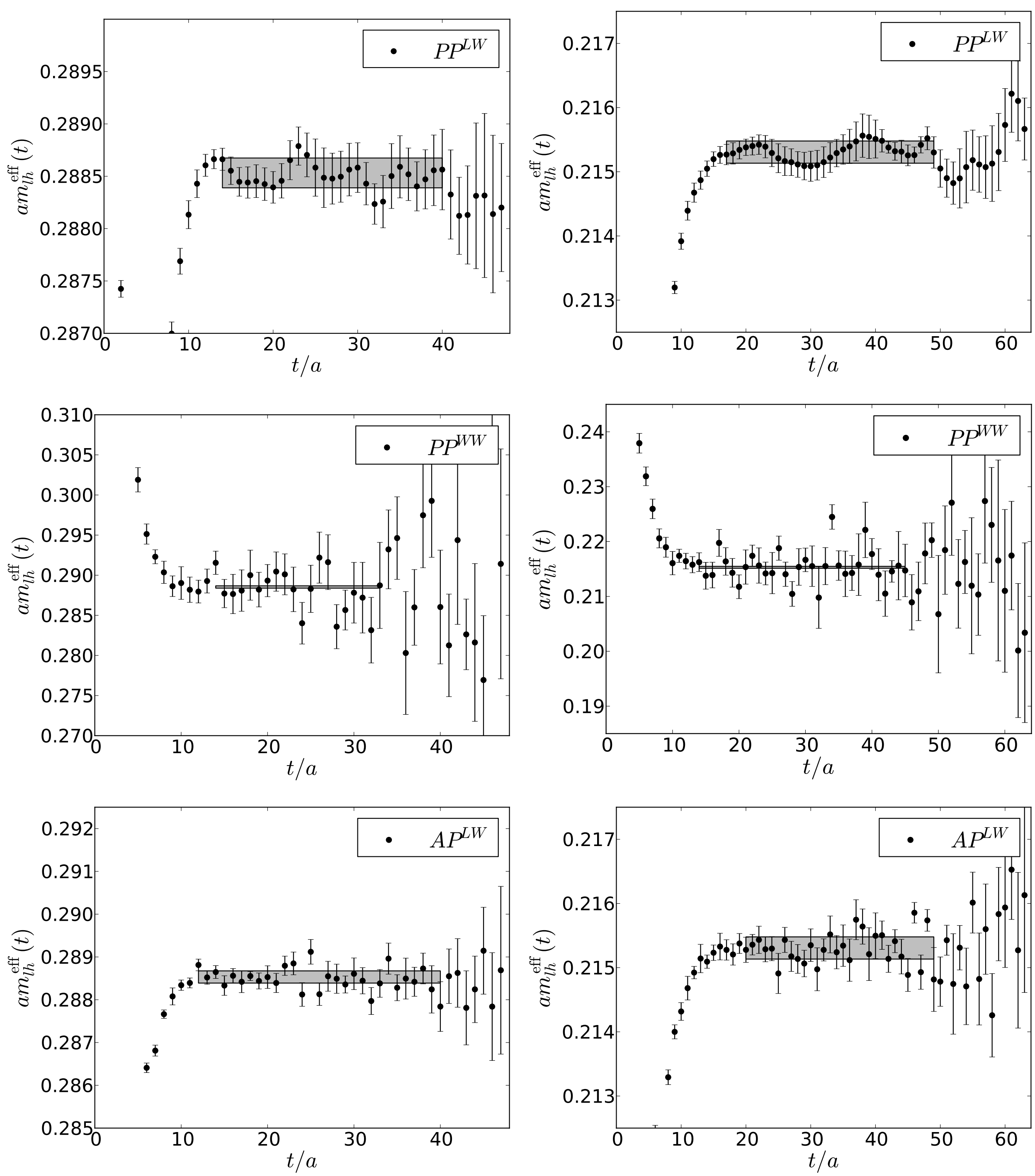


FIG. 10. Effective m_{lh} on the 48I (left column) and 64I (right column) ensembles. We fit a common value of the mass to all three correlators on a given ensemble.

combinations of these correlators one can remove the associated $\mathscr{O}(a)$ error and reduce the $\mathscr{O}(a^2)$ error. $Z_A/Z_{\mathscr{A}}$ is then computed via the following ratio: [29]

$$R(t) = \frac{1}{2}\left[\frac{C(t-\frac{1}{2})+C(t+\frac{1}{2})}{2L(t)} + \frac{2C(t+\frac{1}{2})}{L(t-1)+L(t+1)}\right]. \tag{43}$$

While the 1–2% $m_{\rm res}$ errors associated with the above determination of Z_A could be neglected in our earlier work, where we were far from the chiral limit and the statistical errors were larger than in the current work, in Refs. [6] and [30] it was shown that a better approximation could be obtained via the *vector* current. The local vector current operator formed from the domain wall surface fields is

$$V^a_\mu(x) = \overline{q}(x)\gamma_\mu \lambda^a q(x)\,, \tag{44}$$

which is related to the Symanzik vector current V^{Sa}_μ by a renormalization coefficient Z_V which was shown to be equal to Z_A up to terms $\mathscr{O}(m^2_{\rm res})$ [6]. There is also a five-dimensional conserved vector current $\mathscr{V}^a_\mu$ for which the renormalization factor, $Z_{\mathscr{V}}$, is unity, and we can obtain a significantly better approximation to Z_A by computing $Z_V/Z_{\mathscr{V}}$ on the lattice:

$$Z_A \approx \frac{Z_V}{Z_{\mathscr{V}}} = \frac{\langle 0|\sum_{\vec{x},i}\mathscr{V}^a_i(\vec{x},t)V^a_i(\vec{0},0)|0\rangle}{\langle 0|\sum_{\vec{x},i}V^a_i(\vec{x},t)V^a_i(\vec{0},0)|0\rangle}\,. \tag{45}$$

Below we determine both $Z_A/Z_{\mathscr{A}}$ and $Z_V/Z_{\mathscr{V}}$, but use only the latter to renormalize our decay constants.

1. Determination of $Z_A/Z_{\mathscr{A}}$

We introduce a practical approach to the conserved axial current for Möbius fermions in Appendix A and Ref. [31]. For the numerical determination of Z_A, the explicit construction of the current, used in Eq. (42), can be avoided with an alternate determination that utilizes the ratio of the divergences of the four-dimensional and five-dimensional axial currents:

$$Z_A \approx \frac{Z_A}{Z_{\mathscr{A}}} = \frac{\langle 0|\sum_{\vec{x}}\partial_\mu \mathscr{A}^a_\mu(\vec{x},t)|\pi\rangle}{\langle 0|\sum_{\vec{x}}\partial_\mu A^a_\mu(\vec{x},t)|\pi\rangle} = \frac{2m\langle 0|\sum_{\vec{x}} j^a_5(\vec{x},t)|\pi\rangle + 2\langle 0|\sum_{\vec{x}} j^a_{5q}(\vec{x},t)|\pi\rangle}{\langle 0|\sum_{\vec{x}}\partial_\mu A^a_\mu(\vec{x},t)|\pi\rangle}\,, \tag{46}$$

where the last equality follows from the PCAC relation, Eq. (41). We extract Z_A from our lattice data using the improved ratio

$$\begin{cases} C_{\mathscr{A}}(t) \equiv \langle 0| \sum_{\vec{x}} \partial_\mu \mathscr{A}_\mu^a(\vec{x},t) |\pi\rangle \\ C_A\left(t-\frac{1}{2}\right) \equiv \langle 0| \sum_{\vec{x}} \partial_\mu A_\mu^a(\vec{x},t) |\pi\rangle \\ Z_A^{\rm eff}(t) = \frac{1}{2}\left[\frac{C_{\mathscr{A}}(t-1)+C_{\mathscr{A}}(t)}{2C_A(t-\frac{1}{2})} + \frac{2C_{\mathscr{A}}(t)}{C_A(t+\frac{1}{2})+C_A(t-\frac{1}{2})}\right] \end{cases}, \tag{47}$$

which is also constructed to minimize errors at $\mathscr{O}(a^2)$ [29]. The translation by $\frac{1}{2}$ in the argument of the correlation function associated with A_μ^a arises from the divergence. The five-dimensional current $\mathscr{A}_\mu^a$, by contrast, is defined on the links between lattice sites, so its divergence is centered on the lattice. In Figure 11 we plot the effective Z_A and fit on each ensemble.

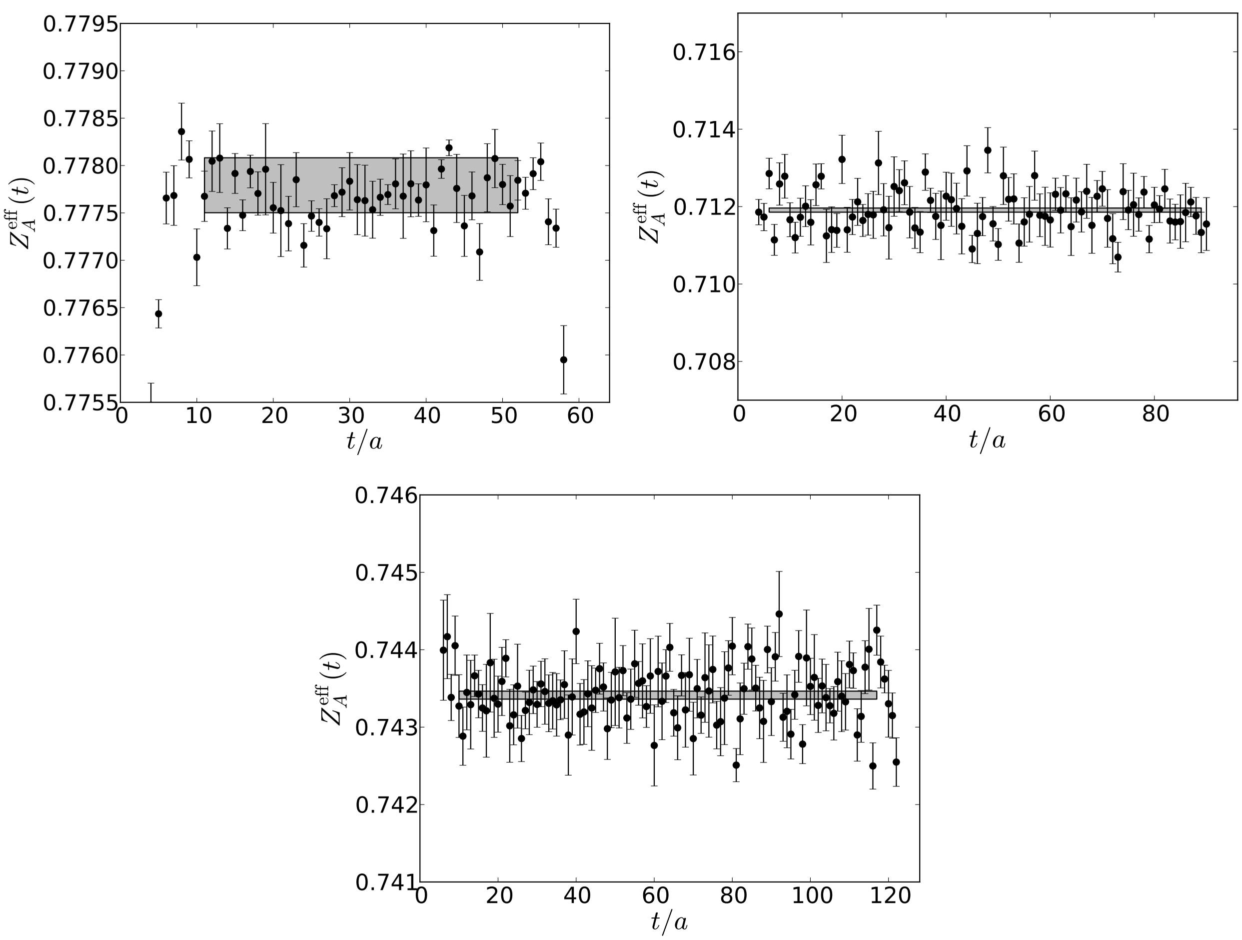


FIG. 11. Effective Z_A on the 32Ifine (top left), 48I (top right), and 64I (bottom) ensembles.

2. Determination of $Z_V/Z_{\mathcal{V}}$

Since the relatively noisy ρ meson is the lightest state to which the vector current couples, computing Z_V accurately requires a different approach from that used for Z_A (Eq. (46)). Instead, we calculate the pion electromagnetic form factors $f_{ll}^{+}(q^2)$ and $f_{ll}^{-}(q^2)$, defined by the matrix element

$$\langle\pi(p_1)|V_\mu|\pi(p_2)\rangle = f_{ll}^{+}(q^2)\,(p_2+p_1)_\mu + f_{ll}^{-}(q^2)\,(p_2-p_1)_\mu\,, \tag{48}$$

where $q = p_2 - p_1$ is the momentum transfer. Current conservation implies $f_{ll}^{-}(q^2) = 0$ for all q^2, leaving only the vector form factor, f_{ll}^{+}. For two pions at rest, $f_{ll}^{+}(0) = 1$, and we can fit Z_V from the temporal component of Eq. (48). We fit to the ratio

$$\frac{\tilde{\mathcal{C}}_{PP}^{WW}(t_{\rm snk})}{\mathcal{C}_{PVP}(t_{\rm src},t,t_{\rm snk})} \overset{t,|t_{\rm src}-t_{\rm snk}|\gg 1}{\simeq} Z_V\,, \tag{49}$$

where

$$\tilde{\mathcal{C}}_{PP}^{WW}(t) = \mathcal{C}_{PP}^{WW}(t) - \frac{1}{2}\mathcal{C}_{PP}^{WW}\left(\frac{N_t}{2}\right)e^{-m_{ll}(N_t/2-t)} \tag{50}$$

is the pion two-point function, Eq. (37), with the around-the-world state removed using the fitted pion mass, and $\mathcal{C}_{PVP}^{WW}(t_{\rm src},t,t_{\rm snk})$ is the three-point function defined by the matrix element, Eq. (48). On the 32Ifine and 48I ensembles, this matrix element was computed for all $\pi-\pi$ separations, $t_{\rm sink}-t_{\rm src}$, that are a multiple of 4. For the 64I ensemble we computed on separations that are multiples of 5. We determine the ranges of $\pi-\pi$ separations to use in the fit by plotting the midpoint of Eq. (49) as a function of the $\pi-\pi$ separation on each ensemble and looking for a plateau: based on this analysis we chose to include $\pi-\pi$ separations in the range 16–32 on the 32Ifine ensemble, 12–24 on the 48I ensemble, and 15–40 on the 64I ensemble. In Figure 12 we illustrate this method by plotting Eq. (49) for a single $\pi-\pi$ separation included in the fit, as well as the fitted value for Z_V, on each ensemble.

3. Determination of the Decay Constants

The light-light pseudoscalar decay constant can be computed from Z_V and the amplitudes of the PP and AP correlators as

$$f_{ll} = Z_V\sqrt{\frac{2}{m_{ll}V}\frac{{\mathcal{N}_{AP}^{LW}}^2}{\mathcal{N}_{PP}^{WW}}}\,, \tag{51}$$

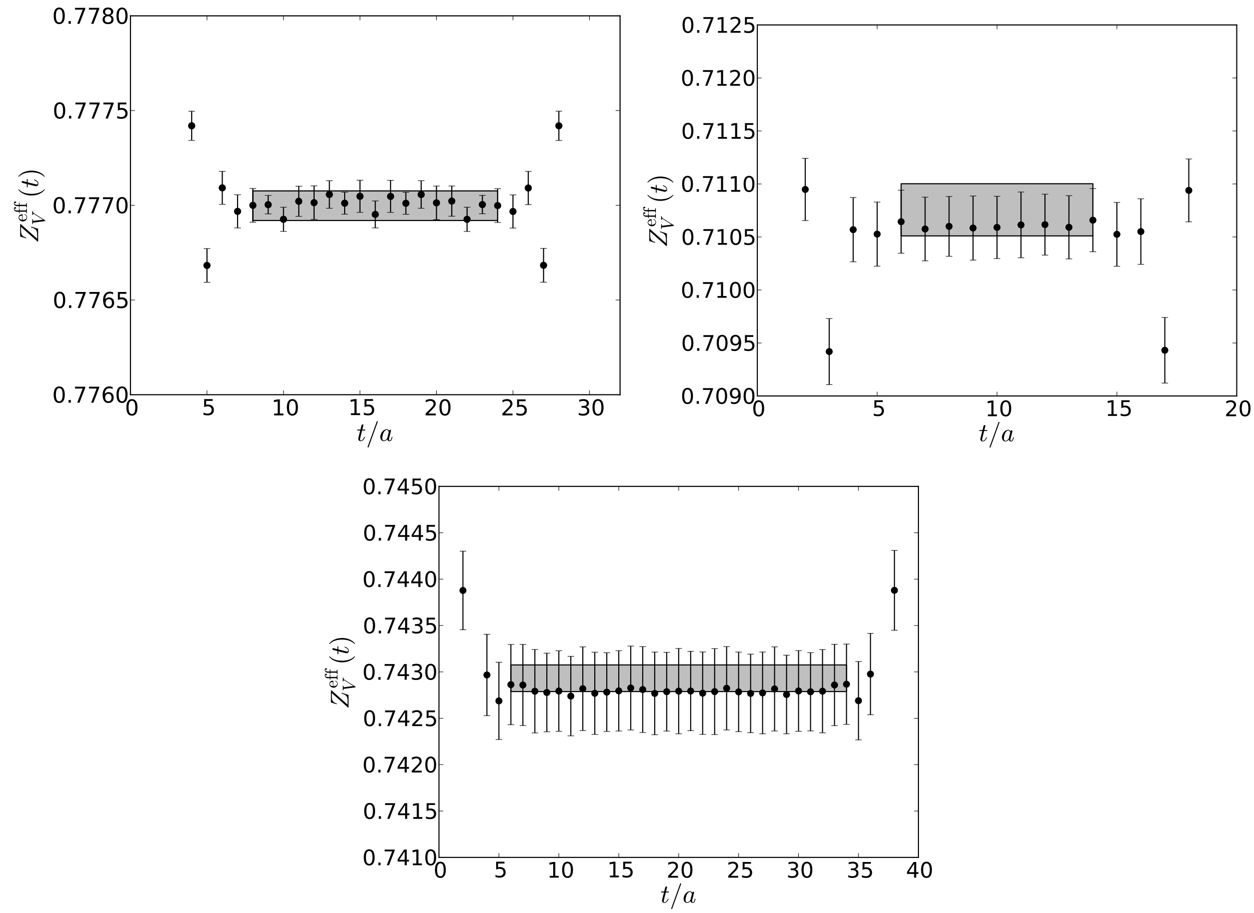


FIG. 12. Effective Z_V on the 32Ifine (top left), 48I (top right), and 64I (bottom) ensembles, for $\pi-\pi$ separations of 32 time units, 20 time units, and 40 time units, respectively. Note that in each case the fit is performed using several $\pi-\pi$ separations, not just the separation plotted here.

and likewise for the heavy-light pseudoscalar. In Figures 13 and 14 we plot the effective amplitudes,

$$\begin{cases} \mathcal{N}_{PP}^{\mathrm{eff}}(t) = \dfrac{\mathscr{C}_{PP}(t)}{\exp\left(-mt\right)+\exp\left(-m(N_t-t)\right)} \\ \mathcal{N}_{AP}^{\mathrm{eff}}(t) = \dfrac{\mathscr{C}_{AP}(t)}{\exp\left(-mt\right)-\exp\left(-m(N_t-t)\right)} \\ m = m^{\mathrm{eff}}(t)\,, \end{cases} \tag{52}$$

associated with f_{ll} and f_{lh}.

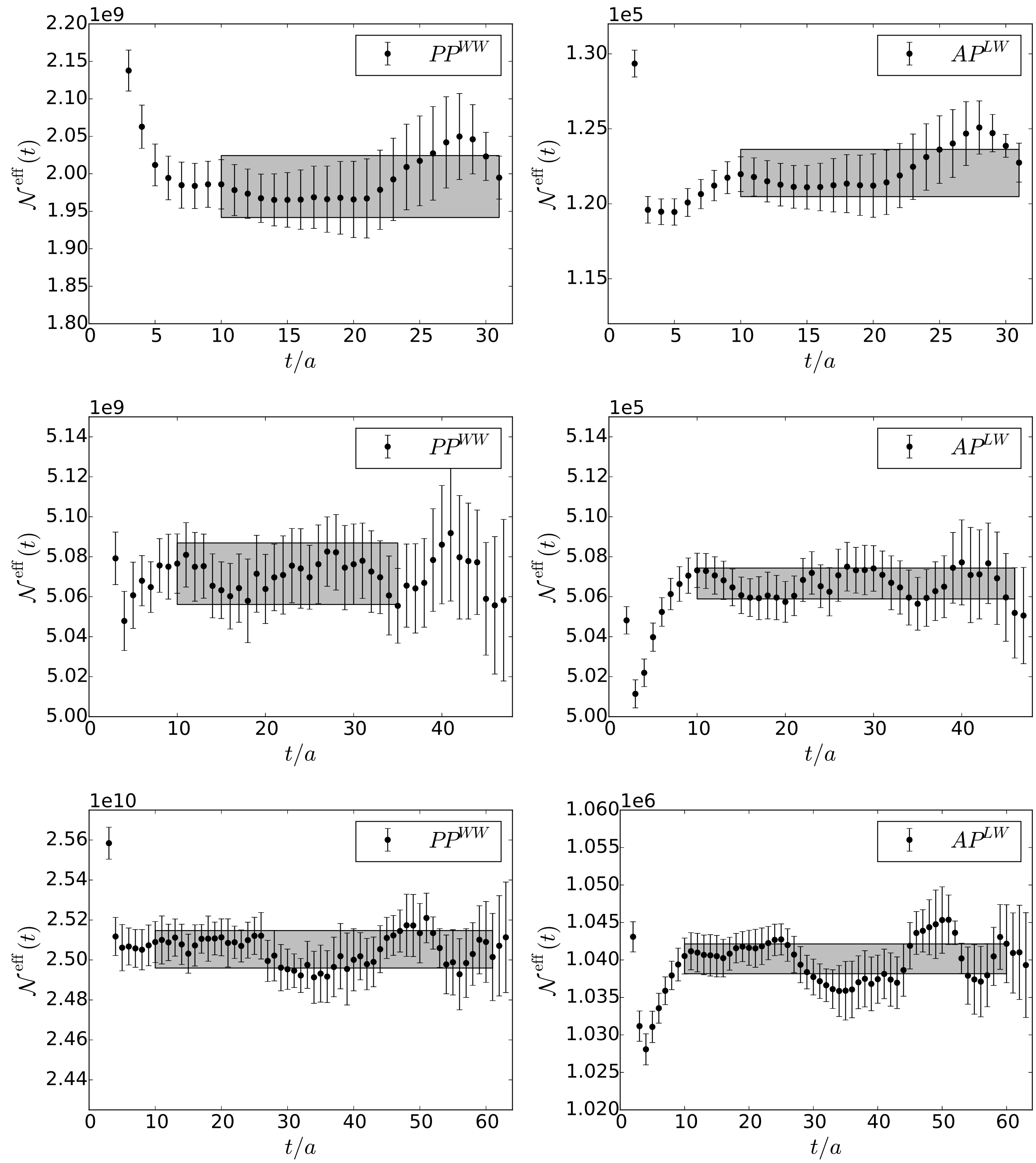


FIG. 13. Effective amplitudes, defined by Eq. (52), associated with f_{ll} on the 32Ifine (top), 48I (middle), and 64I (bottom) ensembles.

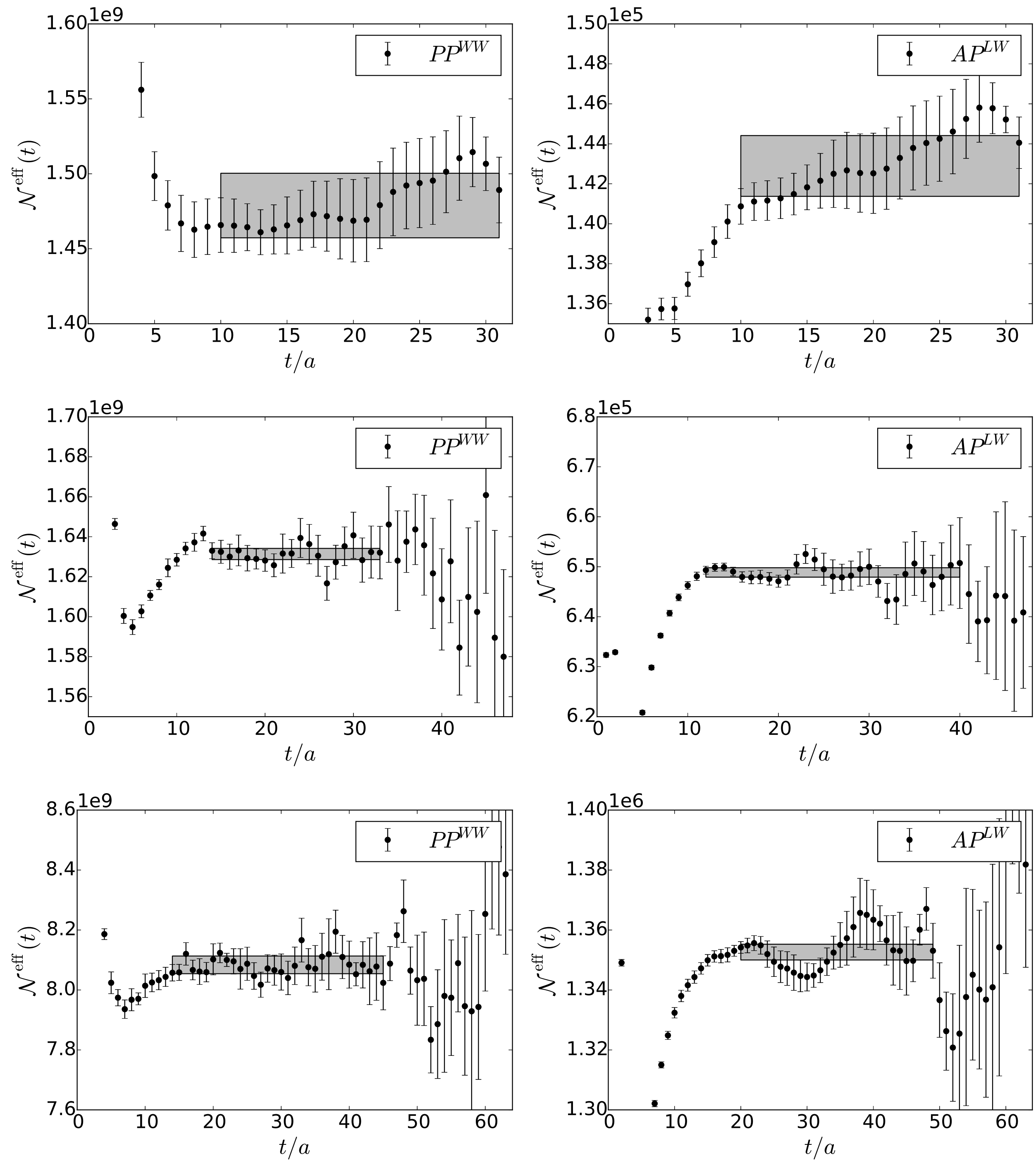


FIG. 14. Effective amplitudes, defined by Eq. (52), associated with f_{lh} on the 32Ifine (top), 48I (middle), and 64I (bottom) ensembles.

D. Neutral Kaon Mixing Parameter

We compute the neutral kaon mixing parameter, B_{lh}, from the ratio

$$\frac{\langle\overline{K^0}|\mathscr{O}_{VV+AA}|K^0\rangle}{\frac{8}{3}\langle K^0|A_0|0\rangle\langle 0|A_0|\overline{K^0}\rangle} \cong B_{lh}\,, \tag{53}$$

where $\mathscr{O}_{VV+AA}$ is the $\Delta S = 2$ four-quark operator responsible for the mixing:

$$\mathscr{O}_{VV+AA} = \bar{s}\gamma_\mu\,(\mathbf{1}-\gamma_5)\,d\cdot\bar{s}\gamma^\mu\,(\mathbf{1}-\gamma_5)\,d\,. \tag{54}$$

The matrix element in the numerator of Eq. (53) was computed for $K-\bar{K}$ separations which are a multiple of 4 (5) on the 32Ifine/48I (64I) ensemble. On the 32Ifine ensemble we use linear combinations of propagators with periodic and antiperiodic boundary conditions in the temporal direction to effectively double the time extent of the lattice for the B_{lh} correlators, a technique we have also employed in previous calculations [5]. We determine appropriate ranges of $K-\bar{K}$ separations to include in the fit using the same procedure as described in the previous section for Z_V. We chose separations of 52 and 56 time units on the 32Ifine ensemble, $20, 24\ldots 40$ on the 48I ensemble, and $25, 30\ldots 40$ on the 64I ensemble. In Figure 15 we plot the B_{lh} effective amplitude for a single $K-\bar{K}$ separation included in the fit, as well as the fitted value for B_{lh}, on each ensemble.

E. Omega Baryon Mass

We measured the Ω-baryon mass m_{hhh} from the two-point correlator

$$\mathscr{C}^{s_1s_2}_{\Omega\Omega}(t) = \sum_{i=1}^{3}\langle 0|\mathscr{O}^{s_1}_{\Omega}(\vec{x},t)_i\overline{\mathscr{O}}^{s_2}_{\Omega}(0)_i|0\rangle\,, \tag{55}$$

using an interpolating operator

$$\mathscr{O}_\Omega(x)_i = \varepsilon_{abc}\left(s^{\mathsf{T}}_a(x)C\gamma_i s_b(x)\right)s_c(x), \tag{56}$$

where C denotes the charge conjugation matrix. We performed measurements using both Coulomb gauge-fixed wall sources and Z_3 box (Z_3B) sources, and, in both cases, a local (point) sink. The correlator, Eq. (55), is a 4×4 matrix in spin space which couples to both positive $(+)$ and negative $(-)$ parity states, and has the asymptotic form

$$\mathscr{C}^{s_1s_2}_{\Omega}(t) \overset{t\gg 1}{\cong} \sum_{\vec{p}}\left(\frac{1}{2}(\mathbf{1}+\gamma_4)\,\mathscr{A}^{s_1s_2}_{+}(\vec{p})e^{-E^{+}_{\vec{p}}t} - \frac{1}{2}(\mathbf{1}-\gamma_4)\,\mathscr{A}^{s_1s_2}_{-}(\vec{p})e^{-E^{-}_{\vec{p}}t}\right) \tag{57}$$

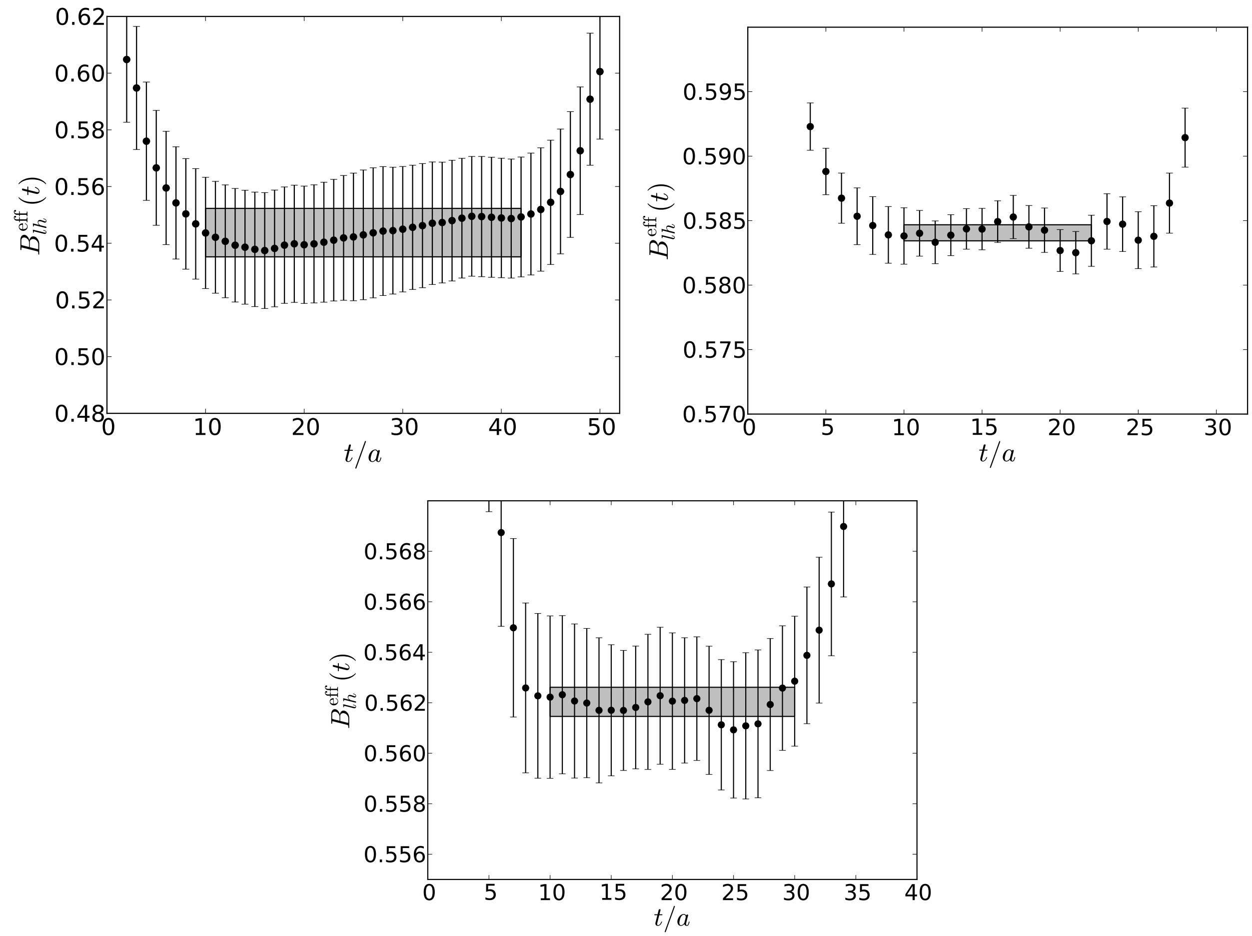


FIG. 15. Effective B_{lh} on the 32Ifine (top left), 48I (top right), and 64I (bottom) ensembles, for $K-K$ separations of 52 time units, 32 time units, and 40 time units, respectively. Note that in each case the fit is performed using several $K-\bar{K}$ separations, not just the separation plotted here.

for large t. The fit to extract m_{hhh} is performed by first projecting onto the positive parity component,

$$\mathscr{P}_{+}\mathscr{C}_{\Omega}^{s_1 s_2} = \frac{1}{4}\,\mathrm{tr}\left\{\frac{1}{2}\left(\mathbf{1}+\gamma_4\right)\mathscr{C}_{\Omega}^{s_1 s_2}\right\}, \tag{58}$$

for each source type, and then performing a simultaneous fit of both correlators to a sum of two exponential functions with common mass terms :

$$\begin{cases} \mathscr{C}_{\Omega\Omega}^{LW}(t) = \mathscr{N}_{\Omega\Omega}^{LW}\, e^{-m_{hhh}t} + \tilde{\mathscr{N}}_{\Omega\Omega}^{LW}\, e^{-m'_{hhh}t} \\ \mathscr{C}_{\Omega\Omega}^{LZ_3B}(t) = \mathscr{N}_{\Omega\Omega}^{LZ_3B} e^{-m_{hhh}t} + \tilde{\mathscr{N}}_{\Omega\Omega}^{LZ_3B} e^{-m'_{hhh}t} \end{cases}. \tag{59}$$

One can also include terms proportional to $e^{-m_{-}(N_t-t)}$, where m_{-} is the mass of the ground state in the negative parity channel, to account for around-the-world contamination effects, but we find that our lattices are sufficiently large and the masses of these states sufficiently heavy that including these terms has no statistically significant influence on the fitted Ω mass. Using multiple source types and double-exponential fits to common masses allows us to reduce the statistical error on the Ω baryon mass m_{hhh}, as well as to also fit the mass of the first excited state in the positive parity channel m'_{hhh}. Figure 16 plots the effective Ω-baryon mass on each ensemble.

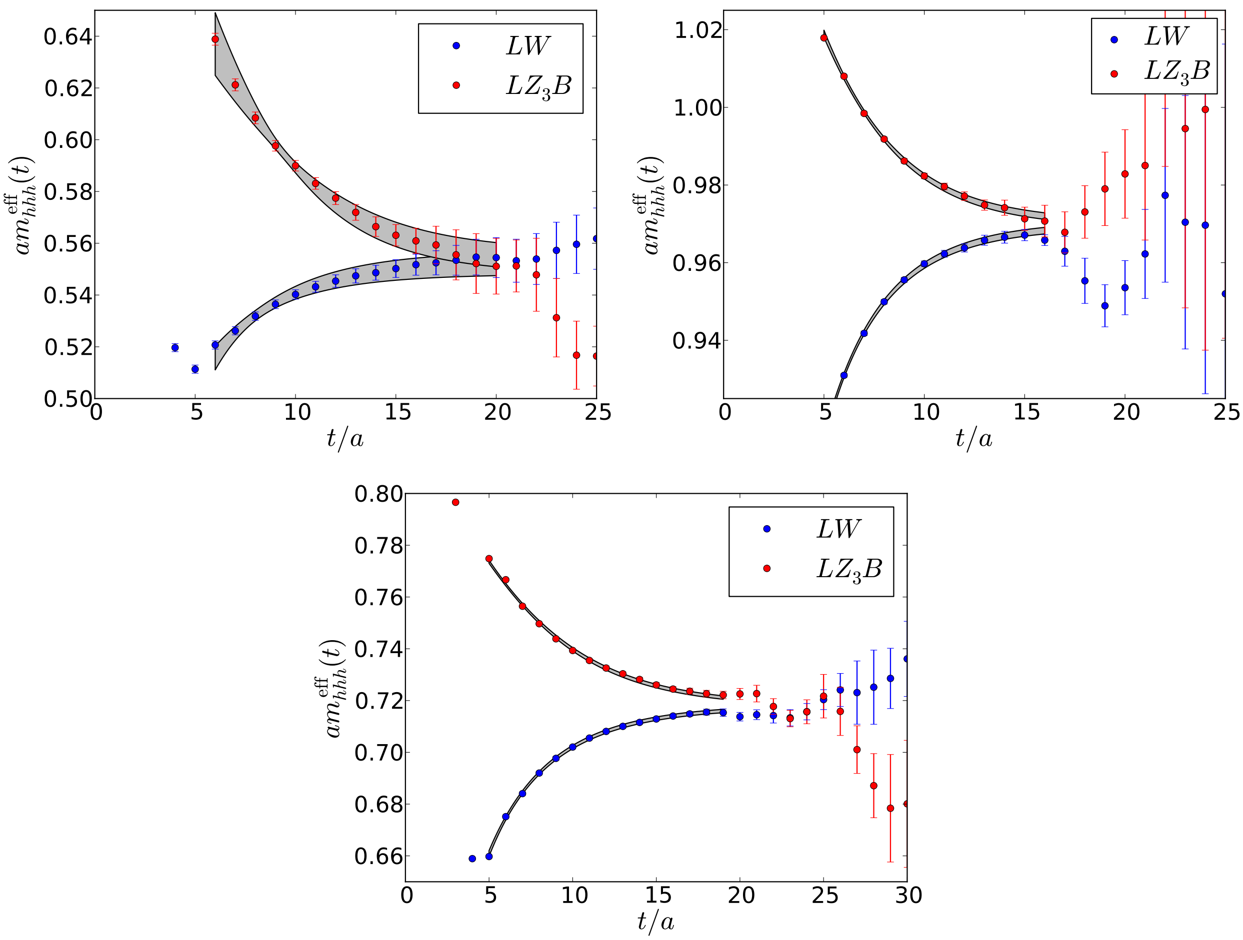


FIG. 16. The effective mass of the Omega baryon obtained using both our wall (LW) and Z_3 box source (LZ_3B) on the 32Ifine (top left), 48I (top right), and 64I (bottom) ensembles. The correlation functions are simultaneously fit to a two-exponential fit form, and the effective mass determined from the fit function (obtained by applying the same technique as used to extract the effective mass from the raw data) is overlayed with the data.

In Figure 17 we plot the dependence of our fitted ground and excited state energies on the lower temporal bound of the fit. The upper bound of the fit window is fixed at 20, 16, and 19 on the 32Ifine, 48I, and 64I ensembles, respectively. We observe excellent stability for bounds above

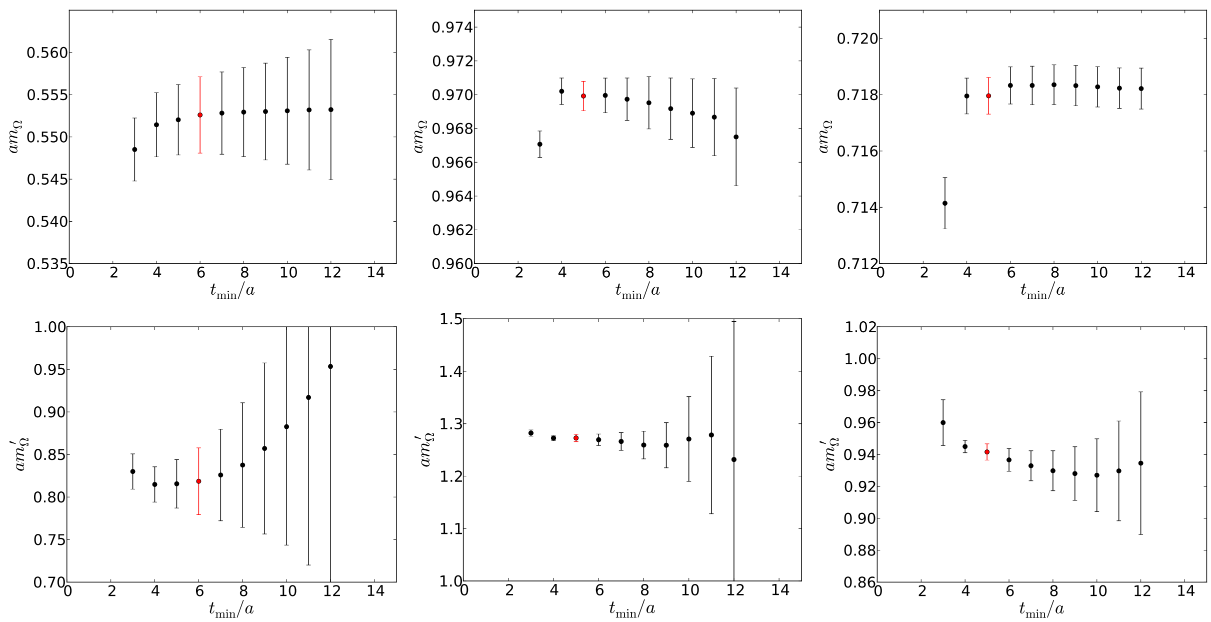


FIG. 17. The stability, as a function of the lower bound on the fit range $t_{\min}/a$, of our fitted Omega baryon ground state (upper row) and first-excited state (lower row) for the 32Ifine (left), 48I (middle) and 64I (right) ensembles. The point in red indicates our final value.

$t_{\min} = 4$, suggesting that we have good resolution on both the ground and excited states, and that contamination of our results by higher-energy excited states can be discounted. In practice we use $t_{\min} = 5$ for both the 48I and 64I ensembles, and $t_{\min} = 6$ for the 32Ifine ensemble.

F. Wilson flow scales

The Wilson flow scales, $t_0^{1/2}$ and w_0, are quantities with the dimension of length defined via the following equations: [32]

$$t^2\langle E(t)\rangle|_{t=t_0} = 0.3\,, \tag{60}$$

and [33]

$$t\frac{d}{dt}(t^2\langle E(t)\rangle)|_{t=w_0^2} = 0.3\,, \tag{61}$$

where E is the discretized Yang-Mills action density,

$$E = \frac{1}{2}\mathrm{tr}(F_{\mu\nu}F_{\mu\nu})\,. \tag{62}$$

We determine the action density using the clover discretization, for which $F_{\mu\nu}$ is estimated at each lattice site from the clover of four 1×1 plaquettes in the $\mu-\nu$ plane. We find that this leads to smaller discretization errors (especially for t_0) than estimating $F_{\mu\nu}$ directly from the plaquette via

$$\langle P\rangle = 1 - \frac{a^4}{36}\langle E\rangle + O(a^6) \tag{63}$$

which is in agreement with some previous experience [32]. In Figure 18 we show an example of the interpolation of the two scales on the 64I ensemble. The final results for all ensembles are listed in Table VI.

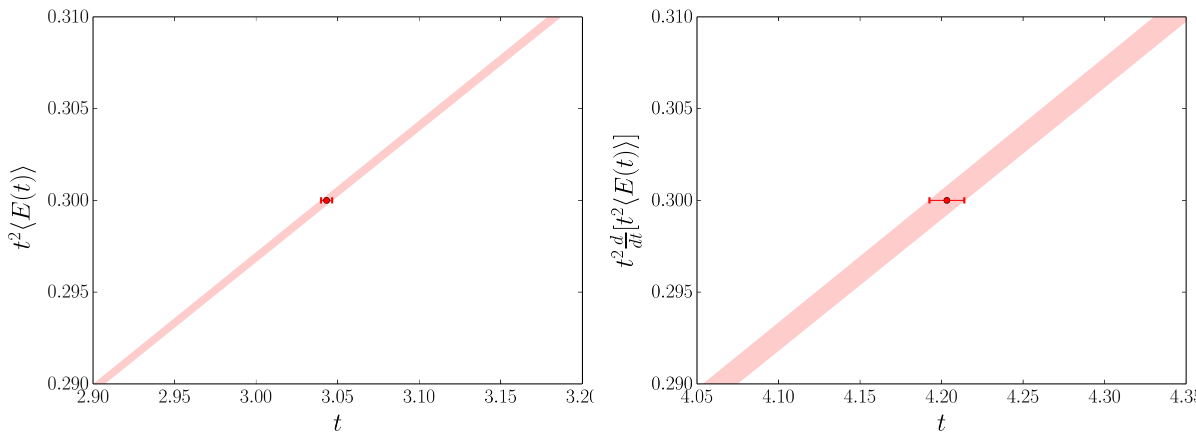


FIG. 18. The interpolation in Wilson flow time t on the 64I ensemble of the functions of the action density used to define t_0 and w_0 respectively. The red point is the interpolated value.

IV. SIMULTANEOUS CHIRAL/CONTINUUM FITTING PROCEDURE

The bare quark masses for the 48I and 64I ensembles were chosen based on the results for the physical quark masses at equivalent bare couplings obtained in Ref. [5]. The simulated values for the dimensionless ratios m_π/m_Ω and m_K/m_Ω are shown in Table VI. Since we are not simulating electromagnetism, we compare to the following physical values: $m_\pi = 135.0$ MeV, $m_K = 495.7$ MeV and $m_\Omega = 1.67225$ GeV. Clearly our simulations are very close to the physical point, yet we must perform the very modest extrapolation in order to obtain precise physical results.

A. Summary of global fit procedure

In Refs. [5, 6] we have detailed a strategy for performing simultaneous chiral and continuum 'global' fits to our lattice data. In this document we perform such fits to the following quantities:

m_π, m_K, f_π, f_K, m_Ω and the Wilson flow scales $t_0^{1/2}$ and w_0. We parametrize the mass dependence of each quantity using three ansätze (where applicable): NLO partially-quenched chiral perturbation theory with and without finite-volume corrections (i.e. infinite volume χPT), which we henceforth refer to as the 'ChPTFV' and 'ChPT' ansätze respectively; and a linear 'analytic' ansatz. For the ChPT and ChPTFV ansätze we use heavy-meson χPT [34, 35] to describe quantities with valence strange quarks. For the convenience of the reader, we have collected the various ChPT and analytic fit forms in Appendix H. In this appendix we also specify the new fit functions that we use to describe the Wilson flow scales, $t_0^{1/2}$ and w_0.

We use the difference between the results obtained for each ansatz to estimate our systematic errors. In order to account for discretization effects, we include in each fit form an a^2 term. As discussed in Ref. [5], we neglect higher order effects including terms in a^4 and $a^2\ln(a^2)$. The fits are performed to dimensionless data, with the parameters determined in the bare normalization of a reference ensemble r. The bare lattice quark masses and data on other ensemble sets are 'renormalized' into this scheme via additional fit parameters: For an ensemble e, these are Z_l^e, Z_h^e for normalizing the light and heavy quark masses respectively, and R_a^e for the scale. These are defined as follows:

$$Z^e_{l/h} = \frac{1}{R^e_a}\frac{(a\tilde{m}_{l/h})^r}{(a\tilde{m}_{l/h})^e} \text{ and } R^e_a = a^r/a^e\,, \tag{64}$$

where a is the lattice spacing and $\tilde{m} = m + m_{\mathrm{res}}$. Note that the scheme used for the quark masses is implicitly mass dependent, hence we allow for different parameters to renormalize the heavy (Z_h) and light (Z_l) quarks. In practice this dependence is very weak and Z_l and Z_h differ only at the percent level even on our coarsest lattices (cf. Table XVII) despite the order of magnitude difference in the mass scales. Within a large range of light quark masses we previously observed no measurable dependence [6], which motivated our choice to obtain these quantities as free parameters in the global fit ('generic scaling') rather than by matching at a single mass ('fixed trajectory').

The procedure for obtaining the general dimensionless fit form for a quantity Q is described in Appendix B.

We choose a continuum scaling trajectory along which m_π/m_Ω and m_K/m_Ω match their physical values. Here we include the Ω baryon mass due to the ease of obtaining an precise lattice measurement and its simple quark mass dependence. This procedure defines m_π, m_K and m_Ω as having no lattice spacing dependence. After performing the fit, we obtain the lattice spacing for the reference ensemble by comparing the value of any of the aforementioned quantities to the corresponding physical value after extrapolating to the physical quark masses. The lattice spacings for the other ensembles are then obtained by dividing this value by R_a^e. An alternate choice of scaling trajectory, for example using f_π in place of m_Ω, would reintroduce the scale dependence on m_Ω and remove it from f_π; the values of each a^2 coefficient are therefore dependent on the choice of scaling trajectory, but the continuum limit is guaranteed to be the same (up to our ability to measure and extrapolate the quantities in question). Note that the inclusion of the Wilson flow data results in significant improvements in the statistical error on the lattice spacings compared to our previous determinations due to its influence on the shared ratios R_a.

While the data on a given ensemble can be expected to be highly correlated, the estimated correlation matrices tend to suffer from having a poor condition number preventing their use in correlated fits. As a result, our global fits are performed assuming a diagonal correlation matrix. This approach can result in larger jackknife statistical errors than for correlated fits, however in the past [35] we have experimented with performing partially-correlated fits where increasingly large numbers of leading eigenvectors were included in the estimate, and found little difference between the uncorrelated and correlated results. With uncorrelated fits the $\chi^2/\mathrm{d.o.f}$ may not be a reliable indicator of the goodness of fit, and to assess their quality we instead generate histograms of the deviation between the data and the fit.

B. Details specific to this calculation

Using our simultaneous fit strategy, we combine our 64I and 48I physical point ensembles with a number of existing domain wall ensembles: the ‘24I’ and ‘32I’ ensembles with lattice volumes $24^3 \times 64 \times 16$ and $32^3 \times 64 \times 16$ and Shamir domain wall fermions with the Iwasaki gauge action at bare couplings $\beta = 2.13$ and 2.25 respectively (equal to the 48I and 64I bare couplings respectively) described in Refs. [35] and [6]; the ‘32ID’ ensembles with volume $32^3 \times 64 \times 32$ and Shamir domain wall fermions with the Iwasaki+DSDR gauge action at $\beta = 1.75$ described in Ref. [5]; and finally the ‘32Ifine’ ensemble with volume $32^3 \times 64 \times 12$ and Shamir domain wall fermions with the Iwasaki gauge action at $\beta = 2.37$ described in this document. For the convenience of the reader, we summarize the input parameters of the 24I, 32I and 32ID ensembles along with a number of relevant quantities including the range of pion masses, the lattice spacing and physical lattice size, in Table IX.

Following our earlier analyses, we use the 32I ensemble set as the reference ensemble against

	32I	24I	32ID
Size	$32^3\times 64\times 16$	$24^3\times 64\times 16$	$32^3\times 64\times 32$
Action	Shamir DWF + I	Shamir DWF + I	Shamir DWF + ID
β	2.13	2.25	1.75
a^{-1}(GeV)	2.383(9)	1.785(5)	1.378(7)
L (fm)	2.649(10)	2.653(7)	4.581(23)
$m_\pi L$	3.122(12)	3.339(15)	3.335(7)
m_π unitary (MeV)	302.4(1.2)–360.1(1.4)	339.7(1.3)–339.7(1.3)	172.4(0.9)–315.5(1.6)
m_π lightest PQ (MeV)	232.4(1.1)	248.3(1.2)	143.8(0.8)

TABLE IX. Input parameters and relevant quantities for the 32I, 24I and 32ID ensembles. For the action, I stands for the Iwasaki gauge action, and ID the Iwasaki action with the DSDR term. Here L is the spatial lattice extent and $m_\pi L$ is given for the lightest partially-quenched pion at the simulated strange quark mass. The last two rows list the range of unitary pion masses and the lightest partially-quenched pion mass (PQ) mass, respectively. The full set of corresponding bare quark masses are given in Table XI. The lattice spacings used here are determined in Section V B

which the 'scaling parameters', $Z_{l/h}$ and R_a, are defined.

1. Ensemble-specific parameters

As discussed in Section II, the Möbius parameters of the 48I and 64I ensembles are chosen to ensure the equivalence of the Möbius and Shamir kernels; as a result, the ensembles with the Iwasaki gauge action can all be described by the same continuum scaling trajectory, i.e. with the same a^2 scaling coefficients. As described in Ref. [5], additional parameters must be introduced to describe the lattice spacing dependence of the 32ID ensembles, which use the Iwasaki+DSDR gauge action to suppress the dislocations that enhance the domain wall residual chiral symmetry breaking on this coarse lattice.

Note that while the 32ID ensemble is the only data set with the Iwasaki+DSDR gauge action, the five additional a^2 terms for f_π, f_K, w_0, $t_0^{1/2}$ and B_K, are completely determined from the overall relative normalization of these data under the χ^2 minimization. This leaves more than sufficient data to determine Z_l, Z_h and R_a on this ensemble set and to help constrain the coefficients of the mass terms that are common to all ensembles in these fits. Although the 32ID ensemble set is coarse ($a^{-1} = 1.38(1)$ GeV), we observe discretization effects only at the 5% level suggesting a discretization systematic error arising from higher order ($\mathcal{O}(a^4)$) terms at the 0.25% scale, small enough to be neglected. The inclusion of these ensembles in the global fits is discussed at length in Ref. [5].

The 48I and 64I ensembles have identical bare couplings to the 24I and 32I ensembles respectively, yet differ in their values of the total quark mass, L_s and Möbius scale parameter α. The change in residual chiral symmetry breaking resulting from the changes in L_s and α gives rise to a shift in the bare mass parameter of the low-energy effective Lagrangian, which we account for at leading order in our fits by renormalizing the quark masses as $\tilde{m} = m + m_{\rm res}$. Higher order effects such as those of order $m_{\rm res}a^2$ are small enough to be ignored. After performing this correction we might assume that the scaling parameters Z_l, Z_h and R_a (or equivalently the lattice spacing) for the 48I and 24I ensembles should be identical, and likewise for the 64I and 32I ensembles. However when we performed our global fits we found that the 48I lattice spacing is $3.2(2)\%$ larger than that of the 24I ensemble, and the 64I lattice spacing is $1.1(2)\%$ larger than the 32I value. We saw no statistically discernible differences in Z_l and Z_h.

As we mentioned in Section II A and discuss in detail in Appendix C, the observed change in the lattice spacings can be expected to originate from the changes in the effective extent of the fifth dimension, $L_s' = \alpha L_s$, which differs by a factor of 3 between the 48I/24I ensembles, and a factor of 1.5 between the 64I/32I. At finite L_s' the Symanzik effective Lagrangian contains the leading-order operator $\beta_{\rm eff}{\rm tr}(F_{\mu\nu}F_{\mu\nu})$. A change in L_s' which causes a 0.0025 change in $m_{\rm res}$ should also be expected to cause a ≈ 0.0025 change $\beta_{\rm eff}$, a change which results in an exponentially-enhanced change in the resulting lattice spacing. Recall that the 5.6% change in the coupling between a $\beta = 2.13$, $a^{-1} = 1.75$ GeV ensemble and a $\beta = 2.25$, $a^{-1} = 2.38$ GeV ensemble, gives rise to a 36% change in the inverse lattice spacing. Thus, we might expect a 3% change in a^{-1} to result from a 0.5% change in the effective coupling, not far from the change we observe. We discuss in Appendix C how changes of this size are not unreasonable, and provide additional numerical evidence for the observed change in lattice scale.

Finite L_s' effects will also give rise to other higher order effects of a similar size. For example, we might expect $\mathcal{O}(0.5\%)$ shifts in the a^2 scaling coefficients of the various quantities included in our global fits. However, in Section V we find that even on the coarser 48I ensemble, the discretization effects are only at the 2-3% level (cf. Table XIII), suggesting negligible, 0.02% finite

Scheme	48I	24I	% diff.	64I	32I	% diff.
$\not{q}$	1.43613(80)	1.4386(12)	0.17%	1.43998(80)	1.4396(37)	0.03%
γ_μ	1.52070(89)	1.5235(13)	0.18%	1.51764(98)	1.5192(39)	0.1%

TABLE X. A comparison of the quark mass renormalization factors Z_m between the 48I/24I and 64I/32I pairs of ensembles, giving the values and their percentage difference. The renormalization scale is 3 GeV and the definitions of the schemes are given in Appendix F alongside details of the computation of the 24I and 32I values. Those for the 48I and 64I are not used later in the analysis and are presented here only for comparison. Note that unlike the 24I and 32I values, those for the 48I and 64I ensembles are not extrapolated to the chiral limit as they are computed at only a single mass but for other ensembles we have observed no significant mass dependence for these non-exceptional schemes.

L_s' effects. We again emphasize that for our large values of L_s', it is only the exponentially enhanced dependence of the lattice spacing upon the Symanzik coefficients that gives rise to observable finite-L_s' dependence in this quantity. We do not expect any other observable effects.

Additional evidence for the closeness of our Möbius and Shamir ensembles can be obtained by comparing the renormalization factors for the quark masses, Z_m, and the kaon bag parameter, Z_{B_K}. The former are computed for the 32I and 24I ensembles in Appendix F, for use in obtaining renormalized physical quark masses later in this document. There we do not present the computation of the corresponding factors for the 48I and 64I ensembles as they are not needed in our later analysis. Nevertheless, we have computed these values, and we list them alongside the 24I and 32I numbers in Table X. We observe only tiny, 0.2% scale differences between the 48I/24I values and even smaller $< 0.1\%$ differences for the 64I/32I ensembles. Comparing the values for Z_{B_K} in Table XLII we again see differences only at the 0.25% scale. This strongly suggests that finite-L_s effects have no significant impact upon the UV physics other than through the exponentially enhanced dependence of the lattice spacing upon a shift in the bare coupling at the 0.5% scale. In addition, these observations justify our fixing both Z_l and Z_h to be the same for the 24I and 48I ensembles, and also for the 32I and 64I ensembles.

2. Weighted global fits

The fits are performed independently for each superjackknife sample J by minimizing χ^2_J under changes in the set of fit parameters $\mathbf{c}_J$ of the function f. χ^2_J is defined as

$$\chi^2_J = \sum_i \frac{[y_{iJ} - f(\mathbf{x}_{iJ}, \mathbf{c}_J)]^2}{\sigma_i^2} \tag{65}$$

were y_{iJ} is the J^{th} superjackknife sample of a measurement i and $\mathbf{x}_{iJ}$ are the associated input parameters (quark masses, etc). σ_i is the error on the measurement, and provides the weight of each data point in the fit.

The naïve χ^2-minimization procedure weights each data point according to just its statistical error, and is therefore unable to account for systematic uncertainties on the *fit function* itself. Given that NLO χPT can only be expected to be accurate to $O(5\%)$ in the 200 - 370 MeV pion-mass range in which the majority of our data lies, the fits over-weight the data in this heavy-mass region resulting in deviations of the fit curve from the light-mass data. In practice the enhanced precision of the near-physical 64I and 48I data partially compensates for the larger number of heavy-mass data points, resulting in only $\mathcal{O}(1\sigma - 2\sigma)$ deviations between these data and the fit curve. However, as the intention of these global fits is only to perform a few-percent mass extrapolation of our near-pristine data, such deviations are unacceptable.While this can be remedied to a certain degree by removing data from the heavy-mass region, there remains pollution from the systematic uncertainty of the fit form. Without going to full NNLO χPT, one might attempt to reduce this uncertainty by introducing physically motivated 'nuisance parameters', perhaps along with Bayesian constraints to confine them within sensible bounds. While this is certainly a valid approach we feel it to be beyond the scope of this work, given that we desire only to perform a small correction to our near-physical data. With this in mind, we instead adopt an alternative approach in which we force the fit curve to pass through our near-physical data by increasing the weight of these data in the χ^2 minimization as follows.

We introduce a measurement-dependent weighting factor ω_i to the χ^2 determination:

$$\chi^2_J = \sum_i \frac{\omega_i [y_{iJ} - f(\mathbf{x}_{iJ}, \mathbf{c}_J)]^2}{\sigma_i^2}\,. \tag{66}$$

Note that only the relative values of ω_i matter as the same parameters that minimize χ^2 will also minimize $r\chi^2$, where r is some common factor. (Of course the algorithm itself has some numerical stopping condition which will need to be adjusted to take into account the change in normalization of χ^2.) In principle one could tune the relative weights based on a combination of the measured statistical error and an estimate of the systematic error of the fit function at each point, but this runs the risk of becoming too complex and arbitrary. Instead, as previously mentioned, we weight the data such that the fit is forced to pass directly through the data points on the 48I and 64I ensembles. To achieve this, we set $\omega_i = \Omega$ for those data, where Ω is assumed to be large, and $\omega_i = 1$ for the remainder. This is performed independently for each superjackknife sample, and does not change the fluctuations on the data *between* superjackknife samples. As a result, the statistical error from the overweighted points is unchanged by this procedure. In Appendix D we demonstrate that the fit results become independent of Ω in the limit $\Omega \to \infty$ and that the procedure has the desired effect of forcing the fit through the physical point data.

For large values of Ω we must choose small values of the numerical stopping condition on the minimization algorithm, increasing the time to perform the fit and making it more susceptible to finite-precision errors. In the aforementioned appendix we determine that $\Omega = 5000$ and a stopping condition of $\delta\chi^2_{\rm min} = 1 \times 10^{-4}$ is sufficient.

We emphasize that this procedure is performed separately for each superjackknife sample of our combined data set, such that the error on the fit function evaluated at the parameters associated with the 64I and 48I data is exactly equal to the error on the corresponding data. This can be seen, for example, in Figure 23 of Section V B, where we see the 1σ width of the fit curve exactly aligns with the error bars for the 48I and 64I data.

V. FIT RESULTS AND PHYSICAL PREDICTIONS

We performed global fits using the ChPTFV, ChPT and analytic ansätze. As discussed in Ref. [5], we attempt to separate the finite-volume and chiral extrapolation effects by performing the analytic fits to data that is first corrected to the infinite-volume using the ChPTFV fit results. Following Ref. [5], the ChPTFV and ChPT fits were performed with a 370 MeV pion mass cut on the data (this is set slightly larger than the value used in that paper, as we wish to include in our fit the 32Ifine data with a 371(5)MeV pion). The criteria for excluding the other fitted data are as follows: For f_π we exclude the data if the pion mass with the same set of partially-quenched quark masses lies above the cut; for f_K and m_K data points with light valence quark mass m_x and heavy mass m_y, we exclude the data if the pion with $m_x = m_y$ on that ensemble is above the pion mass cut; and for m_Ω, $t_0^{1/2}$ and w_0 we exclude the data only if the unitary pion on that ensemble is also excluded.

Ensemble set	m_l	m_y	$\{m_x\}$
32I	0.006	0.006	0.006, 0.004, 0.002
		0.004	0.004, 0.002
		0.002	0.002
	0.004	0.006	0.006, 0.004, 0.002
		0.004	0.004, **0.002**
		0.002	**0.002**
24I	**0.005**	0.005	0.005, 0.001
		0.001	**0.001**
32ID	**0.0042**	**0.008**	0.008, 0.0042, **0.001**, **0.0001**
		0.0042	**0.0042**, **0.001**, **0.0001**
		0.001	**0.001**, **0.0001**
		0.0001	**0.0001**
	0.001	**0.008**	0.008, 0.0042, **0.001**, **0.0001**
		0.0042	**0.0042**, **0.001**, **0.0001**
		0.001	**0.001**, **0.0001**
		0.0001	**0.0001**

TABLE XI. The bare light quark masses for the m_π and f_π data on our older 32I, 24I and 32ID ensembles that we included in our global fits with the 370 MeV pion mass cut. Data in bold are those included in the fits with the lower, 260 MeV cut. Here m_l is the sea light mass, and m_x and m_y are the (partially-quenched) valence masses. The final column gives the full set of available m_x values. Note, each of these points are computed with four different sea strange quark masses that are given in Table XII.

We consider two different pion mass cuts for the analytic fits: the 370 MeV cut used for the ChPTFV and ChPT fits, and a lower, 260 MeV cut. In our previous work we determined that the analytic fits were not able to accurately describe the data over the range from the physical point to the heaviest data, forcing us to use the lower cut. However, in the present analysis the fit predictions are dominated by the near-physical data due to the overweighting procedure, and these data require only a small, percent-scale, chiral extrapolation to correct to the physical light quark mass. This can be seen in Table XIII, in which we list the sizes of the various corrections required to obtain the physical prediction. We therefore also perform analytic fits with the 370 MeV cut,

Ensemble set	m_h^{sim}	$\{m_h^{\mathrm{rw}}\}$	$\{m_h^{\mathrm{val}}\}$
32I	0.03	0.029, 0.028, 0.027	0.03, 0.025
24I	0.04	0.03775, 0.0355, 0.03325	0.04, 0.03
32ID	0.045	0.0455, 0.046, 0.0465	0.035, 0.045, 0.055

TABLE XII. Strange quark masses in the valence and sea sectors on our older 32I, 24I and 32ID ensembles. The second column gives the simulated strange mass, and the third column the subset of reweighted strange masses that are used in our global fits. The final column gives the set of valence strange masses with which we computed the Omega baryon mass, and the kaon mass, decay constant and bag parameters. As discussed in the text, for the 260 MeV pion mass cut, we exclude kaonic data with valence light quark mass m_x if the pion with $m_y = m_x$ is excluded on that ensemble. Similarly, the Omega baryon and the Wilson flow data are excluded if the unitary pion on that ensemble is excluded.

which includes substantially more data, including a third lattice spacing, that may enable a more precise determination of the dominant a^2 scaling behaviour. In practice we find the results to be highly consistent.

Each of the fits with a 370 MeV pion mass cut have 49 free parameters and use 709 data points, giving 660 degrees of freedom; similarly, the analytic fits with the 260 MeV cut have 46 free parameters and use 414 data points, giving 368 degrees of freedom. Note that a substantial amount of the data on the 32ID, 32I and 24I ensembles differ only in their reweighted sea strange quark mass (for which we use four separate values including the simulated value) and are therefore highly correlated. The full set of input quark masses for the 32I, 24I and 32ID data that we include in the global fits for each of our two pion mass cuts are summarized in Tables XI and XII for convenience.

The guesses for the parameters in our global fits were input by hand based on a rough order-of-magnitude estimate obtained from previous fits, and within a reasonable basin of attraction we observed no deviations in the fit result (of course wildly different guesses can lead to false minima, but with much much larger χ^2).

Quantity	Measured value	Ansatz	$a=0$	m_{ud}^{phys}	m_s^{phys}
f_π(48I)	0.075799(84)	ChPTFV	-0.0037(73)	-0.00111(30)	0.00129(30)
		Ana.(370)	-0.0110(67)	-0.00175(20)	-0.00093(44)
		Ana.(260)	-0.0075(80)	-0.00201(24)	-0.00046(33)

$f_\pi(64\mathrm{I})$	0.055505(95)	ChPTFV	-0.0009(39)	-0.00083(41)	0.0001(10)
		Ana.(370)	-0.0059(37)	-0.00179(26)	-0.0039(11)
		Ana.(260)	-0.0040(43)	-0.00211(37)	-0.0020(12)
$f_K(48\mathrm{I})$	0.090396(86)	ChPTFV	-0.0024(58)	-0.00059(14)	-0.00095(68)
		Ana.(370)	-0.0059(54)	-0.00084(10)	-0.00174(73)
		Ana.(260)	-0.0055(62)	-0.00090(12)	-0.00173(75)
$f_K(64\mathrm{I})$	0.066534(99)	ChPTFV	-0.0009(31)	-0.00047(18)	-0.0061(13)
		Ana.(370)	-0.0032(29)	-0.00085(13)	-0.0074(13)
		Ana.(260)	-0.0029(33)	-0.00093(18)	-0.0073(17)
$f_K/f_\pi(48\mathrm{I})$	1.1926(14)	ChPTFV	0.0013(42)	0.00052(16)	-0.00223(49)
		Ana.(370)	0.0051(42)	0.00091(10)	-0.00082(35)
		Ana.(260)	0.0020(47)	0.00111(15)	-0.00127(57)
$f_K/f_\pi(64\mathrm{I})$	1.1987(18)	ChPTFV	0.0000(23)	0.00035(23)	-0.00625(89)
		Ana.(370)	0.0027(23)	0.00093(13)	-0.00346(68)
		Ana.(260)	0.0011(25)	0.00117(22)	-0.0053(13)
$t_0^{1/2}(48\mathrm{I})$	1.29659(39)	ChPTFV	-0.0276(62)	0.000122(20)	0.000204(95)
		Ana.(370)	-0.0260(56)	0.000120(20)	0.000176(84)
		Ana.(260)	-0.0259(68)	0.000140(22)	0.00023(10)
$t_0^{1/2}(64\mathrm{I})$	1.74448(98)	ChPTFV	-0.0150(33)	0.000122(24)	0.00088(24)
		Ana.(370)	-0.0142(30)	0.000124(23)	0.00076(21)
		Ana.(260)	-0.0141(37)	0.000148(32)	0.00097(24)
$w_0(48\mathrm{I})$	1.5013(10)	ChPTFV	0.0063(59)	0.000327(40)	0.00047(20)
		Ana.(370)	0.0080(54)	0.000328(41)	0.00043(19)
		Ana.(260)	0.0076(66)	0.000373(48)	0.00042(18)
$w_0(64\mathrm{I})$	2.0502(26)	ChPTFV	0.0034(32)	0.000322(50)	0.00199(41)
		Ana.(370)	0.0043(29)	0.000335(51)	0.00183(36)
		Ana.(260)	0.0041(36)	0.000388(73)	0.00179(41)

TABLE XIII: Data in lattice units on the 48I and 64I ensembles, along with the relative (fractional) correction to the infinite volume limit, in combination with each of the following: the continuum limit, the physical light quark mass and the physical strange mass. The corrections are shown for the ChPTFV fits, the analytic fit with a 260 MeV pion mass cut (labelled 'Ana.(260)'), and the analytic fit with a 370 MeV cut (labelled 'Ana. (370)'). We include the infinite-volume correction (where applicable) in all of these such that the ChPTFV corrections can be compared directly to those of the analytic fits, where the latter are performed to data that has first been corrected to the infinite volume.

The predicted values of the lattice spacings and (unrenormalized) physical quark masses obtained using the ChPTFV ansatz are listed in Table XV alongside the correlated (superjackknife) differences between those and the results for the other ansätze. A similar listing of the physical predictions can be found in Table XVI. The corresponding fit parameters for all four ansätze are given in Table XVII. For the analytic fit with the 260 MeV cut, the cut excludes the 32Ifine data for which the pion mass is $371(5)$ MeV, and we are therefore unable to directly obtain the scaling parameters associated with the heavy 32Ifine data; instead we first fit without these data and then determine the remaining unknowns, $Z_{l/h}^{32\mathrm{Ifine}}$ and $R_a^{32\mathrm{Ifine}}$, by including the 32Ifine data while freezing the other fit parameters to those obtained without these data.

In Figure 22 we plot the unitary mass dependence of m_π, m_K and m_Ω, which are used to determine the quark masses and overall lattice scale. In this figure we clearly see that the overweighting procedure forces the curve to pass through the near-physical data as desired, and that this procedure does not introduce any significant tension with the heavier data. In Figure 20 we plot a histogram of the deviation of the data from the ChPTFV fit, showing excellent general agreement between the fit and the data, and in Figure 21 we plot the corresponding histograms for the analytic fits. For the analytic fit with the 370 MeV mass cut we observe $\mathcal{O}(3-4)\sigma$ deviations of the 32ID pion mass data from the fit curve, which arise because of chiral curvature in the data: the fit is pinned near the physical point by the overweighting procedure and is strongly influenced by the larger volume of data in the heavy mass regime, leading to deviations from the lighter 32ID data that lies between these extremes. Nevertheless, in Tables XV, XVI and XVII we generally observe better agreement between the analytic fit with the 370 MeV mass cut and the ChPTFV results than for the lower cut. The total (uncorrelated) $\chi^2/d.o.f.$ are given in Table XIV and are sub-unity for all four ansätze.

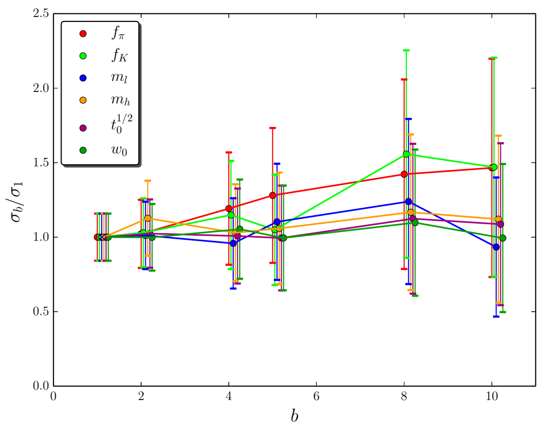


FIG. 19. The dependence of the error for the predicted physical values, obtained from our global fits with the ChPTFV ansatz, of various quantities as a function of the bin size used for the 64I ensemble. The vertical axis plots the ratio σ_b/σ_1 for bin size b along the horizontal axis, where σ is the statistical error and the subscript indicates the 64I bin size for which that error was computed. The upper and lower bounds were obtained by varying σ_b by $1/\sqrt{N}$, where N is the number of samples.

As previously mentioned, the inclusion of the Wilson flow data in these fits has a significant effect on the precision of the lattice spacings via their influence on the shared R_a parameters. This can be seen in Table XVIII, in which we show the various scaling parameters, as well as the unrenormalized quark masses and lattice spacings, obtained using the ChPTFV ansatz with and without the Wilson flow data. For the 48I and 64I ensembles, for which the hadronic measurements are very precise, we see only a small improvement in the statistical error. However, for the 32I, 24I and 32Ifine ensembles we observe factors of three or more improvements in precision. The results themselves are very consistent.

In Figure 19 we plot the dependence of our physical predictions on the bin size used for the 64I data. Here we observe no statistically significant dependence on the bin size, further attesting that our chosen bin size of 5 (5×40 MD time units) is a conservative choice and does not lead to an underestimate in the errors on our physical predictions.

We would like to emphasize that the goal of this analysis is not to extract reliable model parameters but simply to perform a few-percent extrapolation of our pristine near-physical data to the physical point. As we discuss in Section IV B, we are well aware that NLO ChPT can be expected to fail at the 5% level in the 200-370 MeV mass range in which the majority of our data lies (and where the fit would be most heavily weighted if we weighted the data by statistical error alone), and we do not want this model failure to unduly influence the quality of our prediction. The overweighting procedure was chosen to ensure that the fits pass through our 48I and 64I data with the heavier data used only to guide the extrapolation. Despite this, we find that the fits are largely insensitive to the pion mass cut and to the fit ansatz such that all of our results agree to a high degree (including their uncorrelated χ^2/d.o.f.). In order to gauge the quality of our uncorrelated fits, we present histograms of the deviation of the fit from our data in Figures 20, 21 (and 28 for B_K), and we see no spuriously large deviations that cannot be accounted for by higher-order mass dependent terms. Given the high degree of consistency between our results, there is no reason to suggest that any of the fits has converged upon a false minimum. Furthermore, the predictive power of these global fits is highlighted by our numerical discovery of the 3% shift in lattice spacings between the 48I and 24I ensembles and the smaller 1% shift between the 64I and 32I ensembles.

A. Systematic error estimation

In our previous analyses we used the difference between the ChPTFV and ChPT results as a conservative estimate of the higher-order finite-volume errors on our results (recall the ChPTFV formulae incorporate the NLO finite-volume corrections). From a purely χPT perspective this is a considerable over-estimate of the size of the NNLO and above corrections, which are known to be only a small fraction of the NLO values even at smaller volumes. Our prudence was motivated by Ref. [36], in which the authors observed significant deviations between the finite-volume corrections predicted by standard finite-volume chiral perturbation theory and those obtained via a resummed version of the Lüscher formula [37] that relates the finite-volume mass shift of a particle to the infinite-volume Euclidean scattering length of that particle with the pion. Nevertheless, one can conclude from those results that the full finite-volume corrections can be expected to differ from the NLO χPT predictions by only 30–50% for the light pions that we are currently using.

Our present fits are dominated by near-physical data computed on 5.5fm volumes, such that (e.g. in Tables XV and XVI) we observe only very tiny differences between the ChPT and ChPTFV fit results; these differences are typically 10–20% of the size of the statistical error, and hence have negligible impact upon the total error. Given the small size of these differences and that the true sizes of the higher-order finite-volume effects are expected to be several times smaller, we therefore choose to omit the finite-volume systematic from our error estimate.

The estimate of the chiral extrapolation error is made difficult due to the fact that the global fits combine the chiral and continuum extrapolations together, and in this analysis the latter are larger than the former while being less well determined by the fits (the a^2 parameters have typically 50-100% statistical error). As a result, the established procedure of estimating the chiral error from the difference of the ChPTFV and analytic result with a 260 MeV cut is no longer satisfactory.

In this analysis we considered analytic fits with both a 260 MeV and a 370 MeV pion mass cut. The latter is clearly applying the linear ansatz outside of its region of applicability, leading to deviations from the 32ID data at the 3-4σ level. Despite this there is generally excellent agreement between the continuum predictions of this fit and the ChPTFV. The analytic fit with the 260 MeV mass cut does not suffer from this issue, but at the expense of fitting to a considerably smaller amount of data, including one less lattice spacing. The ChPTFV fit on the other hand is theoretically 'clean' in that it is the correct ansatz for the data in the chiral limit, and agrees very well with our data when applied in the 140 to 370 MeV pion mass range. In Table XIII we see that all four ansätze agree at a broad level (given the size of the errors on the a^2 terms) as to the size of the continuum extrapolation, and this is by far the dominant correction. The only significant inconsistencies are in the light quark extrapolation, for which the 260 MeV analytic fit gives a larger correction indicating a stronger slope near the physical point. Nevertheless, the differences between the predicted corrections of the ChPTFV and 260 MeV analytic fits are at most on the 0.1% level.

Given the small size of the observed differences in the corrections to the 48I and 64I data, and our understanding that these are likely a result of deficiencies in the fitting strategies for those ansätze, we choose to take the cleaner ChPTFV ansatz, which describes our data very well, as our final result and treat the systematic error associated with the extrapolation to the physical point as negligible.

Finally, we consider the discretization systematic. For Wilson-style fermions the explicit symmetry breaking allows for a dimension-5 clover term of $\mathcal{O}(a\Lambda_{QCD})$; for domain wall fermions this term is heavily suppressed by the separation of the chiral modes in the fifth dimension, and can be discounted in practice [5]. Our domain wall simulations can be treated as non-perturbatively $\mathcal{O}(a)$ improved, and further chiral symmetry implies that all terms containing an odd power of the lattice spacing ($\mathcal{O}(a\Lambda_{\rm QCD})$, $\mathcal{O}(a^3\Lambda^3_{\rm QCD})$, etc) can be neglected; the leading discretization effects therefore enter at the $\mathcal{O}(a^4\Lambda^4_{\rm QCD})$ level, and these are of a comparable size [5] to logarithmic corrections to lattice artefacts that are regularly considered negligible. In our previous papers and above (cf. Table XIII) we observe that the discretization effects for the coarser 48I ensemble are at the 2%

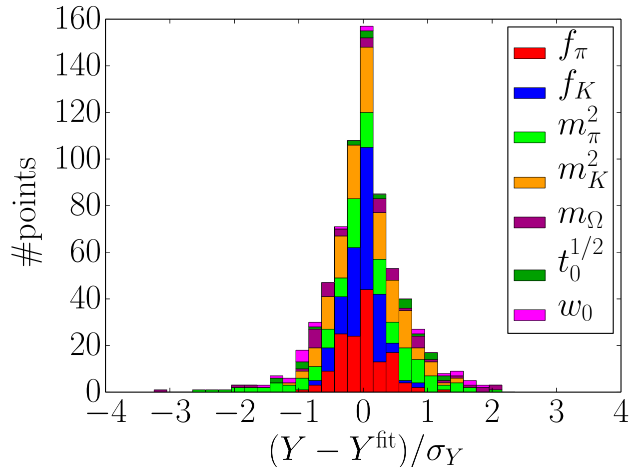


FIG. 20. A stacked (non-overlapping) histogram of the deviation of the ChPTFV fit curve from our data in units of the statistical error. Different coloured blocks are associated with the different quantities given in the legend. The 3σ outlier is the Ω mass on the heavier ($am_l = 0.005$) 24I ensemble at the un-reweighted strange mass of 0.04 in lattice units. The jackknife error on this point (not shown) is such that it is consistent with $(y - y_{\rm fit})/\sigma = -2$.

ChPTFV	ChPT	Analytic (260 MeV)	Analytic (370 MeV)
0.44(13)	0.44(16)	0.49(14)	0.79(18)

TABLE XIV. The χ^2/d.o.f. for each of the four chiral ansätze. Here the χ^2 does not include the over-weighted data, and the number of degrees of freedom has been correspondingly reduced. For the analytic fits, the pion mass cut is given in parentheses.

level, implying a $\mathcal{O}(0.04\%)$ discretization systematic that can be neglected. (For our very coarse 32ID ensemble the discretization effects enter at the 5% level, implying $\mathcal{O}(0.25\%)$ discretization errors that can also be discounted.) We could therefore, in principle, obtain precise continuum results from just two lattice spacings, as we have done in previous publications. However, the fits in this document utilize *three* widely spaced lattice spacings with the Shamir fermion action. In this document we present several plots overlaying our data with the fitted scaling behavior, from which we observe no evidence of deviations from a^2 scaling.

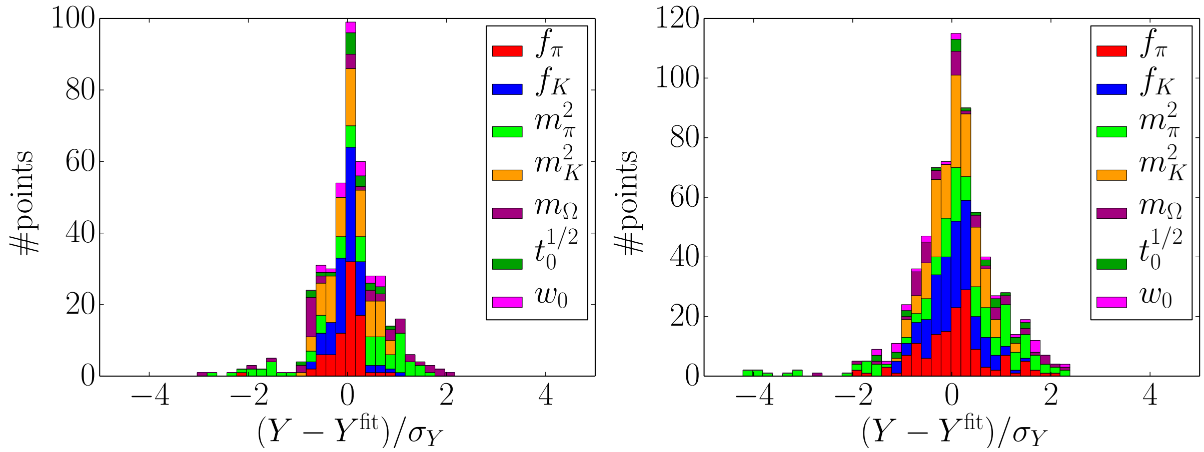


FIG. 21. A stacked (non-overlapping) histogram of the deviation of the analytic fit curves from our data in units of the statistical error. The left figure is for the 260 MeV pion mass cut, and the right plot for the 370 MeV cut. Different coloured blocks are associated with the different quantities given in the legend. The outliers in the right-hand plot are exclusively from m_π on the 32ID ensembles, indicating that the linear curve is deviating from the data due to chiral curvature.

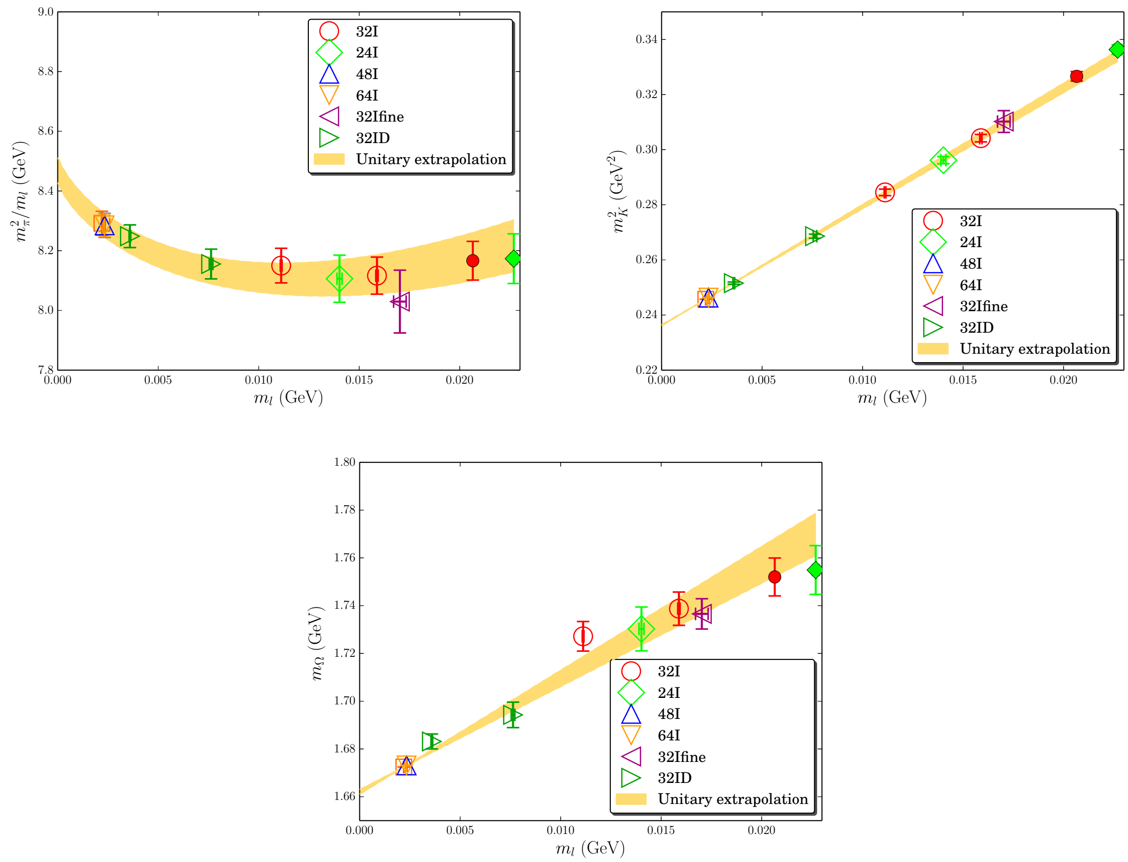


FIG. 22. m_π^2/m_l (upper-left), m_K^2 (upper-right) and m_Ω (lower) unitary data corrected to the physical strange quark mass and the infinite volume limit as a function of the unrenormalized physical quark mass, plotted against the ChPTFV fit curves. Data with hollow symbols are those included in the fit and data with filled symbols are those excluded. The square point is our predicted continuum value. Note the 64I and 48I data lie essentially on top of each other in this figure.

	ChPTFV	Δ(ChPT)	Δ (Analytic [260 MeV])	Δ (Analytic [370 MeV])
am_l(32I)	0.000260(13)	0.00000152(63)	−0.0000054(63)	−0.0000025(58)
am_s(32I)	0.02477(18)	0.000044(15)	0.000032(95)	0.000072(45)
a^{-1}(32I)	2.3833(86) GeV	−0.00234(74) GeV	−0.0001(51) GeV	−0.0043(25) GeV
am_l(64I)	0.0006203(77)	0.00000137(62)	−0.0000047(60)	−0.0000031(56)
am_s(64I)	0.02539(17)	0.000039(14)	0.000054(88)	0.000056(40)
a^{-1}(64I)	2.3586(70) GeV	−0.00181(67) GeV	−0.0021(40) GeV	−0.0027(19) GeV
am_l(24I)	−0.001770(79)	−0.00000048(35)	−0.0000037(21)	−0.0000012(20)
am_s(24I)	0.03224(18)	0.0000209(69)	−0.000054(50)	0.000046(18)
a^{-1}(24I)	1.7848(50) GeV	−0.00074(21) GeV	0.0032(22) GeV	−0.00194(65) GeV
am_l(48I)	0.0006979(81)	−0.00000049(35)	−0.0000020(18)	−0.0000016(19)
am_s(48I)	0.03580(16)	0.0000129(64)	0.000015(25)	0.000017(13)
a^{-1}(48I)	1.7295(38) GeV	−0.00029(16) GeV	−0.00027(59) GeV	−0.00042(33) GeV
am_l(32ID)	−0.000106(17)	−0.0000069(12)	−0.000002(13)	0.0000004(61)
am_s(32ID)	0.04625(48)	−0.000091(27)	−0.00018(28)	−0.00016(11)
a^{-1}(32ID)	1.3784(68) GeV	0.00141(37) GeV	0.0025(38) GeV	0.0020(17) GeV
am_l(32Ifine)	0.000058(16)	0.0000021(20)	0.000024(12)	0.0000040(57)
am_s(32Ifine)	0.01852(30)	0.000044(34)	−0.00019(26)	0.00005(10)
a^{-1}(32Ifine)	3.148(17) GeV	0.0003(14) GeV	0.0100(99) GeV	−0.0020(44) GeV

TABLE XV. The unrenormalized physical quark masses in bare lattice units (without $m_{\rm res}$ included) and the values of the inverse lattice spacing a^{-1} obtained using the ChPTFV ansatz, and the full correlated differences (labelled Δ) between the results obtained using the other ansätze and the ChPTFV result. We present analytic fit results obtained using both the 370 MeV and 260 MeV pion mass cut. The latter fit was performed without the 32Ifine data, and a separate fit with fixed parameters was used to obtain the 32Ifine scaling parameters.

	ChPTFV	Δ(ChPT)	Δ (Analytic [260 MeV])	Δ (Analytic [370 MeV])
f_π	0.1302(9) GeV	−0.000375(53) GeV	−0.00019(45) GeV	−0.00068(20) GeV
f_K	0.1555(8) GeV	−0.000251(52) GeV	−0.00035(43) GeV	−0.00043(17) GeV
f_K/f_π	1.1945(45)	0.00152(12)	−0.0010(21)	0.00297(60)
$t_0^{1/2}$	0.7292(41) GeV^{-1}	0.00098(37) GeV^{-1}	0.0014(23) GeV^{-1}	0.0014(11) GeV^{-1}
w_0	0.8742(46) GeV^{-1}	0.00114(42) GeV^{-1}	0.0013(27) GeV^{-1}	0.0016(12) GeV^{-1}

TABLE XVI. The physical predictions obtained using the ChPTFV ansatz, and the full correlated differences (labelled Δ) between the results obtained using the other ansätze and the ChPTFV result. We present analytic fit results obtained using both the 370 MeV and 260 MeV pion mass cut. The latter fit was performed without the 32Ifine data, and a separate fit with fixed parameters was used to obtain the 32Ifine scaling parameters.

Parameter	ChPT	ChPTFV	Parameter	Analytic (260 MeV)	Analytic (370 MeV)
Z_l^{24I}	0.9727(51)	0.9715(54)		0.9675(70)	0.9686(56)
Z_l^{48I}	0.9727(51)	0.9715(54)		0.9675(70)	0.9686(56)
Z_l^{32ID}	0.9192(67)	0.9156(72)		0.910(13)	0.9105(84)
$Z_l^{32Ifine}$	1.012(19)	1.015(17)		0.971(19)	1.005(15)
Z_h^{24I}	0.9634(38)	0.9628(40)		0.9637(43)	0.9636(36)
Z_h^{48I}	0.9634(38)	0.9628(40)		0.9637(43)	0.9636(36)
Z_h^{32ID}	0.9159(60)	0.9144(63)		0.9174(82)	0.9172(56)
$Z_h^{32Ifine}$	1.004(12)	1.005(12)		1.013(16)	1.005(12)
R_a^{24I}	0.7493(22)	0.7489(24)		0.7503(26)	0.7494(21)
R_a^{48I}	0.7263(27)	0.7257(28)		0.7256(29)	0.7268(25)
R_a^{64I}	0.9898(19)	0.9896(19)		0.9888(16)	0.9903(18)
R_a^{32ID}	0.5795(34)	0.5783(36)		0.5794(45)	0.5802(33)
$R_a^{32Ifine}$	1.3222(44)	1.3208(44)		1.3251(46)	1.3224(43)
B (GeV)	4.233(21)	4.236(21)	$C_0^{m_\pi}$ ([GeV]^2)	0.00037(15)	0.000421(91)
$L_8^{(2)}$	0.000611(41)	0.000631(41)	$C_1^{m_\pi}$ (GeV)	7.982(80)	7.917(51)
$L_6^{(2)}$	−0.000145(36)	−0.000146(36)	$C_2^{m_\pi}$ (GeV)	0.190(32)	0.219(25)
c_{m_π,m_h}	6.8(4.1)	3.7(4.1)	$C_3^{m_\pi}$ (GeV)	−0.036(31)	−0.026(32)
f (GeV)	0.12195(94)	0.12229(96)	$C_0^{f_\pi}$ (GeV)	0.1259(11)	0.12593(88)
c_f^I ([GeV]^2)	0.021(23)	0.017(23)	$C_a^{f_\pi,I}$ ([GeV]^2)	0.023(25)	0.034(21)
c_f^{ID} ([GeV]^2)	−0.027(30)	−0.033(30)	$C_a^{f_\pi,ID}$ ([GeV]^2)	−0.007(31)	0.013(29)
$L_5^{(2)}$	0.000524(78)	0.000513(78)	$C_1^{f_\pi}$	1.082(78)	0.988(45)
$L_4^{(2)}$	−0.000198(64)	−0.000171(64)	$C_2^{f_\pi}$	0.792(75)	0.643(71)
c_{f_π,m_h}	0.084(46)	0.070(46)	$C_3^{f_\pi}$	0.094(54)	0.188(46)
$m^{(K)}$ ([GeV]^2)	0.2363(16)	0.2363(17)	$C_0^{m_K}$ ([GeV]^2)	0.2363(19)	0.2363(15)
λ_2	0.02825(50)	0.02845(50)	$C_1^{m_K}$(GeV)	3.782(77)	3.828(43)
λ_1	0.00367(71)	0.00371(72)	$C_2^{m_K}$ (GeV)	0.54(16)	0.478(95)
c_{m_K,m_y} (GeV)	3.933(16)	3.935(17)	$C_3^{m_K}$ (GeV)	3.923(22)	3.929(15)
c_{m_K,m_h} (GeV)	0.097(86)	0.094(86)	$C_4^{m_K}$ (GeV)	0.11(15)	0.075(83)
$f^{(K)}$ (GeV)	0.15123(94)	0.15146(97)	$C_0^{f_K}$ (GeV)	0.1530(11)	0.15304(89)

$c^{I}_{f^{(K)}}$ ([GeV]^2)	0.012(18)	0.010(18)	$C_a^{f_K,I}$ ([GeV]^2)	0.017(19)	0.018(17)
$c^{ID}_{f^{(K)}}$ ([GeV]^2)	−0.020(27)	−0.024(27)	$C_a^{f_K,ID}$ ([GeV]^2)	−0.006(28)	−0.001(26)
λ_4	0.00620(38)	0.00594(39)	$C_1^{f_K}$	0.343(78)	0.361(34)
λ_3	−0.00383(79)	−0.00335(80)	$C_2^{f_K}$	0.653(86)	0.573(69)
c_{f_K,m_y}	0.2952(51)	0.2959(51)	$C_3^{f_K}$	0.3047(60)	0.2991(51)
c_{f_K,m_h}	0.074(45)	0.080(46)	$C_4^{f_K}$	0.113(62)	0.124(46)
$m^{(\Omega)}$ (GeV)	1.6618(30)	1.6620(33)	$C_0^{m_\Omega}$ (GeV)	1.6612(43)	1.6618(27)
c_{m_Ω,m_l}	4.86(42)	4.75(43)	$C_1^{m_\Omega}$	5.14(75)	4.89(44)
c_{m_Ω,m_y}	5.565(44)	5.583(46)	$C_2^{m_\Omega}$	5.582(63)	5.553(42)
c_{m_Ω,m_h}	1.39(45)	1.60(47)	$C_3^{m_\Omega}$	1.35(74)	1.27(47)
$c_{\sqrt{t_0},0}$ ([GeV]$^{-1}$)	0.7317(39)	0.7307(42)	$c_{\sqrt{t_0},0}$ ([GeV]$^{-1}$)	0.7323(49)	0.7320(37)
$c^{I}_{\sqrt{t_0},a}$ ([GeV]^2)	0.081(18)	0.085(19)	$c^{I}_{\sqrt{t_0},a}$ ([GeV]^2)	0.079(21)	0.080(18)
$c^{ID}_{\sqrt{t_0},a}$ ([GeV]^2)	0.037(13)	0.042(14)	$c^{ID}_{\sqrt{t_0},a}$ ([GeV]^2)	0.035(18)	0.035(13)
$c_{\sqrt{t_0},l}$ ([GeV]$^{-2}$)	−0.655(81)	−0.660(81)	$c_{\sqrt{t_0},l}$ ([GeV]$^{-2}$)	−0.747(84)	−0.640(80)
$c_{\sqrt{t_0},h}$ ([GeV]$^{-2}$)	−0.221(43)	−0.227(43)	$c_{\sqrt{t_0},h}$ ([GeV]$^{-2}$)	−0.262(38)	−0.205(44)
$c_{w_0,0}$ ([GeV]$^{-1}$)	0.8798(46)	0.8787(48)	$c_{w_0,0}$ ([GeV]$^{-1}$)	0.8805(58)	0.8803(43)
$c^{I}_{w_0,a}$ ([GeV]^2)	−0.022(16)	−0.019(17)	$c^{I}_{w_0,a}$ ([GeV]^2)	−0.022(19)	−0.024(16)
$c^{ID}_{w_0,a}$ ([GeV]^2)	0.018(12)	0.023(13)	$c^{ID}_{w_0,a}$ ([GeV]^2)	0.018(17)	0.016(12)
$c_{w_0,l}$ ([GeV]$^{-2}$)	−2.05(12)	−2.06(12)	$c_{w_0,l}$ ([GeV]$^{-2}$)	−2.30(14)	−2.03(12)
$c_{w_0,h}$ ([GeV]$^{-2}$)	−0.597(64)	−0.602(64)	$c_{w_0,h}$ ([GeV]$^{-2}$)	−0.567(67)	−0.580(65)

TABLE XVII: The fit parameters of each of our chiral ansätze. The parameters are given in physical units and with the heavy quark mass expansion point adjusted to the physical strange quark mass *a posteriori*. Analytic fit results are presented with a 370 MeV and 260 MeV pion mass cut. The latter was performed without the 32Ifine data, and a separate fit with fixed parameters was used to obtain the 32Ifine scaling parameters. For the ChPTFV and ChPT fits we use a chiral scale of 1.0 GeV. The fit formulae to which these parameters correspond can be found in Refs. [5, 6].

	With W.flow	Without W.flow
am_l(32I)	0.000260(13)	0.000262(15)
am_s(32I)	0.02477(18)	0.02483(27)
a^{-1}(32I)	2.3833(86) GeV	2.3726(181) GeV
Z_l(64I)	1.0(0)	1.0(0)
Z_h(64I)	1.0(0)	1.0(0)
R_a(64I)	0.9896(19)	0.9953(60)
am_l(64I)	0.0006203(77)	0.0006175(84)
am_s(64I)	0.02539(17)	0.02531(19)
a^{-1}(64I)	2.3586(70) GeV	2.3615(80) GeV
Z_l(24I)	0.9715(54)	0.9702(56)
Z_h(24I)	0.9628(40)	0.9612(43)
R_a(24I)	0.7489(24)	0.7494(42)
am_l(24I)	−0.001770(79)	−0.001767(78)
am_s(24I)	0.03224(18)	0.03236(32)
a^{-1}(24I)	1.7848(50) GeV	1.7779(132) GeV
Z_l(48I)	0.9715(54)	0.9702(56)
Z_h(48I)	0.9628(40)	0.9612(43)
R_a(48I)	0.7257(28)	0.7291(55)
am_l(48I)	0.0006979(81)	0.0006971(85)
am_s(48I)	0.03580(16)	0.03577(18)
a^{-1}(48I)	1.7295(38) GeV	1.7299(40) GeV
Z_l(32ID)	0.9156(72)	0.9122(79)
Z_h(32ID)	0.9144(63)	0.9107(70)
R_a(32ID)	0.5783(36)	0.5791(52)
am_l(32ID)	−0.000106(17)	−0.000099(18)
am_s(32ID)	0.04625(48)	0.04649(53)
a^{-1}(32ID)	1.3784(68) GeV	1.3741(75) GeV
Z_l(32Ifine)	1.015(17)	0.998(30)
Z_h(32Ifine)	1.005(12)	0.989(21)

R_a(32Ifine)	1.3208(44)	1.308(16)
am_l(32Ifine)	0.000058(16)	0.000078(30)
am_s(32Ifine)	0.01852(30)	0.01907(68)
a^{-1}(32Ifine)	3.148(17) GeV	3.104(45) GeV

TABLE XVIII: A comparison of the scaling parameters and the predictions for the lattice spacings and unrenormalized quark masses obtained by fitting using the ChPTFV ansatz with and without the Wilson flow data.

B. Physical predictions

In this section we present our predictions.

1. *χPT parameters*

The LO and NLO SU(2) partially-quenched χPT low-energy constants are given in Table XVII. These can be combined into the standard SU(2) χPT LECs, $\bar{l}_3$ and $\bar{l}_4$, giving

$$\bar{l}_3 = 2.73(13) \ \text{ and } \ \bar{l}_4 = 4.113(59)\,. \tag{67}$$

We can also compute the ratio of the decay constant to the LO SU(2) χPT parameter f, for which we obtain:

$$F_\pi/F = 1.0645(15)\,. \tag{68}$$

The errors on the above are statistical only; we make no attempt to estimate the systematic errors on these numbers due to higher-order effects or indeed the reliability of χPT in general. These issues will be investigated in a forthcoming publication.

2. *Lattice spacings*

For the lattice spacings we obtain the following values:

$$\begin{aligned}
a_{32\mathrm{I}}^{-1} &= 2.3833(86)\ \mathrm{GeV}\,, \\
a_{64\mathrm{I}}^{-1} &= 2.3586(70)\ \mathrm{GeV}\,, \\
a_{24\mathrm{I}}^{-1} &= 1.7848(50)\ \mathrm{GeV}\,, \\
a_{48\mathrm{I}}^{-1} &= 1.7295(38)\ \mathrm{GeV}\,, \\
a_{32\mathrm{Ifine}}^{-1} &= 3.148(17)\ \mathrm{GeV}\,, \\
a_{32\mathrm{ID}}^{-1} &= 1.3784(68)\ \mathrm{GeV}\,,
\end{aligned} \tag{69}$$

where we quote the statistical error in parentheses. Our previous values [5] for the lattice spacings of the 32I, 24I and 32ID ensembles are as follows:

$$\begin{aligned}
a_{32\mathrm{I}}^{-1} &= 2.310(37)(17)(9)\ \mathrm{GeV}\,, \\
a_{24\mathrm{I}}^{-1} &= 1.747(31)(24)(4)\ \mathrm{GeV}\,, \\
a_{32\mathrm{ID}}^{-1} &= 1.3709(84)(56)(3)\ \mathrm{GeV}\,,
\end{aligned} \tag{70}$$

where the errors are statistical, chiral and finite-volume. We observe a 1.8σ tension between the new and old values of the 32I lattice spacing, which appears to arise from the introduction of the physical point data; if we look at Figure 22 we see that the physical point data appears to favor a stronger light quark mass slope than one would obtain from the heavier data. Nevertheless there do not seem to be any clear discrepancies, except for those that might be attributed to statistical effects. Other than this, our new results are consistent with these values, and are significantly more precise due to the inclusion of the Wilson flow data.

3. *Decay constants*

In Table XVI we list the predicted values of f_π, f_K and f_K/f_π obtained using the ChPTFV ansatz, as well as the differences between those results and those of the other ansätze. As we now have data at several lattice spacings, we can examine the scaling of both f_π and f_K in order to ensure that their dependence on the lattice spacing can be described by a quadratic form. In Figure 23 we plot the data, corrected to the physical quark masses and the infinite volume using the ChPTFV fit, as a function of the lattice spacing. In addition we show the scaling curve for the Iwasaki ensembles. We observe excellent consistency between the data and the fit curve for both quantities.

In Figure 24 we show the chiral extrapolation in the continuum/infinite-volume limits with the ChPTFV ansatz, again showing excellent agreement between the data and the fit.
We obtain the following physical predictions:

$$\begin{aligned} f_\pi &= 0.13019(89)\ \text{GeV}\,, \\ f_K &= 0.15551(83)\ \text{GeV}\,, \\ f_K/f_\pi &= 1.1945(45)\,, \end{aligned} \tag{71}$$

where, as above, the statistical errors are given in parentheses. Previously [5] we obtained

$$\begin{aligned} f_\pi &= 0.1271(27)(9)(25)\ \text{GeV}\,, \\ f_K &= 0.1524(30)(7)(15)\ \text{GeV}\,, \\ f_K/f_\pi &= 1.1991(116)(69)(116)\,. \end{aligned} \tag{72}$$

Here we see that the inclusion of the 48I and 64I data, giving statistically precise data at simulated masses very near the physical quark masses, has led to a highly significant improvement in our results.
In our first global fit analysis [6], performed only to the 32I and 24I ensembles over a (unitary) pion mass range of 290–420 MeV, we obtained a value for f_π from our NLO χPT fit that was 6.6% (9 MeV) lower than the experimental value. We concluded that this discrepancy was due to systematic errors in the chiral extrapolation, and introduced the analytic fits as a means of estimating this systematic. When we included the 32ID ensembles into the global fit [5] we observed a marked improvement in the results for the decay constants and a corresponding reduction in the size of the chiral systematic (as estimated by taking the difference between the ChPTFV and analytic fit results).
Now, with the inclusion of the 48I and 64I data we have essentially eliminated the chiral extrapolation error, and have obtained values for both decay constants that are in excellent agreement with the Particle Data Group (PDG) values [38], $f_{\pi^-} = 0.1304(2)$ GeV and $f_{K^-} = 0.1562(7)$ GeV. Here, f_{π^-} is determined experimentally using the measured branching fraction and pion lifetime, with $|V_{ud}|$ computed very precisely via nuclear β decay, such that the error is dominated by higher order terms in the decay width formula. On the other hand, the value for f_{K^-} requires $|V_{us}|$ as input, which, for the quoted result, is computed using $|V_{us}|f_+(0)$ determined via semileptonic kaon decays and lattice input for $f_+(0)$. The consistency of our f_K with the PDG value could therefore be taken as both representing the consistency of experiment with the Standard Model, and the quality of the lattice QCD determinations of both the kaon semileptonic form factor and our

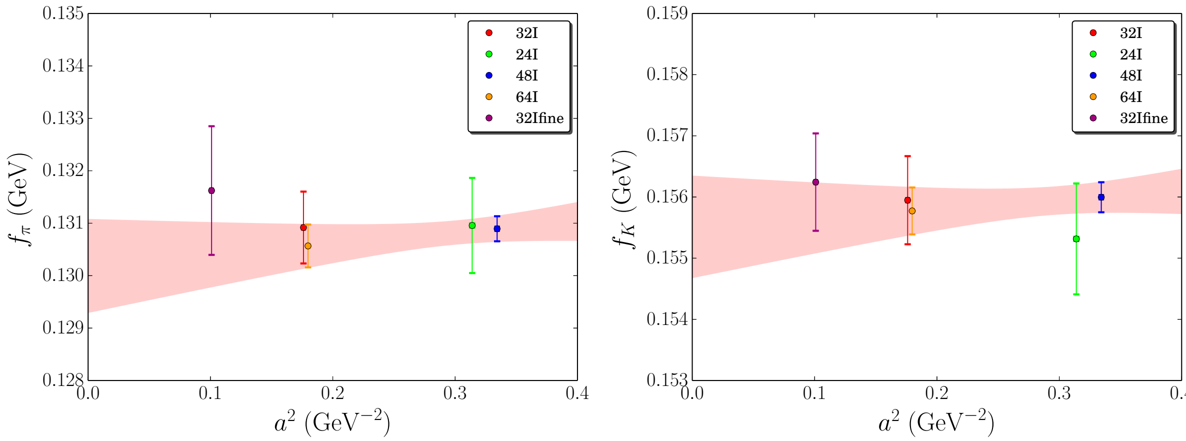


FIG. 23. f_π (left) and f_K (right) data corrected to the physical up/down and strange quark masses and the infinite-volume as a function of the square of the lattice spacing. The curve shows the continuum extrapolation for the Iwasaki action with the ChPTFV ansatz. Here we have not shown the 32ID data point as it has a different gauge action.

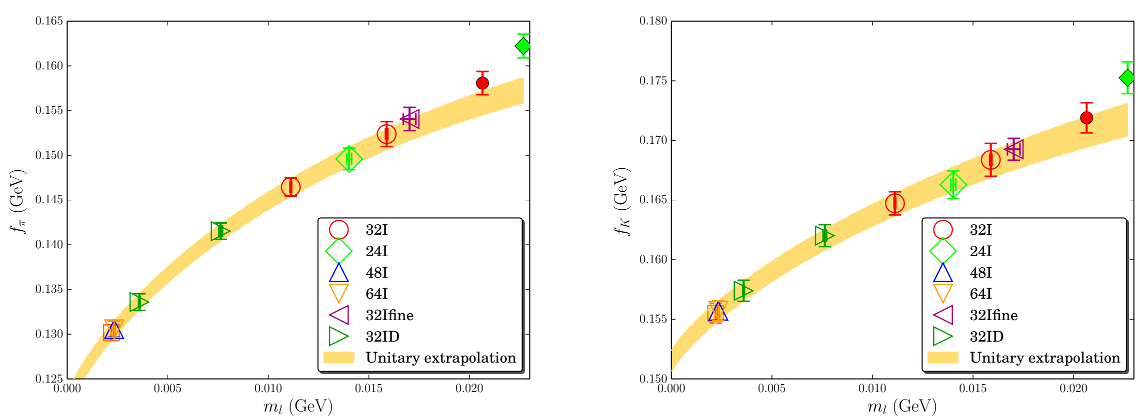
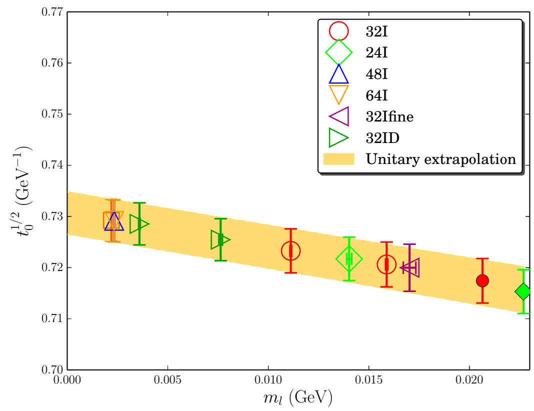


FIG. 24. f_π (left) and f_K (right) unitary data corrected to the physical strange quark mass and the continuum and infinite-volume limits as a function of the unrenormalized physical quark mass, plotted against the ChPTFV fit curves. Data with hollow symbols are those included in the fit and data with filled symbols are those excluded. The square point is our predicted continuum value. Note the 64I and 48I data lie essentially on top of each other in this figure.

determination of the kaon decay constant.



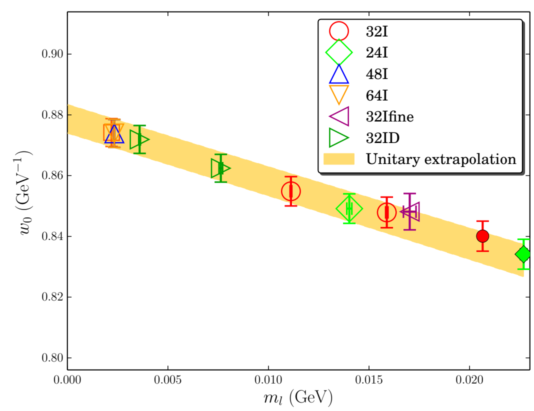


FIG. 25. $t_0^{1/2}$ (left) and w_0 (right) unitary data corrected to the physical strange sea quark mass and the continuum limit as a function of the unrenormalized physical quark mass, plotted against the ChPTFV fit curves. Data with hollow symbols are those included in the fit and data with filled symbols are those excluded. The square point is our predicted continuum value. Note the 64I and 48I data lie essentially on top of each other in this figure.

4. *Wilson flow scales*

In Table XVI we list the predicted values of the Wilson flow scales, $t_0^{1/2}$ and w_0, in the continuum limit. The unitary mass dependencies are plotted in Figure 25 and the a^2 dependencies in Figure 26. For our final results, we obtain the following continuum predictions:

$$\begin{aligned} t_0^{1/2} &= 0.7292(41)\ \mathrm{GeV}^{-1}\,, \\ w_0 &= 0.8742(46)\ \mathrm{GeV}^{-1}\,, \end{aligned} \tag{73}$$

where the statistical error is quoted in parentheses.

The above values can be compared to the following results obtained using 2+1f 2HEX-smeared Wilson fermions [33]:

$$\begin{aligned} t_0^{1/2} &= 0.1465(25)\ \mathrm{fm} &= 0.7425(127)\ \mathrm{GeV}^{-1}\,, \\ w_0 &= 0.1755(18)\ \mathrm{fm} &= 0.8894(91)\ \mathrm{GeV}^{-1}\,, \end{aligned} \tag{74}$$

where we have combined the statistical and systematic errors in quadrature. We find excellent agreement between these and our results.

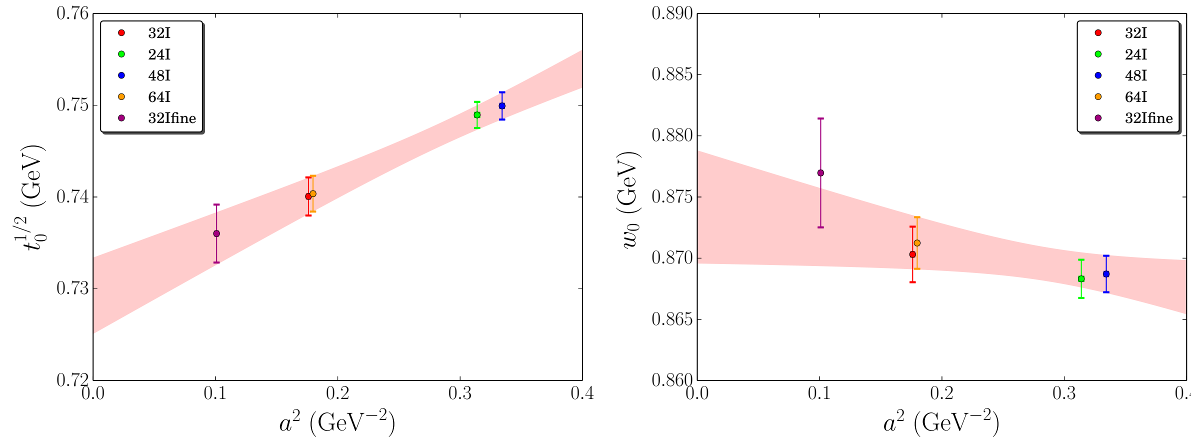


FIG. 26. $t_0^{1/2}$ (left) and w_0 (right) data corrected to the physical up/down and strange sea quark masses as a function of the square of the lattice spacing. The curve shows the continuum extrapolation for the Iwasaki action with the ChPTFV ansatz. Here we have not shown the 32ID data point as it has a different gauge action.

5. Unrenormalized physical quark masses

The quark masses in bare lattice units on the 32I reference ensemble are given in Table XV. In physical units, and including the residual mass, the unrenormalized physical quark masses are given in Table XIX. Combining these results we obtain the following:

$$\begin{aligned} m_{ud}^{\text{unrenorm.}} &= 2.198(11)\text{ MeV}\,, \\ m_s^{\text{unrenorm.}} &= 60.62(24)\text{ MeV}\,, \end{aligned} \tag{75}$$

where the errors are statistical.

Ansatz	$m_{ud}^{\text{unrenorm.}}$ (MeV)	m_s^{unrenorm} (MeV)
ChPTFV	2.198(11)	60.62(24)
ChPT	2.199(10)	60.67(22)
analytic (260 MeV)	2.185(16)	60.70(27)
analytic (370 MeV)	2.188(13)	60.69(21)

TABLE XIX. Unrenormalized physical quark masses. For the analytic fits, the corresponding pion mass cut is given in parentheses.

C. Renormalized physical quark masses and the chiral condensate

The quark masses presented above are defined in the bare lattice normalization of the 32I reference ensemble. On each of the 32I and 24I ensembles independently, we calculate the non-perturbative renormalization factors that are necessary to convert quark masses in the corresponding bare normalization into a variant of the Rome-Southampton RI-MOM scheme [39] that can be related to $\overline{\text{MS}}$ via perturbation theory. The procedure applied below is identical to that used in Refs. [6] and [5], and the determination of the renormalization coefficients is documented in Appendix F; below we provide only a brief outline.

We compute amputated, projected bilinear vertex functions,

$$\Lambda_{\mathscr{O}}(q^2) = \text{tr}\left[\Pi_{\mathscr{O}}(q^2)\Gamma^{(s)}_{\mathscr{O}}\right] , \tag{76}$$

where $\mathscr{O}$ is an operator, Π are the matrix-valued amputated vertex functions and $\Gamma^{(s)}$ are projection operators, for which the superscript s indexes the particular renormalization scheme (where applicable). We use the 'symmetric' RI-MOM schemes, defined by the following condition on the incoming and outgoing quark momenta, p_{in} and p_{out} respectively: $p_{\text{in}}^2 = p_{\text{out}}^2 = q^2 \equiv (p_{\text{in}} - p_{\text{out}})^2$. We define renormalization factors by matching to the tree-level amplitude at the scale $\mu^2 = q^2$:

$$\frac{Z_O}{Z_q^{n/2}}(\mu,a) \times \Lambda_O^{bare}(\mu,a) = \Lambda_O^{tree} . \tag{77}$$

In order to cancel the factors of the quark field renormalization in the denominator, we use

$$Z_m^{(s)}(\mu,a) = \frac{\bar{\Lambda}_S(\mu,a)}{Z_V \times \bar{\Lambda}_V^{(s)}(\mu,a)} , \tag{78}$$

where $\bar{\Lambda}_{\mathscr{O}} \equiv \Lambda_{\mathscr{O}}^{bare} \times (\Lambda_{\mathscr{O}}^{tree})^{-1}$, S and V are the scalar and vector operators repectively, and Z_V is the vector-current renormalization computed using hadronic variables via the procedure given in

Section III C 2. We use two different choices of projection operator for the vector vertex, formed from the quantities $\not{q}q^\mu/q^2$ and γ^μ; these define the SMOM and SMOM$_{\gamma^\mu}$ schemes respectively. More details on the projection operators and the numerical determination of these quantities can be found in Appendix F.

We now describe the procedure by which we obtain the renormalized quark masses given the renormalization factors. In Section V B 5 we present quark masses normalized according to the bare lattice units of the 32I reference ensemble. For any other ensemble e, the quark masses in the associated bare normalization can be obtained simply by dividing the values of m_{ud} and m_s given in Eq. (75) by Z_l^e and Z_h^e respectively. For each ensemble, the masses renormalized in the RI-SMOM schemes can therefore be computed as

$$(m_f^{\mathrm{SMOM}*})^e = (Z_m^{\mathrm{SMOM}*})^e m_f^{\mathrm{unrenorm.}}/Z_f^e\,, \tag{79}$$

where $f \in \{l,h\}$. These measurements contain finite lattice spacing errors associated with the vertex functions used in the conversion to $\overline{\mathrm{MS}}$. In order to convert our continuum quark masses to the RI-SMOM scheme, and thence to $\overline{\mathrm{MS}}$, we linearly extrapolate the ratio

$$Z_{mf}^e = Z_m^e/Z_f^e \tag{80}$$

in a^2 to the continuum. This extrapolation is performed using only two lattice spacings, potentially introducing additional systematic effects. In practice we find that the linear continuum fit results in a 4% shift in the central values from those computed on our finest ensemble (32I). The good chiral symmetry of the action heavily suppresses $\mathcal{O}(a^3)$ terms in the Symanzik effective theory and higher order corrections enter only at the $\mathcal{O}(a^4)$ level. This suggests systematic effects on the order of $(4\%)^2 \sim 0.16\%$, which we treat as negligible. Applied to the quark masses, the products

$$m_f^{\mathrm{SMOM}*} = (Z_{mf}^{\mathrm{SMOM}*})^{contm.} m_f^{\mathrm{unrenorm.}}\,, \tag{81}$$

are then free from $\mathcal{O}(a^2)$ scaling errors and have negligible higher order discretization systematics. Fixing the renormalization coefficients to a particular scale requires the input of the lattice spacings from the main analysis in order to convert the lattice momenta to physical units; for this we used only the central values of the ChPTFV fits. In order to account for the effect of the statistical and systematic uncertainties on the lattice spacings, we repeated the determination of the renormalization coefficients using two different values of the lattice spacings that differed slightly in value, and from these we estimated the slope of the renormalization coefficients with respect to the input lattice spacing. For each chiral ansatz, we then used the slope to shift the central values of the renormalization coefficients to the lattice spacings determined via that ansatz, and also to inflate the statistical errors of the superjackknife distribution to reflect the uncertainty on those values. The continuum extrapolations of Z_{ml} and Z_{mh} were performed independently for each ansatz, enabling us to determine the full effect of the systematic errors in the final step. The values of Z_m and Z_{mf} thus determined are given in Table XX.

Applying the renormalization factors to the masses from the previous section, we obtain the values given in Table XXI. Converting to the $\overline{\text{MS}}$ scheme and including the additional systematic errors associated with the perturbative matching, we find

$$\begin{aligned} m_{ud}(\overline{\text{MS}}, 3.0\,\text{GeV}) &= 2.997(36)(33)\ \text{MeV}\,, \\ m_s(\overline{\text{MS}}, 3.0\,\text{GeV}) &= 81.64(77)(88)\ \text{MeV}\,. \end{aligned} \tag{82}$$

where the errors are statistical and from the perturbative truncation respectively. In the RGI scheme, these correspond to

$$\begin{aligned} \hat{m}_{ud} &= 8.62(10)(9)\ \text{MeV}\,, \\ \hat{m}_s &= 235.0(22)(25)\ \text{MeV}\,. \end{aligned} \tag{83}$$

The quark mass ratio is

$$m_s/m_{ud} = 27.34(21)\,, \tag{84}$$

for which there is no systematic error associated with the perturbative matching as it cancels in the ratio.

For comparison, in our previous work [5] we obtained

$$\begin{aligned} m_{ud}(\overline{\text{MS}}, 3.0\,\text{GeV}) &= 3.05(8)(6)(1)(2)\ \text{MeV}\,, \\ m_s(\overline{\text{MS}}, 3.0\,\text{GeV}) &= 83.5(1.7)(0.8)(0.4)(0.7)\ \text{MeV}\,. \end{aligned} \tag{85}$$

and

$$m_s/m_{ud} = 27.36(39)(31)(22)\,, \tag{86}$$

for which the errors are statistical, chiral and finite-volume. Our new results are highly consistent with these values and again show a substantial improvement in the systematic error as a result of including the near-physical data.

We can also compute the chiral condensate,

$$\Sigma = -\langle \bar{u}u \rangle_{m_u, m_d \to 0} = BF^2 = Bf^2/2\,, \tag{87}$$

by combining the leading-order SU(2) χPT parameters from Table XVII. Like the quark masses, this quantity must be renormalized. Again we first convert to our intermediate SMOM schemes and subsequently perturbatively convert each to $\overline{\mathrm{MS}}$, using the difference as an estimate of the perturbative truncation systematic. The appropriate renormalization factor can be determined by noting that the leading-order χPT formula for the pion mass must be renormalization-scheme independent:

$$(m_\pi^2)_{\mathrm{LO}} = 2B^{\mathrm{unrenorm}} m_{ud}^{\mathrm{unrenorm}} = 2B^{\mathrm{SMOM}*} m_{ud}^{\mathrm{SMOM}*} = 2B^{\mathrm{SMOM}*} (Z_{ml}^{\mathrm{SMOM}*})^{contm.} m_{ud}^{\mathrm{unrenorm.}} . \quad (88)$$

This suggests that

$$B^{\mathrm{SMOM}*} = B^{\mathrm{unrenorm.}} / (Z_{ml}^{\mathrm{SMOM}*})^{contm.} . \quad (89)$$

The subsequent conversion to the $\overline{\mathrm{MS}}$ scheme at 3 GeV can be performed by further dividing by the appropriate scheme change factor.

It is customary to quote the dimension-one quantity $(\Sigma)^{1/3}$. We obtain

$$\begin{aligned} \Sigma^{1/3}(\mathrm{SMOM}, 3.0\ \mathrm{GeV}) &= 0.2837(19)\ \mathrm{GeV} \\ \Sigma^{1/3}(\mathrm{SMOM}_{\gamma^\mu}, 3.0\ \mathrm{GeV}) &= 0.2791(19)\ \mathrm{GeV} , \end{aligned} \quad (90)$$

which, after converting to $\overline{\mathrm{MS}}$ and combining, gives

$$\Sigma^{1/3}(\overline{\mathrm{MS}}, 3.0\ \mathrm{GeV}) = 0.2853(20)(10)\ \mathrm{GeV} , \quad (91)$$

where the errors are statistical and from the perturbative matching respectively.

Scheme	Lattice	Ansatz	Z_m	Z_{ml}	Z_{mh}
SMOM	24I	ChPTFV	1.4386(12)	1.4808(82)	1.4942(63)
SMOM	24I	ChPT	1.4385(12)	1.4788(79)	1.4932(60)
SMOM	24I	analytic (260 MeV)	1.4390(12)	1.4874(108)	1.4931(68)
SMOM	24I	analytic (370 MeV)	1.4383(12)	1.4849(86)	1.4927(57)
SMOM	32I	ChPTFV	1.4396(37)	1.4396(37)	1.4396(37)
SMOM	32I	ChPT	1.4393(37)	1.4393(37)	1.4393(37)
SMOM	32I	analytic (260 MeV)	1.4396(37)	1.4396(37)	1.4396(37)
SMOM	32I	analytic (370 MeV)	1.4391(37)	1.4391(37)	1.4391(37)
SMOM	cont.	ChPTFV	-	1.3870(122)	1.3699(100)
SMOM	cont.	ChPT	-	1.3888(120)	1.3704(100)
SMOM	cont.	analytic (260 MeV)	-	1.3780(145)	1.3706(103)
SMOM	cont.	analytic (370 MeV)	-	1.3805(128)	1.3705(99)
$\mathrm{SMOM}_{\gamma^\mu}$	24I	ChPTFV	1.5235(13)	1.5682(87)	1.5824(67)
$\mathrm{SMOM}_{\gamma^\mu}$	24I	ChPT	1.5234(13)	1.5661(83)	1.5813(64)
$\mathrm{SMOM}_{\gamma^\mu}$	24I	analytic (260 MeV)	1.5240(13)	1.5752(115)	1.5813(72)
$\mathrm{SMOM}_{\gamma^\mu}$	24I	analytic (370 MeV)	1.5232(13)	1.5725(91)	1.5808(60)
$\mathrm{SMOM}_{\gamma^\mu}$	32I	ChPTFV	1.5192(39)	1.5192(39)	1.5192(39)
$\mathrm{SMOM}_{\gamma^\mu}$	32I	ChPT	1.5189(39)	1.5189(39)	1.5189(39)
$\mathrm{SMOM}_{\gamma^\mu}$	32I	analytic (260 MeV)	1.5192(39)	1.5192(39)	1.5192(39)
$\mathrm{SMOM}_{\gamma^\mu}$	32I	analytic (370 MeV)	1.5186(39)	1.5186(39)	1.5186(39)
$\mathrm{SMOM}_{\gamma^\mu}$	cont.	ChPTFV	-	1.4567(126)	1.4386(103)
$\mathrm{SMOM}_{\gamma^\mu}$	cont.	ChPT	-	1.4585(125)	1.4389(103)
$\mathrm{SMOM}_{\gamma^\mu}$	cont.	analytic (260 MeV)	-	1.4470(150)	1.4392(106)
$\mathrm{SMOM}_{\gamma^\mu}$	cont.	analytic (370 MeV)	-	1.4496(134)	1.4390(102)

TABLE XX. The non-perturbative renormalization factors calculated at $\mu = 3.0$ GeV that are used to convert bare quark masses (Z_m) and quark masses in the normalization of the 32I reference ensemble (Z_{ml},Z_{mh}). Values are given on the 32I and 24I ensembles and in the continuum limit for the latter quantity.

Scheme	Ansatz	$m_{u/d}$ (GeV)	m_s (GeV)
SMOM	ChPTFV	0.003049(37)	0.08305(80)
SMOM	ChPT	0.003055(36)	0.08314(79)
SMOM	analytic (260 MeV)	0.003011(50)	0.08319(87)
SMOM	analytic (370 MeV)	0.003021(42)	0.08317(76)
SMOM$_{\gamma^\mu}$	ChPTFV	0.003202(38)	0.08721(83)
SMOM$_{\gamma^\mu}$	ChPT	0.003208(37)	0.08730(81)
SMOM$_{\gamma^\mu}$	analytic (260 MeV)	0.003162(51)	0.08735(89)
SMOM$_{\gamma^\mu}$	analytic (370 MeV)	0.003172(43)	0.08733(79)

TABLE XXI. The physical quark masses renormalized at $\mu = 3.0$ GeV in the two intermediate RI-SMOM schemes for each of the chiral ansätze. The quoted errors are statistical only.

D. Neutral kaon mixing parameter, B_K

The neutral kaon mixing parameter is renormalization scheme dependent, and as such the fits must be performed using renormalized data. As this introduces additional systematic errors, we follow our established procedure of performing these fits separately from the main global fit analysis. Below we first summarize our non-perturbative renormalization procedure for B_K and then present the results of the chiral/continuum fit and finally our physical predictions.

1. *Renormalization of B_K*

In this section we provide a brief outline of the procedure for determining the renormalization coefficients; for more details we refer the reader to Appendix F and Refs. [40] and [5].

As with the quark mass renormalization, we make use of ‘symmetric’ regularization-invariant momentum schemes (RI-SMOM for short), defined by the condition $\mu^2 = p_1^2 = p_2^2 = q^2 \equiv (p_1 - p_2)^2$, where p_1 and p_2 are the momenta of the incoming and outgoing quarks: $d(p_1)\bar{s}(-p_2) \to \bar{d}(-p_1)s(p_2)$. We compute the amputated and projected Green’s function of the relevant four-quark operator, $\mathscr{O}_{LL}$, describing the $K-\bar{K}$ mixing, normalized by the square of the average between the vector and axial bilinear:

$$Z_{B_K}^{(s_1,s_2)}(\mu,a) \times \frac{\bar{\Lambda}_{VV+AA}^{(s_1)}(\mu,a)}{\bar{\Lambda}_{AV}^{(s_2)}(\mu,a)^2} = 1\,, \tag{92}$$

where

$$\Lambda_{AV} = \frac{1}{2}(\Lambda_V + \Lambda_A), \tag{93}$$

and $\bar{\Lambda}_{\mathscr{O}} \equiv \Lambda_{\mathscr{O}}^{bare} \times (\Lambda_{\mathscr{O}}^{tree})^{-1}$ for the operator $\mathscr{O}$, as before.

Note that the quark wave function renormalization factor cancels in the ratio. In Appendix F we show that the difference between Λ_V and Λ_A at 3 GeV is numerically negligible, and therefore the above choice of normalization is irrelevant. The superscript (s_i) refers to choice of projector (cf. [5]): either γ_μ or $\not{q}$. The choices $s_1 = s_2 = \gamma^\mu$ and $s_1 = s_2 = \not{q}$ define the so-called SMOM(γ^μ, γ^μ) and SMOM$(\not{q}, \not{q})$ schemes respectively.

We perform the full analysis separately for each scheme and use the difference to estimate the systematic error associated with the $\overline{\text{MS}}$ matching. While treating the two schemes in an equal fashion is the most rigorous estimate we can make with the current data, we have indications that this might overestimate the error on the SMOM$(\not{q}, \not{q})$ result: A preliminary study [41] of step scaling to higher momentum scales suggests that the scale evolution in this scheme agrees with perturbation theory over the full range of scales, whereas the SMOM(γ^μ, γ^μ) scheme evolves into better agreement as the scale is raised. The perturbative truncation error is therefore greater for the SMOM(γ^μ, γ^μ) scheme than for the SMOM$(\not{q}, \not{q})$ scheme. The complete study of the evolution to higher energy scales requires careful treatment of the charm threshold, and is the subject of further work by RBC and UKQCD. These observations are consistent with our earlier results at lower scales, and the better agreement with the perturbative scale evolution for the SMOM$(\not{q}, \not{q})$ scheme was the reason we have, in this work and previously, taken our central values for B_K from this scheme [40].

We compute Z_{B_K} on each ensemble at a number of q^2, and interpolate to a chosen high momentum scale at which the matching to $\overline{\text{MS}}$ can be performed. We choose to perform the matching at 3.0 GeV as before. The values of the renormalization coefficients at the various lattice momenta and further details of the analysis are given in Appendix F.

All matrix elements included in the global fit must be renormalized to a common scale of 3.0 GeV in order that the global fit can extrapolate these to a shared, universal continuum limit. As described in Ref. [5], due to the coarseness of the 32ID ensemble we are unable to renormalize directly at 3 GeV without introducing potentially sizeable lattice artifacts. Instead we renormalize with a lower momentum scale of $\mu_0 = 1.4363$ GeV, and apply the continuum non-perturbative running $\sigma_{B_K}^{(s_1,s_2)}(\mu, \mu_0)$, extracted from the 32I and 24I lattices (and extrapolated to the continuum), to convert this value to $\mu = 3$GeV. More details of this conversion are given in Appendix F.

Scheme	Lattice	ChPTFV	ChPT	Analytic (260 MeV)	Analytic (370 MeV)
$(\not{q},\not{q})$	32I	0.9787(3)	0.9787(3)	0.9787(3)	0.9786(3)
	24I	0.9568(3)	0.9568(3)	0.9570(3)	0.9568(3)
	48I	0.9545(1)	0.9544(1)	0.9544(1)	0.9544(1)
	64I	0.9782(2)	0.9781(2)	0.9781(2)	0.9781(2)
	32Ifine	0.9995(4)	0.9995(4)	0.9998(5)	0.9995(4)
	32ID	0.9284(45)	0.9286(45)	0.9276(45)	0.9289(45)
(γ^μ,γ^μ)	32I	0.9409(2)	0.9408(2)	0.9409(2)	0.9408(2)
	24I	0.9161(5)	0.9161(5)	0.9162(5)	0.9160(5)
	48I	0.9140(1)	0.9140(1)	0.9140(1)	0.9140(1)
	64I	0.9411(1)	0.9410(1)	0.9410(1)	0.9410(1)
	32Ifine	0.9617(3)	0.9617(2)	0.9619(3)	0.9617(2)
	32ID	0.8824(25)	0.8824(25)	0.8824(26)	0.8824(25)

TABLE XXII. Z_{B_K} at 3 GeV in the two intermediate schemes, with the central values shifted and errors inflated to account for the different values of the lattice spacings obtained via each chiral ansatz.

Determining the lattice momentum corresponding to the 3 GeV match point requires the input of the lattice spacings determined in the previous sections. The effects of the uncertainties on the lattice spacings are incorporated by shifting the central values and inflating the errors according to the lattice spacings determined via each of the chiral ansätze, using the procedure outlined in the Section V C. The resulting values of Z_{B_K} are given in Table XXII.

2. *Chiral/continuum fit to B_K*

As above, we describe the chiral dependence using chiral perturbation theory, with and without finite-volume corrections, as well as a linear ansatz with a 260 MeV and 370 MeV pion mass cut. The chiral/continuum fit forms can be found in Ref. [40]. As before, we use separate parameters to describe the lattice spacing dependence of the Iwasaki and Iwasaki+DSDR actions. The fit parameters can be found in Table XXIV, and in Figure 27 we show examples of the unitary and continuum extrapolations. In Figure 28, in which we plot a histogram of the statistical deviations of the data from the ChPTFV fit curve, we see excellent consistency between the data and the fit.

Scheme	ChPTFV	ChPT	Analytic (260 MeV)	Analytic (370 MeV)
$(\not{q},\not{q})$	0.55(38)	0.70(42)	0.46(35)	0.51(34)
(γ^μ,γ^μ)	0.62(43)	0.78(46)	0.52(40)	0.58(39)

TABLE XXIII. The χ^2/d.o.f. for each of the four chiral ansätze and the two intermediate renormalization schemes. Here the χ^2 does not include the overweighted data, and the number of degrees of freedom has been correspondingly reduced. For the analytic fits, the pion mass cut is given in parentheses.

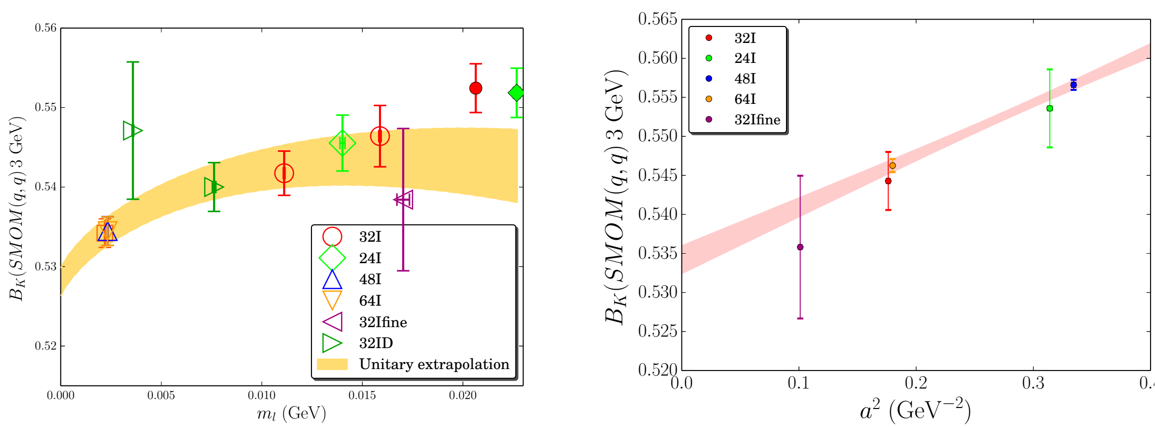


FIG. 27. The left figure shows the unitary light quark mass dependence of B_K in the SMOM$(\not{q},\not{q})$ at 3 GeV. The quark masses are in physical units and in the native normalization of the 32I reference ensemble. Data with hollow symbols are those included in the fit and data with filled symbols are those excluded. The right figure shows the lattice spacing dependence of those data. Here we have not included the 32ID ensemble as it lies on a different scaling trajectory.

The total χ^2/d.o.f. for each of the four ansätze are given in Table XXIII.

The fits to B_K with a 370 MeV pion mass cut have 7 free parameters (the remainder having been determined in our earlier fits, above) and use 163 data points, giving 156 degrees of freedom; for the 260 MeV cut have 7 parameters and 90 data points, giving 83 degrees of freedom.

3. *Predicted values*

In Table XXV we list the continuum predictions for B_K, renormalized in each of the two intermediate schemes, that we obtained using the ChPTFV ansatz, as well as the sizes of the differences between those and the other chiral ansätze. In contrast to the other quantities, for B_K we observe

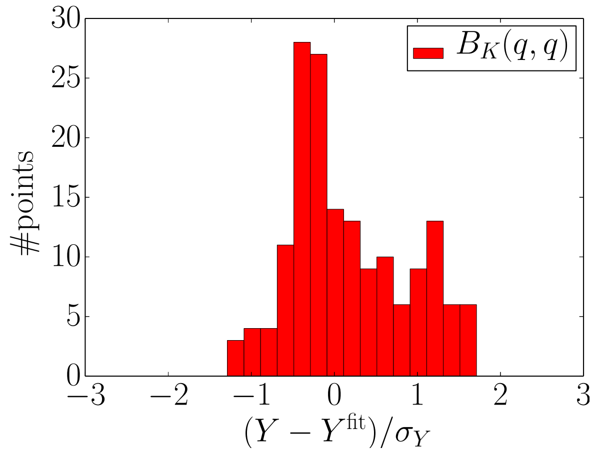


FIG. 28. A histogram of the deviation of the ChPTFV fit curve from our data in units of the statistical error for the $(\not{q},\not{q})$ intermediate scheme.

Parameter	ChPT	ChPTFV	Parameter	Analytic (260 MeV)	Analytic (370 MeV)
B_K^0	0.5280(16)	0.5278(16)	$C_0^{B_K}$	0.5316(28)	0.5322(17)
$c^I_{B_K,a}$	0.125(12)	0.128(12)	$C_a^{B_K,I}$	0.145(19)	0.129(12)
$c^{ID}_{B_K,a}$	0.148(15)	0.153(15)	$C_a^{B_K,ID}$	0.201(33)	0.164(15)
c_{B_K,m_x}	0.00492(64)	0.00420(64)	$C_1^{B_K}$	$-1.0(1.1)$	0.37(19)
c_{B_K,m_l}	$-0.00809(94)$	$-0.00728(95)$	$C_2^{B_K}$	0.58(68)	0.38(28)
c_{B_K,m_y}	1.316(32)	1.324(32)	$C_3^{B_K}$	1.547(96)	1.331(32)
c_{B_K,m_h}	$-0.13(18)$	$-0.06(18)$	$C_4^{B_K}$	0.50(55)	0.07(18)

TABLE XXIV. The B_K fit parameters for each of our chiral ansätze in the SMOM$(\not{q},\not{q})$ scheme at 3.0 GeV. The parameters are given in physical units and with the heavy quark mass expansion point adjusted to the physical strange quark mass. For the ChPT and ChPTFV ansatzë the chiral scale Λ_χ has been adjusted to 1 GeV.

that the differences between the ChPTFV and analytic ansätze are of the same order as the statistical error, although those differences are poorly resolved. Nevertheless, we choose to continue to neglect the chiral systematic error for the following reasons: We previously chose to treat the chiral extrapolation error as small not just because the differences between the analytic and ChPTFV forms are small, but because we have good evidence to believe that the ChPTFV fits are correctly capturing this behavior in addition to their strong theoretical motivation. This was not the case in our former works where we were extrapolating from heavier masses. There the analytic

	ChPTFV	Δ ChPT	Δ Analytic (260 MeV)	Δ Analytic (370 MeV)
$B_K(\not{q},\not{q})$	0.5341(18)	0.00020(11)	−0.0035(25)	−0.00029(21)
$B_K(\gamma^\mu,\gamma^\mu)$	0.5166(18)	0.00027(12)	−0.0037(24)	−0.00029(21)
$B_K(\overline{\text{MS}}$ via $\not{q},\not{q})$	0.5293(17)	0.00020(11)	−0.0035(24)	−0.00029(21)
$B_K(\overline{\text{MS}}$ via $\gamma^\mu,\gamma^\mu)$	0.5187(18)	0.00027(12)	−0.0037(24)	−0.00029(21)

TABLE XXV. The physical predictions for B_K in the two intermediate schemes and in the $\overline{\text{MS}}$ scheme (via the intermediate schemes) obtained using the ChPTFV ansatz, and the full correlated differences (labelled Δ) between the results obtained using the other ansätze and the ChPTFV result. Analytic fit differences are presented with a 370 MeV and 260 MeV pion mass cut.

fits were motivated by the apparent linearity in the available data with full knowledge that they do not correctly describe any underlying chiral curvature and are therefore not applicable over large mass ranges. Given that both fit forms were deficient in different ways, we conservatively took their full difference as an estimate of the error. On the other hand, in our new analysis we have a large amount of data in the light mass regime and the fits are forced to pass through data essentially at the physical point. As a result there is no longer any reason to distrust the ChPTFV results, especially given that they are only being used to perform a 4 MeV extrapolation in the pion mass. On the other hand there is now good evidence of chiral curvature in our results and therefore good reason to discount the analytic results. In fact, it is a testament of the robustness of our procedure that, despite this deficiency, the results obtained using these two ansätze differ only at the fraction-of-a-percent level.

We use the SMOM$(\not{q},\not{q})$ result for our central value, giving us a final continuum result in a non-perturbative MOM scheme with 0.3% total error after all sources of error are accounted for:

$$B_K(\not{q},\not{q},3\text{GeV}) = 0.5341(18)\,. \tag{94}$$

This final prediction, and the result in the SMOM(γ^μ,γ^μ) scheme, can be converted into the $\overline{\text{MS}}$ scheme using the following one-loop matching coefficients [40]:

$$C(\not{q},\not{q}\to\overline{\text{MS}}) = 0.99113\,,\quad C(\gamma^\mu,\gamma^\mu\to\overline{\text{MS}}) = 1.00408\,, \tag{95}$$

using $\alpha_s(3\text{ GeV}) = 0.24544$. The resulting $\overline{\text{MS}}$ values are also listed in Table XXV.

For the reasons discussed above, we use the value obtained via the SMOM$(\not{q},\not{q})$ scheme for our final $\overline{\text{MS}}$ result. The matching introduces a perturbative truncation error, which we estimate by taking the full difference between the results obtained using the two RI-SMOM intermediate schemes. We obtain:

$$B_K(\overline{\text{MS}}, 3\text{ GeV}) = 0.5293(17)(106)\,, \tag{96}$$

where the errors are statistical and from the perturbative matching to $\overline{\text{MS}}$ respectively.
In the renormalization group invariant (RGI) scheme, the above corresponds to

$$\hat{B}_K = 0.7499(24)(150)\,. \tag{97}$$

Previously [5] we obtained:

$$B_K(\overline{\text{MS}}, 3\text{ GeV}) = 0.535(8)(7)(3)(11)\,, \tag{98}$$

for which the errors are statistical, chiral, finite-volume and from the perturbative matching respectively. Comparing with the above, we see excellent agreement. Our new result offers a considerable improvement in the statistical error, but the truncation effects are the same as we have not changed the scale, and dominate the final error.

VI. CONCLUSIONS

Combining decades of theoretical, algorithmic and computational advances, we are finally able to perform $2+1$ flavor simulations with an essentially chiral action directly at the physical masses of the up, down and strange quarks in isospin symmetric QCD with both fine lattice spacings and large physical volumes. In this paper we report on two such ensembles; a $48^3 \times 96 \times 24$ (48I) ensemble and a $64^3 \times 128 \times 12$ (64I) ensemble, both using Möbius domain wall fermions. The inverse lattice spacings are $a^{-1} = 1.730(4)$ GeV and 2.359(7) GeV, respectively, and these ensembles have $m_\pi L = 3.863(6)$ and 3.778(8). We make use of the Möbius kernel with parameters chosen such that the Möbius and Shamir (traditional domain wall) kernels are identical, but the approximation to the sign-function of the four-dimensional effective action is improved in the former, resulting in a smaller residual chiral symmetry breaking for the same computational cost. The simulated pion masses are 139.2(4) and 139.2(5) MeV for the 48I and 64I ensembles respectively. These are slightly above the physical value, requiring a small extrapolation that we performed by combining these ensembles with several of our older Shamir domain wall ensembles in a simultaneous chiral/continuum 'global fit', specifically the $24^3 \times 64 \times 16$ (24I) and $32^3 \times 64 \times 16$ (32I) ensembles with the Iwasaki gauge action at $\beta = 2.13$ and 2.2 respectively, and the $32^3 \times 64 \times 32$ (32ID) ensemble with the Iwasaki+DSDR gauge action at $\beta = 1.75$. We also include a new $32^3 \times 64 \times 12$ (32Ifine) Shamir domain wall ensemble with the Iwasaki gauge action at $\beta = 2.37$, corresponding to $a^{-1} = 3.148(17)$ GeV, and a heavier 371(5) MeV pion mass; this enables us to examine the scaling behaviour of our data in the 1.75-3.15 GeV range of inverse lattice spacings to look for deviations from the leading a^2 scaling behavior. These ensembles give us access to a wide range of unitary and partially-quenched data ranging from the physical point up to the imposed 370 MeV pion-mass cut. As we use the same kernel for our Möbius and Shamir simulations, we are able to describe all of these ensembles using the same continuum scaling curve, apart from the 32ID ensemble which has a different gauge action.

The global fits are performed using the techniques developed in Refs. [6] and [5]. We fit to the following quantities: m_π, m_K, f_π, f_K, m_Ω and the Wilson flow scales w_0 and $t_0^{1/2}$. A separate fit is performed to the neutral kaon mixing parameter, B_K. To describe the mass dependence of these quantities we use NLO partially-quenched chiral perturbation theory with and without finite-volume corrections (referred to as the 'ChPTFV' and 'ChPT' ansätze) and also a linear 'analytic' ansatz.

Despite the significantly improved precision of the 48I and 64I data, we found that the fits missed these data by 1-2σ; this is an artifact of the large number of data points in the heavy-mass regime where χPT is only reliable to $\mathscr{O}(5\%)$. We resolve this issue by over-weighting the 48I and 64I data in order that the fit is forced to pass through these points. We emphasize that, while these global fits combine a large amount of data from various sources, the overweighting procedure guarantees that the predictions (and their statistical errors) are dominated by the near-physical data. A simpler procedure in which we simply treated the quark mass mistuning as an additional systematic error, would also obtain a similar statistical precision; the global fits essentially just remove these systematic effects.

The 48I and 64I ensembles each have the same gauge coupling as the corresponding 24I and 32I ensembles, but with smaller residual chiral symmetry breaking (significantly so for the former). We found that the differences in the fermion action between these two pairs of ensembles, each evaluated at the same gauge coupling, resulted in a 3.2(2)% difference between the 48I and 24I lattice scales, and a 1.1(2)% difference between that of the 64I and 32I ensembles. In Appendix C we show that this can be understood as an unexpectedly large effect of the changes in L_s and the Möbius scale parameter α which distinguish these ensembles, and provide added numerical evidence that these effects are accurately described by such shifts in the lattice scales.

We showed that due to the dominance of the 48I and 64I data, which were measured with near-physical pion masses on large, 5.5fm boxes, the systematic errors associated with the chiral extrapolation and finite-volume can be neglected. The errors on our final results, which we take from the ChPTFV fits, are dominated by statistics, and are themselves very small. For the pion and kaon decay constants we obtain $f_\pi = 130.2(9)$ MeV and $f_K = 155.5(8)$ MeV; for the average up/down quark mass and strange quark mass in the $\overline{\mathrm{MS}}$-scheme at 3 GeV, 2.997(49) and 81.64(1.17) MeV; the neutral kaon mixing parameter B_K in the RGI scheme, 0.750(15) and the $\overline{\mathrm{MS}}$-scheme at 3 GeV, 0.530(11); and the Wilson flow scales $t_0^{1/2} = 0.729(4)$ GeV^{-1} and $w_0 = 0.874(5)$ GeV^{-1}. In Table XXVI we compare our numbers to the $N_f = 2+1$ results compiled by the Flavor Lattice Averaging Group (FLAG) in their Review of Lattice Results [42].

Our results for the light and strange quark masses, obtained in Section V C, are renormalized in the $\overline{\mathrm{MS}}$ scheme at 3 GeV. The only remaining uncertainties on these quantities are statistical and perturbative matching errors, roughly 1% each. The renormalization and running of the quark masses were computed nonperturbatively, details of which can be found in Appendix F. The masses are quite consistent with our previous determinations, but show significant improvement due to the inclusion of the physical point ensembles. Our masses agree with the FLAG averages, but have errors that are both smaller than those of the average as well as those of any of the individual results used therein [5, 46–49]. The ratio of strange to light quark masses, shown in Eq. (86), is also consistent with the FLAG average [42], but here the error is slightly larger since systematic errors mostly cancel, though it is as small as any individual result used in the average [5, 46–49].

The FLAG average for the standard model kaon bag parameter is largely dominated by the Budapest-Marseille-Wuppertal collaboration (BMWc) result [45], $\hat{B}_K = 0.7727(81)_{\mathrm{stat}}(34)_{\mathrm{sys}}(77)_{\mathrm{PT}}$, where the errors are statistical, systematic and from perturbation theory, respectively. We would like to stress the difficulties one encounters in reliably assessing truncation errors, a point also emphasized by BMWc [45]. Among other checks, the BWMc showed that the NLO-perturbative and their non-perturbative running in the RI-MOM scheme agree between 1.8 and 3.5 GeV within statistical errors (of 2%), and quote 1% for the error due to perturbation theory, 2% being the size of the NLO term in the perturbative expansion. We proceed differently, by evaluating the difference between two different intermediate SMOM schemes, and estimate an error of 2%. We believe our procedure is more robust than those that have fed into the FLAG average, since multiple intermediate schemes were used to assess the truncation error. This error can certainly be reduced further in the future by performing the matching to $\overline{\mathrm{MS}}$ at higher scale or by computing

Quantity	This Work	FLAG Average	
f_π	130.19 ± 0.89 MeV	130.2 ± 1.4 MeV	[5, 43, 44]
f_K	155.51 ± 0.83 MeV	156.3 ± 0.9 MeV	[5, 43, 44]
f_K/f_π	1.1945 ± 0.0045	1.194 ± 0.005	[5, 43–45]
$m_u = m_d(\overline{\text{MS}}, 3\text{ GeV})$	$2.997 \pm 0.036 \pm 0.033$ MeV		
$m_s(\overline{\text{MS}}, 3\text{ GeV})$	$81.64 \pm 0.77 \pm 0.88$ MeV		
$m_s/m_u = m_s/m_d$	27.34 ± 0.21	27.46 ± 0.15	[5, 46–49]
$m_u = m_d(\overline{\text{MS}}, 2\text{ GeV})$	$3.315 \pm 0.040 \pm 0.036$ MeV	3.42 ± 0.06 MeV	[5, 47–49]
$m_s(\overline{\text{MS}}, 2\text{ GeV})$	$90.29 \pm 0.85 \pm 0.97$ MeV	93.8 ± 1.5 MeV	[5, 46, 48, 49]
$t_0^{1/2}$	$0.7292 \pm 0.0041\ \text{GeV}^{-1}$		
w_0	$0.8742 \pm 0.0046\ \text{GeV}^{-1}$		
$B_K(\text{SMOM}(\not{q}, \not{q}), 3\text{ GeV})$	0.5341 ± 0.0018		
$B_K(\overline{\text{MS}}, 3\text{ GeV})$	$0.5293 \pm 0.0017 \pm 0.0106$		
$\hat{B}_K$	$0.7499 \pm 0.0024 \pm 0.0150$	0.7661 ± 0.0099	[5, 45, 50, 51]
F_π/F	1.0645 ± 0.0015	1.0624 ± 0.0021	[46, 52, 53]
$[\Sigma(\overline{\text{MS}}, 3\text{ GeV})]^{1/3}$	$285.3 \pm 2.0 \pm 1.0$ MeV		
$[\Sigma(\overline{\text{MS}}, 2\text{ GeV})]^{1/3}$	$275.9 \pm 1.9 \pm 1.0$ MeV	271 ± 15 MeV	[6, 47, 52]
$\bar{l}_3$	2.73 ± 0.13	3.05 ± 0.99	[5, 47, 52, 53]
$\bar{l}_4$	4.113 ± 0.059	4.02 ± 0.28	[5, 47, 52, 53]

TABLE XXVI. Summary of results from the simulations reported here. The first error is the statistical error, which for most quantities is much larger than any systematic error we can measure or estimate. The exception is for the quantities in $\overline{\text{MS}}$ and $\hat{B}_K$. For these quantities, the second error is the systematic error on the renormalization, which is dominated by the perturbative matching between the continuum RI-MOM scheme and the continuum $\overline{\text{MS}}$ scheme. Comparison of our results to the averages compiled by the Flavor Lattice Averaging Group [42] for $N_f = 2+1$ flavor isospin symmetric QCD. Note that for $\hat{B}_K$, a direct comparison of the perturbative error is not possible since we use a different, and we believe more robust, method to estimate it. This perturbative error is common to our calculation and to the calculations dominating the FLAG average. In the rightmost column we provide the references to the original work that entered the quoted FLAG-averages. Light quark masses and the chiral condensate are given in the $\overline{\text{MS}}$ scheme, evaluated at 2 GeV. Results from this work have been run down from 3 GeV to 2 GeV using the running factor 1.106 from the FLAG review [42] and do not include the FLAG-estimated systematic error due to the omission of the charm sea quark.

the matching coefficient at NNLO. We want to emphasize that the errors quoted are different because the subjective procedures to estimate these errors are different. For completeness, we also compare the non-perturbative scale evolution to the NLO running between 2 and 3 GeV. We find a deviation of around 1.5% for the RI-SMOM(γ_μ,γ_μ) and for the RI-MOM schemes, and of 0.5% for the RI-SMOM$(\not{q},\not{q})$ scheme.

It is useful to compare our results with Ref. [45] in the intermediate MOM schemes (before converting to $\overline{\mathrm{MS}}$) as these numbers are purely non-perturbative:

$$B_K^{\mathrm{RI}}(3.5\ \mathrm{GeV}) = 0.5308(56)_{\mathrm{stat}}(23)_{\mathrm{sys}} \qquad \text{BMWc} \tag{99}$$

$$B_K^{(\not{q},\not{q})}(3\ \mathrm{GeV}) = 0.5341(18)_{\mathrm{stat}} \qquad \text{this work}\,, \tag{100}$$

where we neglect the various sources of systematic errors in our result since they are considerably smaller than the statistical error. These results are in different non-perturbative schemes and at different scales, and are therefore not directly comparable. However, we can compare their relative total errors: our result and that of BMWc have a 0.3% and a 1.1% relative error, respectively. We emphasize that in terms of objective statistical errors, and every systematic effect for which there is a theoretical framework for estimation (e.g. discretization, mass extrapolation, and finite volume), our new result is more precise than those entering the FLAG average. This is reflected in the 0.3% total relative error on results in a non-perturbatively defined $\not{q}$ RI scheme. Our assessment of the (subjective) perturbative systematic uncertainty on the conversion to $\overline{\mathrm{MS}}$ is more pessimistic than that of FLAG and BMWc, but we believe that it is better founded on the evidence of multiple intermediate schemes.

Predictions of $\hat{B}_K$ in lattice QCD have now reached a level of precision where other ingredients in its utilization for SM-tests are limiting progress (e.g. our knowledge on $|V_{cb}|$).

The results for the kaon and pion decay constant and their ratio are compatible with the FLAG average and amongst the most precise $N_f = 2+1$ predictions that have been made. Our results will certainly allow for further constraining CKM-unitarity tests [42].

The most significant remaining differences between our simulations and the physical world are isospin breaking and EM effects and the effect of quenching the charm quark.

Including isospin breaking effects requires using non-degenerate masses for the up and down quarks. This is possible within the domain wall fermion framework with current technology, for example using the rational quotient action or the one-flavor action developed by TWQCD [54]. However, these techniques are computationally demanding, and the effects in question are expected to be similar in size to the electromagnetic effects, hence there is limited value in considering these in isolation.

The RBC and UKQCD collaborations have performed exploratory calculations using QCD domain wall configurations with quenched electromagnetic interactions [55, 56] and have performed unquenched simulations using reweighting techniques [57]. There is increasing effort in the lattice community to control these effects, from more precise electro-quenched calculations [58, 59] (i.e. with EM included only in the valence sector) up to full QCD+QED simulations [60]. Adding QED to lattice simulations is challenging for many reasons. Firstly, adding a coupling constant to the theory, especially in the context of non-degenerate light quarks, considerably increases the cost of the simulations, particularly when using a chiral action close to the physical point. Secondly, the absence of mass gap in QED implies finite-size effects with power-law dependence on the lattice spatial extent, which are potentially large compared to the QED contributions [60, 61]. Finally, it is still not clear how to define quantities such as decay constants in QCD+QED, because the matrix elements are infrared divergent and gauge dependent [62]. Because of these issues, the addition of isospin-breaking effects and electromagnetism remains an important and challenging topic for our future calculations.

Dynamical charm effects are expected to be small for the majority of the quantities studied in this paper, but for quantities such as the $K_L - K_S$ mass difference and $K \to \pi\pi$ amplitudes they can have significant contributions. This is therefore the most promising avenue for RBC and UKQCD to take, allowing us to address these systematic errors on our flagship calculations. The biggest hurdle for including the charm is the requirement of simulating with finer lattice spacings, which tends to incur freezing of topology as well as requiring large computing power to obtain sufficiently large physical volumes. RBC and UKQCD have developed the 'dislocation enhancing determinant' (DED) method [63] to overcome the effects of the topology freezing, and have already commenced large-scale physical simulations with dynamical charm.

ACKNOWLEDGMENTS

The generation of the $48^3 \times 96$ and $64^3 \times 128$ Möbius DWF+Iwasaki ensembles was performed using the IBM Blue Gene/Q (BG/Q) "Mira" machines at the Argonne Leadership Class Facility (ALCF) provided under the Incite Program of the US DOE, on the STFC funded "DiRAC" BG/Q system in the Advanced Computing Facility at the University of Edinburgh, and on the

BG/Q machines at Brookhaven National Laboratory (BNL). The BG/Q computers of the RIKEN-BNL Research Center were used to generate the $32^3 \times 64$ fine 3.14 GeV ensemble. The DiRAC equipment was funded by BIS National E-infrastructure capital grants ST/K000411/1, STFC capital grant ST/H008845/1, and STFC DiRAC Operations grants ST/K005804/1 and ST/K005790/1. DiRAC is part of the National E-Infrastructure. Most of the measurements were also performed on the DiRAC and Mira machines, with the remainder performed using the BG/Q computers at the BNL.

The research leading to these results has also received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) ERC grant agreement #279757.

The software used includes the CPS QCD code (`http://qcdoc.phys.columbia.edu/cps.html`), supported in part by the USDOE SciDAC program; and the BAGEL (`http://www2.ph.ed.ac.uk/~paboyle/bagel/Bagel.html`) assembler kernel generator for high-performance optimized kernels and fermion solvers [64]. The gauge fixing for the 48I ensemble was performed using the CUTH cluster at Columbia University using the "GLU" (Gauge Link Utility) codebase (`https://github.com/RJhudspith/GLU`).

T.B is supported by U.S. DOE grant #DE-FG02-92ER41989. G.M, N.H.C, R.D.M, and D.J.M are supported in part by U.S. DOE grant #de-sc0011941. A.S, C.J, T.I and C.L are supported in part by US DOE Contract #AC-02-98CH10886(BNL). T.I is also supported by Grants-in-Aid for Scientific Research #26400261. C.T.S, T.J and A.P acknowledge the STFC grants ST/J000396/1 and ST/L000296/1. P.A.B, R.D.K, N.G and J.F acknowledge support from STFC grants ST/L000458/1 and ST/J000329/1. N.G also acknowledges support from STFC under the grant ST/J000434/1 and from the EU Grant Agreement #238353 (ITN STRONGnet). C.K is supported by a RIKEN foreign postdoctoral research (FPR) grant. R.J.H is supported by the Natural Sciences and Engineering Research Council of Canada.

Appendix A: Conserved currents of the Möbius domain wall action

The connection of the Möbius formulation to overlap fermions can be made at the propagator level and with the familiar DWF physical fields q_L and q_R. In the following subsection we repeat known but important results connecting the surface-to-surface and surface-to-bulk propagators of the Möbius domain wall action (in our conventions) with the four dimensional overlap propagator.

These results are then used to establish a practical implementation of the conserved axial and vector currents for the Möbius case.

1. Domain wall and overlap propagators, and contact terms

The approximate overlap operator can be written in terms of our four dimensional Schur complement matrices as

$$D_{ov} = S_\chi(m=1)^{-1} S_\chi(m). \tag{A1}$$

Observe that if we solve the following 5-D system of equations,

$$D^5_\chi(m=1)^{-1} D^5_\chi(m)\phi = \begin{pmatrix} q \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{A2}$$

and substitute the UDL decomposition, this yields

$$D_S^{-1}(m=1) D_S(m) L(m)\phi = L(m=1) \begin{pmatrix} q \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{A3}$$

Since $\left(L(m)(q,0,\ldots,0)^T\right)_1 = q$ and $(L(m)\phi)_1 = \phi_1$, the topmost row of our 5-D system of equations gives the overlap propagator:

$$S_\chi(m=1)^{-1} S_\chi(m) = \left(D^5_\chi(m=1)^{-1} D^5_\chi(m)\right)_{11}. \tag{A4}$$

This approximate overlap operator can however be expressed in terms of the $\bar{\psi}$ basis fields, and

$$\begin{aligned}
D_{ov} &= S_\chi(m=1)^{-1} S_\chi(m) && \text{(A5)}\\
&= \left[\mathscr{P}^{-1}\mathscr{P} D^5_\chi(m=1)^{-1} Q_-^{-1}\gamma_5\gamma_5 Q_- D^5_\chi(m)\mathscr{P}^{-1}\mathscr{P}\right]_{11} && \text{(A6)}\\
&= \left[\mathscr{P}^{-1} D^5_{GDW}(m=1)^{-1} D^5_{GDW}(m)\mathscr{P}\right]_{11}. && \text{(A7)}
\end{aligned}$$

The cancellation the Pauli-Villars term can be expressed in terms of *unmodified* generalized domain wall matrix D^5_{GDW}. The overlap contact term can be subtracted from the overlap propagator.

Here we define

$$\tilde{D}_{ov}^{-1} = \frac{1}{1-m}\left[D_{ov}^{-1} - 1\right] \tag{A8}$$

$$= \frac{1}{1-m}\left[\mathscr{P}^{-1} D^5_{GDW}(m)^{-1} D^5_{GDW}(m=1)\mathscr{P} - 1\right]_{11} \tag{A9}$$

$$= \frac{1}{1-m}\left\{\mathscr{P}^{-1} D^5_{GDW}(m)^{-1}\left[D^5_{GDW}(m=1) - D^5_{GDW}(m)\right]\mathscr{P}\right\}_{11}. \tag{A10}$$

Now, the difference $\left[D^5_{GDW}(m=1) - D^5_{GDW}(m)\right]_{ij} = (1-m)\left[P_-\delta_{i,L_s}\delta_{j1} + P_+\delta_{i,1}\delta_{j,L_s}\right]$. This relation is simpler to interpret in our convention than with the convention from Ref. [12]: the mass term is applied to our five dimensional surface fields without field rotation. With this,

$$\tilde{D}_{ov}^{-1} = \left\{\mathscr{P}^{-1} D^5_{GDW}(m)^{-1} R_5 \mathscr{P}\right\}_{11}. \tag{A11}$$

This is just the normal valence propagator of the physical DWF fields $q = (\mathscr{P}^{-1}\psi)_1$ and $\bar{q} = (\bar{\psi}R_5\mathscr{P})_1$. We see that the usual domain wall valence propagator has always contained both the contact term subtraction and the appropriate multiplicative renormalization of the overlap fermion propagator. As a result, the issues of lattice artifacts in NPR raised in Ref. [65] have never been present in domain valence analyses. This was guaranteed to be the case because Shamir's 5-D construction is designed to exactly suppress chiral symmetry breaking in the limit of infinite L_s, including any contact term.

For later use, we may also consider the propagator into the bulk from a surface field q for Möbius fermions,

$$\langle Q_s \bar{q}\rangle = \left[\mathscr{P}^{-1} D^5_{GDW}(m)^{-1} R_5 \mathscr{P}\right]_{s1} \tag{A12}$$

$$= \frac{1}{1-m}\left\{\mathscr{P}^{-1} D^5_{GDW}(m)^{-1} D^5_{GDW}(1)\mathscr{P} - \mathbb{1}\right\}_{s1} \tag{A13}$$

$$= \frac{1}{1-m}\left\{D^5_\chi(m)^{-1} D^5_\chi(1) - \mathbb{1}\right\}_{s1} \tag{A14}$$

$$= \frac{1}{1-m}\left\{L^{-1}(m) D^{-1}(m) D(1) L(1) - \mathbb{1}\right\}_{s1} \tag{A15}$$

$$= \frac{1}{1-m}\left\{L^{-1}(m)\left(\begin{array}{c|c} S_\chi^{-1}(m)S_\chi(1) & 0 \\ \hline 0 & \mathbb{1}\end{array}\right) L(1) - \mathbb{1}\right\}_{s1}. \tag{A16}$$

Now,

$$L(m) = \left(\begin{array}{c|c} 1 & 0 \\ \hline -T^{-(L_s-1)}(P_+ - mP_-) & \\ \vdots & \mathbb{1} \\ -T^{-1}(P_+ - mP_-) & \end{array}\right) \quad ; \; L(m)^{-1} = \left(\begin{array}{c|c} 1 & 0 \\ \hline T^{-(L_s-1)}(P_+ - mP_-) & \\ \vdots & \mathbb{1} \\ T^{-1}(P_+ - mP_-) & \end{array}\right) \tag{A17}$$

and so we have,

$$\langle Q_s\bar{q}\rangle = \frac{1}{1-m}\left(\begin{array}{c|c} D_{ov}^{-1}(m)-\mathbb{1} & 0 \\ \hline T^{-(L_s-1))}[(P_+ - mP_-)D_{ov}^{-1}(m)-\gamma_5] & \\ \vdots & 0 \\ T^{-1}[(P_+ - mP_-)D_{ov}^{-1}(m)-\gamma_5] & \end{array}\right)_{s1} \tag{A18}$$

$$= \left(\begin{array}{c|c} \left[P_+ + P_-T^{-L_s}\right] & 0 \\ \hline T-(L_s-1) & \\ \vdots & 0 \\ T^{-1} & \end{array}\right)_{s1} \left[1+T_1^{-1}\cdots T_{L_s}^{-1}\right]^{-1} D_{ov}^{-1}(m). \tag{A19}$$

Finally, applying the permutation matrix, we have the five dimensional propagator from a physical field,

$$G_q = \mathscr{P}\langle Q_s\bar{q}\rangle = [P_+ + P_-T^{-1}]\begin{pmatrix} T^{-(L_s-1)} \\ T^{-(L_s-2)} \\ \vdots \\ T^{-1} \\ 1 \end{pmatrix}[1+T^{-L_s}]^{-1}D_{ov}^{-1}. \tag{A20}$$

The connection between domain wall systems and the overlap, well established in the literature and reproduced in this section, is useful in understanding the relation of domain wall fermions to their 4-D effective action.

2. Conserved vector and axial currents

The standard derivation of lattice Ward identities proceeds as follows. A change of variables of the fermion fields ψ and $\bar{\psi}$ at a single site y is performed:

$$\psi_y' = \psi_y - i\alpha\psi_y \qquad ; \qquad \bar{\psi}_y' = \bar{\psi}_y + i\bar{\psi}_y\alpha\,. \tag{A21}$$

Under the path integral, the Jacobian is unity, and the partition function is left invariant:

$$Z' = \int d\bar{\psi}d\psi e^{-S[\bar{\psi},\psi]}\left\{1-i\alpha\left[\frac{\delta S}{\delta\psi_y}\psi_y - \bar{\psi}_y\frac{\delta S}{\delta\bar{\psi}_y}\right]\right\} = Z. \tag{A22}$$

Hence,

$$\langle\frac{\delta S}{\delta\psi_y}\psi_y - \bar{\psi}_y\frac{\delta S}{\delta\bar{\psi}_y}\rangle = 0. \tag{A23}$$

The Wilson action gives eight terms from varying $\bar\psi_y$ and eight terms from varying ψ_y due to the 4-D hopping stencil:

$$\bar\psi\delta_y(D_W)\psi = \Delta^-_\mu J^W_\mu(y) = \sum_\mu \left[\begin{array}{c} -\bar\psi_y\frac{1-\gamma_\mu}{2}U_\mu(y)\psi_{y+\hat\mu} + \bar\psi_{y-\hat\mu}\frac{1-\gamma_\mu}{2}U_\mu(y-\hat\mu)\psi_y \\ -\bar\psi_y\frac{1+\gamma_\mu}{2}U^\dagger_\mu(y-\hat\mu)\psi_{y-\hat\mu} + \bar\psi_{y+\hat\mu}\frac{1+\gamma_\mu}{2}U_\mu(y)^\dagger\psi_y \end{array} \right] \tag{A24}$$

$$= \Delta^-_\mu\left[\bar\psi_y\frac{1-\gamma_\mu}{2}U_\mu(y)\psi_{y+\hat\mu} - \bar\psi_{y+\hat\mu}U^\dagger_\mu(y)\frac{1+\gamma_\mu}{2}\psi_y\right] = 0, \tag{A25}$$

where Δ^-_μ is the backwards discretized derivative.

An equivalent alternate approach may be taken, however, and this is a better way to approach non-local actions such as the chiral fermions. Gauge symmetry leaves the action invariant at $O(\alpha)$ under the simultaneous active substitution, for a fixed site y of

$$U_\mu(y)\to(1+i\alpha)U_\mu(y) \qquad ; \qquad U_\mu(y-\hat\mu)\to U_\mu(y-\hat\mu)(1-i\alpha) \tag{A26}$$

and

$$\psi_y\to(1+i\alpha)\psi_y \qquad ; \qquad \bar\psi_y\to\bar\psi_y(1-i\alpha).. \tag{A27}$$

A change of variables on the fermion fields at site y may be performed simultaneously to absorb the phase on the fermions:

$$\psi^\prime_y = (1+i\alpha)\psi_y \qquad ; \qquad \bar\psi^\prime_y = \bar\psi_y(1-i\alpha). \tag{A28}$$

Under the path integral, the Jacobian is again unity, and the phase associated with the fermion is absorbed. We can now view the change in action as being associated with the *unabsorbed* phases on the eight gauge links connected to site y:

$$Z^\prime = Z = \int d\bar\psi^\prime d\psi^\prime e^{-S[\bar\psi^\prime,\psi^\prime,U]}\left\{1+i\alpha\sum_\mu\left[\frac{\delta S}{\delta U_\mu(y)^{ij}}U_\mu(y)^{ij} - \frac{\delta S}{\delta U_\mu(y-\mu)^{ij}}U_\mu(y-\mu)^{ij}\right]\right\}. \tag{A29}$$

For a gauge invariant Lagrangian we can *always* use a picture where the same change in action, and same current conservation law may be arrived at by differentiating with respect to the eight links connected to a site:

$$\langle\sum_\mu\left[\frac{\delta S}{\delta U_\mu(y)^{ij}}U_\mu(y)^{ij} - \frac{\delta S}{\delta U_\mu(y-\mu)^{ij}}U_\mu(y-\mu)^{ij}\right]\rangle = 0. \tag{A30}$$

This arises because the phase freedom of fermions and of gauge fields are necessarily coupled and inseparable in a gauge theory. For the nearest-neighbor Wilson action, this generates the same eight terms entering $\Delta^-_\mu J_\mu = 0$.

In the case of non-local actions, the Dirac matrix, whatever its form, can be viewed as a sum of gauge covariant paths. When generating a current conservation law from $U(1)$ rotation of the fermion field at site y, we sum over all fields $\bar{\psi}(x)$ and $\psi(x)$ connecting through the Dirac matrix $D(x,y)$ to the fixed site $\psi(y)$ and $\bar{\psi}(y)$. The following sum is always constrained to be zero for all y, and is identical to that found by Kikukawa and Yamada [66]:

$$\sum_x \bar{\psi}_x D(x,y)\psi_y - \bar{\psi}_y D(y,x)\psi_x = 0. \tag{A31}$$

The partitioning of this sum of terms, into a paired *discrete divergence operator* and *current* is not obvious, and it is cumbersome to generate Kikukawa and Yamada's *non-local kernel*.

It is instructive to consider what happens if we derive the same sum of terms by differentiating with respect to the 8 links connected to site y.

$$\langle \sum_\mu \left[\frac{\delta S}{\delta U_\mu(y)^{ij}} U_\mu(y)^{ij} - \frac{\delta S}{\delta U_\mu(y-\mu)^{ij}} U_\mu(y-\mu)^{ij} \right] \rangle = 0 \tag{A32}$$

The structure of Eq. (A32) *always* lends itself interpretation as a backwards finite difference. For a non-local action, the differentiation Eq. (A32) appears to generate a lot more terms than the fermion field differentiation Eq. (A31). The reason is clear: these extra terms are constrained by gauge symmetry to sum to zero, but only after cancellation between the different terms in Eq. (A32). Specifically, we consider an action constructed as the product of Wilson matrices:

$$S = \sum_{xyzw} \bar{\psi}_x D_W(x,y) D_W(y,z) D_W(z,w)\psi(w). \tag{A33}$$

The link variation approach gives three terms, each of which are conserved under a nearest-neighbor difference divergence: varying with respect to the 8 links we obtain, via the product rule,

$$\delta_y(\bar{\psi} D_W D_W D_W \psi)\psi = \psi\left[(\delta_y D_W)D_W D_W + D_W(\delta_y D_W)D_W + D_W D_W(\delta_y D_W)\right]\psi. \tag{A34}$$

Each of these contributions contain a backwards difference operator, and it is trivial to split this into a divergence and corresponding conserved current using Eq. (A24).

The above comment is generally applicable to any function of the Wilson matrix. We take this approach to establish the exactly-conserved vector current of an approximate overlap operator, where the approximation is represented by a rational function. We will also establish that matrix elements of this current are identical to those of the Furman and Shamir approach [9] in the case of domain wall fermions. The Furman and Shamir approach will then be used to also establish an axial Ward identity for our generalized Möbius domain wall fermions under which an explicitly known defect arises. This is important in both renormalizing lattice operators and also in determining the most appropriate measure of residual chiral symmetry breaking in our simulations. We construct the conserved vector current by determining the variation in the overlap Dirac operator, $\delta_y D_{ov}$:

$$\begin{aligned}
\delta_y D_{ov} &= \frac{1-m}{2}\gamma_5 \left\{ \delta_y(\frac{1}{1+T^{-Ls}})[1-T^{-Ls}] + \frac{1}{1+T^{-Ls}}\delta_y(1-T^{-Ls}) \right\} \\
&= \frac{1-m}{2}\gamma_5 \left\{ \delta_y(\frac{1}{1+T^{-Ls}}) - \frac{1}{1+T^{-Ls}}\delta_y(T^{-L_s})\left(1-\frac{T^{-L_s}}{1+T^{-Ls}}\right) \right\} \\
&= (1-m)\gamma_5\delta_y\left(\frac{1}{1+T^{-Ls}}\right).
\end{aligned} \tag{A35}$$

We can similarly find the variation in T^{-1} induced by a variation in D_W, where the variation in D_W is just the backwards divergence of the standard Wilson conserved current operator. Denoting,

$$\begin{aligned}
T^{-1} &= -(\tilde{Q}_-)^{-1}\tilde{Q}_+ \\
\tilde{Q}_- &= D_+^s P_- - D_- P_+ = D_-\gamma_5 Q_- \\
\tilde{Q}_+ &= D_+^s P_+ - D_- P_- = D_-\gamma_5 Q_+,
\end{aligned} \tag{A36}$$

we see that

$$\begin{aligned}
\delta_y(T^{-1}) &= -\tilde{Q}_-^{-1}\left\{-\delta_y(\tilde{Q}_-)\tilde{Q}_-^{-1}\tilde{Q}_+ + \delta_y(\tilde{Q}_+)\right\} \\
&= -\tilde{Q}_-^{-1}\left\{\delta_y(\tilde{Q}_-)T^{-1} + \delta_y(\tilde{Q}_+)\right\} \\
&= -\tilde{Q}_-^{-1}\delta_y(D_W)\left\{(bP_- + cP_+)T^{-1} + bP_+ + cP_-\right\}.
\end{aligned} \tag{A37}$$

Since

$$\begin{aligned}
\tilde{Q}_- P_- &= (1+bD_W)P_- \;;\; \tilde{Q}_+ P_- = (cD_W - 1)P_- \\
\tilde{Q}_- P_+ &= (cD_W - 1)P_+ \;;\; \tilde{Q}_+ P_+ = (1+bD_W)P_+,
\end{aligned}$$

we may re-express the identity

$$(b+c)(P_+ + P_-) = c\tilde{Q}_- P_- - b\tilde{Q}_+ P_- + c\tilde{Q}_+ P_+ - b\tilde{Q}_- P_+ \tag{A38}$$

$$\tilde{Q}_-^{-1}(P_+ + P_-) = \frac{\tilde{Q}_-^{-1}}{b+c}\left[\tilde{Q}_+(cP_+ - bP_-) + \tilde{Q}_-(cP_- - bP_+)\right], \tag{A39}$$

and this lets us find a symmetrical form:

$$(b+c)\delta_y(T^{-1}) = \left[b[P_+ - T^{-1}P_-] + c[T^{-1}P_+ - P_-]\right]\delta_y(D_W)\left[b[P_+ + P_- T^{-1}] + c[P_+ T^{-1} + P_-]\right].$$

We may now look at the variation of the term T^{-L_s}

$$\delta_y(T^{-L_s}) = \sum_{s=1}^{L_s} T^{-(s-1)} \left[\begin{array}{c} b[P_+ - T^{-1}P_-] \\ +c[T^{-1}P_+ - P_-] \end{array} \right] \delta_y(D_W) \left[\begin{array}{c} b[P_+ + P_-T^{-1}] \\ +c[P_+T^{-1} + P_-] \end{array} \right] T^{-(L_s-s)}. \tag{A40}$$

Compiling these results, we find

$$\delta_y D_{ov} = -\frac{1-m}{b+c}\gamma_5 \frac{1}{1+T^{-Ls}} \left(\sum_{s=1}^{L_s} T^{-(s-1)} \delta_y(T^{-1}) T^{-(L_s-s)} \right) \frac{1}{1+T^{-Ls}}. \tag{A41}$$

The terms may be expanded until insertions of the the backwards divergence of the Wilson current are reached (Eq. (A24)). Gauge symmetry then implies the conservation of the obvious current and the vector Ward identities can be constructed. For example, we may take as source $\eta^{jj'\alpha\alpha'}(z) = \delta_{jj'}\delta_{\alpha\alpha'}\delta^4(z-x)$ and a two-point function of the conserved current may be constructed as

$$\Delta_\mu^- \langle \bar{\psi}\gamma_\nu\psi(x)|\mathscr{V}_\mu(y)\rangle = \mathrm{Tr}\gamma_\nu\gamma_5\eta^\dagger D_{ov}^{-\dagger}\gamma[1+T^{-L_s}]^{-1} \left\{ \sum_{s=0}^{L_s-1} T^{-s}\delta_y(T^{-1})T^{-(L_s-1-s)} \right\} [1+T^{-L_s}]^{-1} D_{ov}^{-1}\eta\ . \tag{A42}$$

Note that when $c=0$, the insertion of Eq. (A40) contains only terms such as

$$[P_-T^{-1} + P_+], \tag{A43}$$

which are also present in the surface to bulk propagator Eq. (A20). As one would expect, when we take b and c to represent domain wall fermions, the two-point function of our exactly conserved vector current - derived from the four dimensional effective action - exactly matches the matrix element of the vector current constructed by Furman and Shamir [9], Eq. (2.21), from a five dimensional interpretation of the action.

Since the Furman and Shamir current was easily constructed from the five dimensional propagator Eq. (A20), one might hope to do the same in the generalized approach to domain wall fermions. To play a similar trick for the c term, we would need to generate the terms

$$P_-[1+T^{-L_s}]^{-1}D_{ov}^{-1}, \tag{A44}$$

and

$$P_+T_1^{-1}[1+T^{-L_s}]^{-1}D_{ov}^{-1}. \tag{A45}$$

These are not manifestly present in Eq. (A19). However, the presence of the contact term on the $s=0$ slice can be removed after a propagator calculation. We define this slice as

$$S(x) = \langle Q_0\bar{q}\rangle = \frac{1}{1-m}\left(D_{ov}^{-1}(m) - \mathbb{1}\right). \tag{A46}$$

In a practical calculation, the source vector η may be used to eliminate the contact term by forming

$$(1-m)S(x)\eta+\eta=D_{ov}^{-1}(m)\eta=[1+T^{-L_s}][1+T^{-L_s}]^{-1}D_{ov}^{-1}\eta. \tag{A47}$$

By applying P_+ and P_- we find we have the following set of vectors

$$\begin{pmatrix} P_+ \\ P_-T^{-L_s}P_+[1+T^{-L_s}] \\ P_-[1+T^{-L_s}] \end{pmatrix}[1+T^{-L_s}]^{-1}D_{ov}^{-1}, \tag{A48}$$

and we may eliminate to form a L_s+1 vectors from a 4-D source η

$$T(s)=\begin{pmatrix} 1 \\ T^{-1} \\ \vdots \\ T^{-L_s} \end{pmatrix}\left[1+T_1^{-1}\cdots T_{L_s}^{-1}\right]^{-1}D_{ov}^{-1}(m)\eta. \tag{A49}$$

This may be used to construct

$$\left[b[P_++P_-T^{-1}]+c[P_+T^{-1}+P_-]\right]T^s, \tag{A50}$$

for $s\in\{0\ldots L_s-1\}$, and by contracting these vectors through the Wilson conserved current the matrix element, Eq. (A42), can be formed in a *very* similar manner to the standard DWF conserved vector current. When $c=0$ the matrix element reduces to being *identical* to that for the Furman and Shamir vector current.

A flavor non-singlet axial current, almost conserved under a backwards difference operator, can now also be constructed following Furman and Shamir. We associate a fermion field rotation

$$\psi(x,s)\to\begin{cases} e^{i\alpha\Gamma(s)}\psi(x,s) & ;\ x=x_0 \\ \psi(x,s) & ;\ x\neq x_0 \end{cases}, \tag{A51}$$

where

$$\Gamma(s)\to\begin{cases} -1 & ;\ 0\leq s<L_s/2 \\ 1 & ;\ L_s/2\leq s \end{cases}. \tag{A52}$$

We acquire a related (almost-) conserved axial current, whose pseudoscalar matrix element is

$$\begin{array}{c}\Delta_\mu^-\langle\bar{\psi}\gamma_5\psi(x)|\mathscr{A}_\mu(y)\rangle= \\ \mathrm{Tr}[\eta^\dagger\tilde{D}_{ov}^{-\dagger}\gamma_5][1+T^{-L_s}]^{-1}\left\{\sum_{s=0}^{L_s-1}T^{-s}\Gamma(s)\delta_y(T^{-1})T^{-(L_s-1-s)}\right\}[1+T^{-L_s}]^{-1}D_{ov}^{-1}\eta\end{array}. \tag{A53}$$

The exact vector current conservation induces the same J_{5q} midpoint density defect that arose for DWF, and the Ward identity is

$$\Delta_\mu^- \langle \bar{\psi}\gamma_5\psi(x)|\mathscr{A}_\mu(y)\rangle = \langle \bar{\psi}\gamma_5\psi(x)|2mP(y)+2J_{5q}(y)\rangle. \tag{A54}$$

This allows us to retain the usual definition of the residual mass in the case of Möbius domain wall fermions. We emphasize that the definition,

$$m_{res} = \left.\frac{\langle \pi(\vec{p}=0)|J_{5q}\rangle}{\langle \pi(\vec{p}=0)|P\rangle}\right|_{m=-m_{res}},$$

via the zero-momentum pion matrix element of J_{5q} is particularly important, because then our PCAC relation,

$$\langle \pi(\vec{p}=0)|2mP+2J_{5q}\rangle = 0,$$

guarantees that the low momentum lattice pions are massless. This is the appropriate measure of chiral symmetry breaking for the analysis of the chiral expansion.
Section III C discusses methods of using the vector and axial ward identities to measure the renormalization of the local vector and axial currents, and their use in our analysis.

Appendix B: Deriving dimensionless global fit forms

In this section we briefly describe how to obtain the appropriate dimensionless global fit function describing the lattice data for a quantity Q of mass dimension D on a general ensemble e. The procedure is as follows:

1. Write down the fit formula for Q in *physical units* on the reference ensemble, including an a^2 term. For example, a linear ansatz might have the following form:

$$Q = c_{Q,0}(1+c_{Q,a}a_r^2) + c_{Q,m_l}\tilde{m}_l^r + c_{Q,m_h}\tilde{m}_h^r,$$

   where we have assumed that there are no partially-quenched data points for simplicity. Here the superscript r on the quark masses indicates that they are in the normalization of the reference ensemble.

2. To derive the fit form for Q on ensemble e, first replace a_r with the lattice spacing, a_e, appropriate for that ensemble, then rewrite a_e as $a_e = a_r/R_a^e$:

$$Q = c_{Q,0}(1+c_{Q,a}a_r^2(R_a^e)^{-2}) + c_{Q,m_l}\tilde{m}_l^r + c_{Q,m_h}\tilde{m}_h^r.$$

3. Multiply by a_r^D and redefine the fit parameters in terms of dimensionless quantities (denoted with a prime superscript):

$$a_r^D Q = c'_{Q,0}(1 + c'_{Q,a}(R_a^e)^{-2}) + c'_{Q,m_l}(a_r \tilde{m}_l^r) + c'_{Q,m_h}(a_r \tilde{m}_h^r)\,.$$

4. Using $a_r = R_a^e a_e$, rewrite the function in terms of the lattice spacing on the ensemble e:

$$(R_a^e)^D (a_e^D Q) = c'_{Q,0}(1 + c'_{Q,a}(R_a^e)^{-2}) + c'_{Q,m_l} R_a^e (a_e \tilde{m}_l^r) + c'_{Q,m_h} R_a^e (a_e \tilde{m}_h^r)\,.$$

5. Finally, use $\tilde{m}^r = Z_l^e \tilde{m}^e$ to move the quark masses into the native normalization of ensemble e, and divide by $(R_a^e)^D$:

$$(a_e^D Q) = (R_a^e)^{-D} c'_{Q,0}(1 + c'_{Q,a}(R_a^e)^{-2}) + c'_{Q,m_l}(R_a^e)^{1-D} Z_l^e (a_e \tilde{m}_l^e) + c'_{Q,m_h}(R_a^e)^{1-D} Z_h^e (a_e \tilde{m}_h^e)\,.$$

This fit function now describes the data in lattice units for the ensemble e.

Appendix C: Dependence of the lattice spacing on the fermion action

In Sec. IV we described that, contrary to our expectations, combining the 24I and 48I ensembles into a single global fit required that two lattice spacings, differing by 3.2(2)%, be used for these two, nominally similar ensembles. (Similar but smaller discrepancies between the lattice spacings for the 32I and 64I ensembles were also found.) In this appendix we will discuss this phenomenon in greater detail and describe additional measurements that we performed in order to verify that this assignment of different lattice spacings is correct. For clarity we will focus on the 24I and 48I ensembles, since the explanation for both cases is the same. For the 24I ensemble set we consider only the ensemble with the lighter input quark mass of $m_f = 0.005$

The 24I and 48I ensembles are very similar. Each uses the same Iwasaki gauge action with the same value of $\beta = 2.13$. They differ in the fermion formulation used (Shamir and Möbius respectively), the total light quark mass ($m_f + m_{\rm res} = (5.0 + 3.154(15)) \times 10^{-3} = 8.154(15) \times 10^{-3}$ and $m_f + m_{\rm res} = (7.8 + 6.102(40)) \times 10^{-4} = 13.999(40) \times 10^{-4}$, respectively) and the degree of residual chiral symmetry breaking, which is suggested by the differences in the values of the residual quark masses just quoted. For a comparison of the $m_f = 0.004$ 32I and 64I ensembles, the corresponding numbers are $m_f + m_{\rm res} = (4.0 + 0.6664(76)) \times 10^{-3} = 4.6664(76) \times 10^{-3}$ and $m_f + m_{\rm res} = (6.78 + 3.116(23)) \times 10^{-4} = 9.896(23) \times 10^{-4}$ respectively.

If we were to describe the low energy Green's functions computed on the 24I and 48I ensembles as corresponding to separate Symanzik effective theories, these two effective theories would be essentially identical, except for differences in their low energy constants of order $(ma)^n$. For example, in a theory with chiral fermions the dimension-4 $(F^{\mu\nu})^2$ term, closely related to the lattice scale, would have coefficients which differed by terms of order $(ma)^2$, terms much too small to be relevant here. Of course, had such a term been important, our global fitting procedure would have included its effects by describing both the 24I and 48I ensembles with a single Symanzik effective theory, with a single lattice spacing, whose mass-dependent coefficients were represented by explicit mass-dependent terms in the fit. In this framework both the 24I and 48I ensembles would be described by the same lattice spacing a and the same value of R_a.

It may be useful to briefly review the meaning of the lattice spacing a as it is generally defined in field theory and specifically defined in the calculation presented here. Perhaps the simplest way to define the cut-off scale is by specifying the value of a "physical" quantity, such as the Wilson flow or three-gluon coupling, at a sufficiently short flow time or large gluon momentum that the process can be understood in perturbation theory. Theories with identical lattice actions but with different quark masses will give the same value for the lattice scale up to terms of order $(ma)^2$ if we introduce the lattice scale a as the natural lower/upper limit on the flow times or momentum scales that are available for such a short-distance definition. From this perspective, such mass dependent effects are much too small to result in the 3% discrepancy we find. In our actual approach, we define the lattice spacing through the mass of the Ω^-. This requires our global fitting procedure and an explicit extrapolation to a specific value of input quark masses, specifically those which give physical values for m_π/m_Ω and m_K/m_Ω, in order that such a low-energy definition of the lattice scale be well defined. Necessarily, in this approach the 24I and 48I ensembles are assigned a common lattice spacing and their different input quark masses are completely accounted for in the global fitting procedure (up to negligible systematic effects). For our low-energy definition of the lattice spacing, it is not possible to interpret the 3% difference in a between the 24I and 48I ensembles as resulting from their different input masses.

Instead, the change in the lattice spacing between the 24I and 48I ensembles must be attributed to some other change in the lattice action. We are left to conclude that this effect must be a result of the change in fermion formulation. As discussed in Section II, we can consider this change as being accomplished in two steps: we first change L_s from 16 to 48 using the Shamir formulation, and then change from the Shamir ($L_s = 48$, $b+c = 1$ to the Möbius ($L_s = 24$, $b+c = 2$) formulation at fixed $L_s(b+c)$. Since all 4-dimensional Green's functions related by this final change are expected to agree at the 0.1% level, the Shamir to Möbius change is inconsistent with a 3% change in the lattice spacing, which would naturally result in a 3% change in such Green's functions. (For example, a change in the Omega mass of 3% would result in at least a 3% change in the Omega propagator.)

Thus, we expect that this 3% change in lattice spacing would have been observed even if we had continued to use the Shamir action and simply increased L_s from 16 to 48. While this is a surprisingly large effect for such a change in L_s, we believe that it is a plausible explanation. The effect of the smaller $L_s = 16$ value is usually characterized by the value of $m_{\rm res}a = 3.154(15) \times 10^{-3}$, which is substantially less than 3%. However, considerable effort has been devoted to reducing the size of $m_{\rm res}$, including a careful choice for the domain wall parameter M_5 and the choice of the Iwasaki gauge action. It is possible that, while these choices have significantly reduced $m_{\rm res}$, they have not correspondingly reduced the size of other L_s-dependent effects.

For example, the value of the lattice spacing, which is determined by the strength of QCD interactions at the scale of $\Lambda_{\rm QCD}$, is a strong function of the anti-screening produced by QCD vacuum polarization. The quarks act to reduce this anti-screening, and the Pauli-Villars determinant was originally included in the domain wall fermion action [8] to regulate what would have been a divergent contribution to QCD vacuum polarization coming from the increasing number of fermion species as $L_s \to \infty$. While, as can be seen by the relation with overlap fermions discussed in Sec. II, these effects have a well defined $L_s \to \infty$ limit, we cannot rule out the possibility that they appear at the 3% level for $\beta = 2.13$ and $L_s = 16$. Instead, we interpret this large shift in a as providing new information about the potential effects of finite L_s, and a warning that simple estimates can occasionally be misleadingly low. In this spirit, we should recognize that the earlier arguments about the insensitivity of the coefficients of the $O(a^2)$ Symanzik correction terms to our change in action may underestimate these effects. Of course, in this case, if our few tenths of a percent estimate were to become even a 5% effect, it would not interfere with our current continuum extrapolations.

Since the conclusion, implied by our global fits, that the lattice spacing did indeed change by 3.2% and 1% when going from the 24I to 48I and 32I to 64I ensembles respectively, was a surprise, it was important to test this hypothesis. For that purpose, we generated two additional MDWF+I ensembles with input parameters set equal to those of the lightest 24I and 32I ensembles (i.e. those with $am_l = 0.005$ and $am_l = 0.004$, respectively), but using the Möbius parameters and L_s values that were used for the 48I and 64I ensembles respectively. We compensated for the reduction in the residual mass by increasing the input bare quark mass in order that the total quark masses remain equal to the 24I and 32I values. If the observed differences in the lattice spacings can indeed be attributed to the change in L_s (that which would have been required if the new ensembles were generated with the Shamir action), then the lattice scales for these new ensembles should match those determined for the 48I and 64I ensembles.

We refer to these new ensembles as the '24Itest' and '32Itest' ensembles. They were generated with Möbius domain wall fermions and the Iwasaki gauge action at $\beta = 2.13$ and 2.25 respectively, and with lattice sizes of $24^3 \times 64 \times 24$ and $32^3 \times 64 \times 12$. Both ensembles use Möbius parameters of $\alpha = b + c = 2$ and $b - c = 1$, making them equivalent to Shamir domain wall ensembles with $L_s = 48$ and 24 respectively. On the 24Itest ensemble, we measured the residual mass and Wilson flow scales on configurations in the range 120 to 550; the residual mass was measured every 40 configurations, and the Wilson flow scales every 10, and we binned the latter over four successive measurements. Similarly, for the 32Itest ensemble, we performed measurements in the configuration range 200 to 610, measuring the residual mass every 20 and the Wilson flow scales every 10, binning the latter over two successive measurements.

The values of the average plaquette, residual mass, total quark mass and Wilson flow scales are listed in Table XXVII. From the table we can immediately observe that, while the total quark masses of the 24Itest and 24I ensembles are closely matched, there are clear differences in the average plaquette and Wilson flow scales; smaller differences are also observable between the 32Itest and 32I measurements. The differences in the Wilson flow scales are $\sim 3\%$ between the 24I and 24Itest ensembles and $\sim 1\%$ between the 32I and 32Itest, which are very similar to the differences in lattice scales observed between the 24I/48I and 32I/64I ensembles respectively.

We cannot directly compare the computed values on the test ensembles in Table XXVII with the corresponding 48I and 64I values, due to the measurements being performed with different quark masses. For a definitive test, we instead include the test ensembles in the global fits. For each ensemble there are associated three free parameters: the scaling parameters Z_l, Z_h and R_a. The observed differences in the fermion action appear to result in negligible changes to Z_l and Z_h, hence we are able to fix those values to those of the 24I/48I (for the 24Itest) and 32I/64I (for the 32Itest ensemble); this leaves only R_a as a free parameter for each ensemble. In Table XXVIII we list the values of R_a that we obtain, alongside the corresponding values for the 24I, 32I, 48I and 64I ensembles. We observe excellent agreement between R_a on the 24Itest ensemble and that on the 48I, and similarly between the 32Itest and 64I. This offers clear evidence that the change in L_s

Quantity	24I (0.005)	24Itest	32I (0.004)	32Itest
$\langle P\rangle$	0.588053(4)	0.587035(6)	0.615587(3)	0.615318(8)
m_l	0.005	0.00746	0.004	0.00437
m_h	0.04	0.04246	0.03	0.03037
$m_{\rm res}$	0.003154(15)	0.000666(25)	0.0006697(34)	0.000306(9)
$m_l+m_{\rm res}$	0.008154(15)	0.008126(25)	0.0046697(34)	0.004676(9)
$m_h+m_{\rm res}$	0.043154(15)	0.043126(25)	0.0306697(34)	0.030676(9)
$t_0^{1/2}$	1.3163(6)	1.2766(19)	1.7422(11)	1.7226(24)
w_0	1.4911(15)	1.4485(46)	2.0124(26)	1.9937(57)

TABLE XXVII. Comparison of various quantities in lattice units between the test ensembles and the original ensembles. For the 24I and 32I ensembles we quote values for the residual mass computed at unitary light quark masses (not extrapolated to the chiral limit). These and the average plaquette values were determined in Ref. [6]. The Wilson flow scales on these ensembles are discussed in Appendix E. For comparison, the residual masses for the 48I and 64I ensembles are 0.000610(4) and 0.000312(2) respectively.

$\beta=2.13$			$\beta=2.25$		
24I	48I	24Itest	32I	64I	32Itest
0.7491(23)	0.7259(27)	0.7243(28)	1.0(0)	0.9897(19)	0.9877(19)

TABLE XXVIII. The values of the lattice spacing ratio $R_a^i = a^{32\mathrm{I}}/a^i$ for ensembles i with $\beta=2.13$ and $\beta=2.25$, including the two test ensembles.

is responsible for the observed differences in lattice spacing. It provides further confidence in our global fitting procedure, which was sufficiently reliable to produce strong evidence for this effect even though it was not expected in advance.

Note, in this explanation we continue to assume the near equality of the Möbius and Shamir 4-D theories for fixed $L_s(b+c)$, and to view the difference in a as what would have been observed had we used only the Shamir action, increasing L_s from 16 to 48 (for 24I/48I) and 16 to 24 (for 32I/64I). While we believe that this assumption has a strong theoretical justification, the numerical experiment just described does not provide direct evidence for its validity.

Appendix D: Weighted fits

We define a weighted χ^2 as

$$\chi^2 = \sum_i \frac{\omega_i [y_i - f(\mathbf{x}_i, \mathbf{c})]^2}{\sigma_i^2}, \tag{D1}$$

where i indexes the measurements, y_i and σ_i are the measured value and statistical error, $\mathbf{x}_i$ the associated coordinates, and $\mathbf{c}$ the set of parameters of the fit function f. The quantities ω_i are set to a value Ω for some subset of the data, where Ω is assumed to be large, and to unity for all other data. We demonstrate below that the dependence on Ω vanishes in the limit $\Omega \to \infty$ and that this limit is sensible.

The minimum of χ^2 satisfies

$$\begin{aligned} \frac{\partial \chi^2}{\partial c_\kappa} &= \sum_i \frac{\omega_i}{\sigma_i^2} \frac{\partial \Delta_i^2(\vec{x}_i, \vec{c})}{\partial c_\kappa} \\ &= 0 . \end{aligned} \tag{D2}$$

Writing out the derivative explicitly and dividing both sides by Ω gives the following expression:

$$\sum_i \frac{\omega_i}{\Omega} \frac{\Delta_i(\vec{x}_i, \vec{c})}{\sigma_i^2} \cdot \frac{\partial f(\vec{x}_i, \vec{c})}{\partial c_\kappa} = 0 . \tag{D3}$$

If we naïvely take the $\Omega \to \infty$ limit of this equation, it appears that all of the data with $\omega_i = 1$ drop out entirely and hence do not contribute to the fit. This is certainly true in those cases in which the number of data points with weight $\omega_i = \Omega$ is sufficient to determine the full set of parameters $\vec{c}$. However when there are fewer points, there is no solution that satisfies Eq. (D3) in the $\Omega \to \infty$ limit. We argue that if one first determines the solution for *finite* Ω, either analytically or numerically, then afterwards take the limit $\Omega \to \infty$, the solution remains valid and in fact depends on the data with $\omega_i = 1$. The resolution of this apparent paradox is that when the overweighted points are insufficient to determine the parameters, the fit has (almost-)unconstrained directions with infinitesimally small curvature arising from the vanishing unweighted data, and hence there is a well defined minimum.

a. Simple example

It is straightforward to demonstrate the behavior discussed above via a simple example in which we are attempting to determine the parameters of the function

$$f(x) = a + bx \tag{D4}$$

by minimizing

$$\chi^2 = \sum_{i=0}^{N-1} (r_i - f(x_i))^2 w_i = 0\,, \tag{D5}$$

where r_i are a series of N data points with coordinates x_i and unit variances for simplicity.

Let us first consider a scenario in which we have three data points ($N = 3$), two of which are overweighted: $w_0 = w_1 = \Omega$, and the third is assigned $w_2 = 1$. Here the overweighted data points are sufficient to determine both parameters and the result of solving for the minimum of Eq. (D5) in the limit of large Ω, and the result of solving at finite Ω and taking the limit afterwards, are identical:

$$a = (r_1 x_0 - r_0 x_1)/(x_0 - x_1) \;\text{ and }\; b = (r_0 - r_1)/(x_0 - x_1)\,. \tag{D6}$$

Notice that this result does not contain the unit-weight data point, r_2.

Let us now consider just two data points ($N = 2$), and take $w_0 = \Omega$ and $w_1 = 1$ such that the number of overweighted data points is no longer sufficient to determine both parameters. The equations for the minimum of χ^2 are:

$$\begin{aligned} \frac{\partial \chi^2}{\partial a} &= -2\Omega\,(r_0 - f(x_0)) - 2\,(r_1 - f(x_1)) = 0 \qquad \text{and} \\ \frac{\partial \chi^2}{\partial b} &= -2\Omega\,(r_0 - f(x_0))\,x_0 - 2\,(r_1 - f(x_1))\,x_1 = 0\,. \end{aligned} \tag{D7}$$

Taking the large Ω limit gives

$$-2\Omega\,(r_0 - f(x_0)) = 0 \;\text{ and }\; -2\Omega\,(r_0 - f(x_0))\,x_0 = 0\,. \tag{D8}$$

These are identical up to a trivial normalization, hence we have two unknowns and only one equation; no unique solution can be found. (Note that the fact that the equations are the same will not be true in a general case with multiple over-constrained data points; there one would instead find expressions that cannot be simultaneously satisfied.) On the other hand we *can* solve for the minimum at finite Ω; the solutions are identical to those given in Equation (D6), and are *independent of* Ω, allowing us to take the large Ω limit *a posteriori* without issue.

Finally we consider one further example, again with three data points but this time with only one over-weighted: $N = 3$, $w_0 = \Omega$ and $w_1 = w_2 = 1$. Here, as above, the number of overweighted points is insufficient to determine both parameters, but all three points together are more than enough to constrain the parameters (with one degree of freedom). We might therefore expect that the solutions at finite Ω would be Ω-dependent unlike in the previous example. Indeed this is the case, but it is straightforward to show that the solutions are finite in the limit $\Omega \to \infty$ and

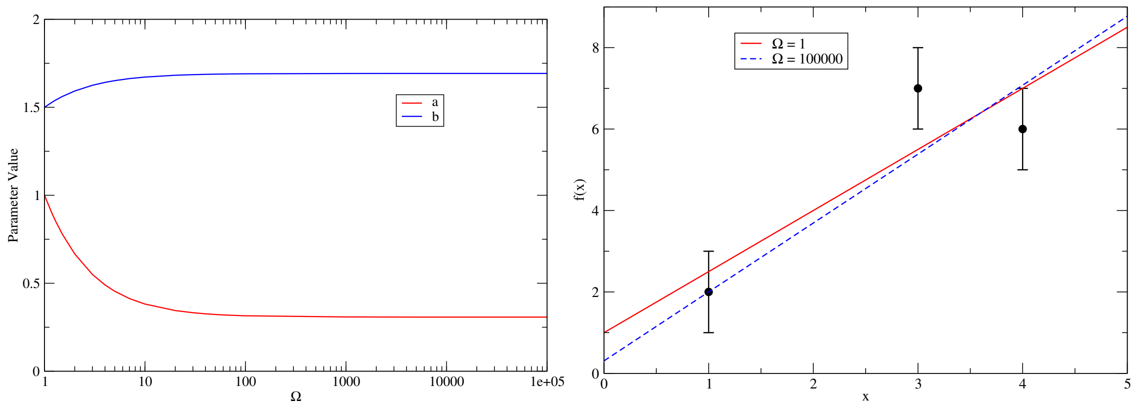


FIG. 29. (Left) The fit parameters of the function $f(x) = a + bx$ determined from arbitrarily chosen data points, $r_0(x_0 = 1) = 2$, $r_1(x_1 = 3) = 7$ and $r_2(x_2 = 4) = 6$, plotted against the weighting Ω of the first point. (Right) The fit curves with $\Omega = 1$ (red full line) and $\Omega = 100000$ (dashed blue) overlaying the data.

furthermore that they are functions of *all three data points* in this limit. The expressions are somewhat lengthy and we have not reproduced them here, but we have plotted the Ω dependence of the solutions for a particular set of data points and parameters in Figure 29. In the figure we also plot the function before and after the weighting, demonstrating that it does indeed pass through the over-weighted data point.

b. Determination of the optimal Ω value in the global fits

It remains to demonstrate the limiting behaviour in the more complex environment of the global fits. As the minimization is performed numerically via the Marquardt-Levenberg algorithm we must be careful in our choice of algorithmic parameters; the algorithm terminates when the change in χ^2 under a shift of the fit parameters is less than some chosen value, $\delta\chi^2_{\rm min}$. As we increase Ω at fixed $\delta\chi^2_{\rm min}$, the relative effects of fluctuations in the unweighted data are reduced and the fit becomes more tolerant to increasingly large deviations of the fit from the unit-weight data. This manifests as an increase in the jackknife statistical error of our predictions. We must therefore choose a value of $\delta\chi^2_{\rm min}$ that is small enough to properly take into account the constraints from the unit-weight data. The choice is limited by the increased time for the fit to reach its minimum coupled with the inevitable limits of finite precision. For fixed $\delta\chi^2_{\rm min}$, the time to perform the fit also naturally increases with Ω due to the increase in the overall scale of the fluctuations. We must therefore determine an optimal value for Ω that is large enough that our predictions are no longer

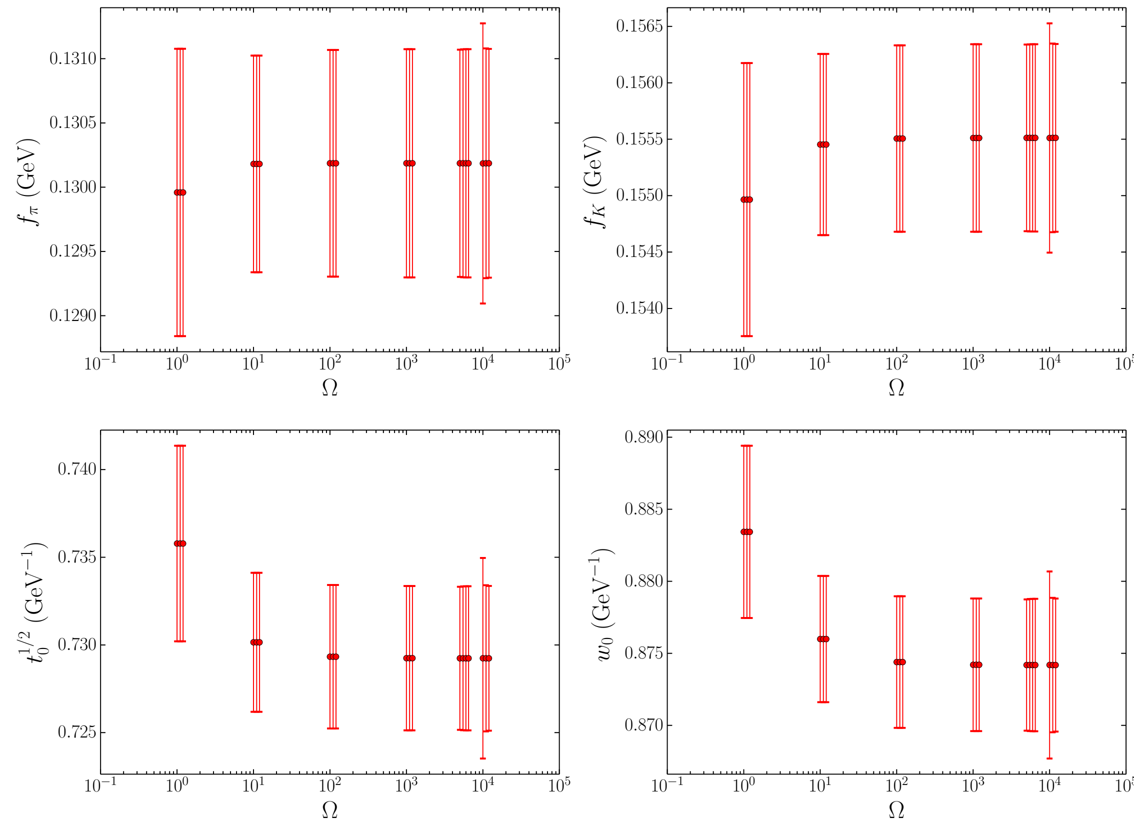


FIG. 30. Plots of the predicted continuum value for f_π (upper-left), f_K (upper-right), $t_0^{1/2}$ (lower-left) and w_0 (lower-right) as a function of the weight Ω applied to the physical point ensembles in the fit. Fits were performed with $\Omega = 1, 10, 100, 1000, 5000, 10000$ and 50000. We also considered three different values of the stopping condition $\delta\chi^2_{\rm min}$: 1×10^{-3}, 1×10^{-4} and 1×10^{-5}. For the point at $\Omega = 5000$ we also considered a fourth value, $\delta\chi^2_{\rm min} = 1\times10^{-6}$, and we only consider two values for $\Omega = 50000$ where the errors are clearly less well controlled. For each choice of Ω, the results for each value of $\delta\chi^2_{\rm min}$ have been offset for clarity, with the largest value the left-most point of each cluster, with the largest error.

noticeably dependent on its value while small enough for the fits to complete in a reasonable time and to be unaffected by finite precision errors.

In Figure 30 we show examples of the Ω dependence on the predicted values of f_π, f_K, w_0 and $t_0^{1/2}$. The plots also show the result of reducing the stopping condition $\delta\chi^2_{\rm min}$ by several orders of magnitude. We observe percent-scale shifts in the central values of these quantities from the unweighted fit results, and we clearly see the behavior flattens out at around $\Omega = 1000$. We choose $\Omega = 5000$ as a value large enough to be well within the flat region while small enough to avoid the difficulties discussed above. For the chosen value of Ω we observed no significant dependence of the results on $\delta\chi^2_{\min}$, but to be conservative we chose 1×10^{-4} as our final value. Note that we observed stronger dependence of our results on $\delta\chi^2_{\min}$ for some alternate choices of guess parameters, but with tighter stopping conditions the results stabilized and agreed with those presented in this document. To be certain, all fits presented within the body of this work were repeated with tighter stopping conditions, and no significant changes from the given values were observed.

Appendix E: Additional measurements on the 32I, 24I and 32ID ensembles

In this work we include additional data for the 24I and 32I ensembles, specifically measurements of the Wilson flow scales, $t_0^{1/2}$ and w_0, and also an improved measurement of the vector current renormalization coefficient that we use to normalize our decay constants. To remind the reader, these ensembles have lattice volumes of $24^3\times 64\times 16$ and $32^3\times 64\times 16$, and use the Shamir domain wall fermion action with the Iwasaki gauge action at bare couplings $\beta = 2.13$ and 2.25 respectively, and were originally described in Refs. [35] and [6]. We also perform measurements of the Wilson flow scales on the 32ID ensemble, which has a lattice volume of $32^3\times 64\times 32$, Shamir domain wall fermions with the Iwasaki+DSDR gauge action at $\beta = 1.75$, and was described in Ref. [5].

1. Wilson flow scales

The procedure for determining the Wilson flow scales is described in Section III F. We have three 32I ensembles with bare light quark masses of $am_l = 0.004$, 0.006 and 0.008, upon which we perform measurements using 300, 312 and 252 configurations respectively (separated by 10 MD time units) following our earlier analyses. The measurements are binned over four successive configurations to take account of autocorrelations. For the 24I ensemble set, we have two ensembles with $am_l = 0.005$ and 0.01, and we measure on 202 and 178 configurations respectively (separated by 40 MD time units) and use a bin size of 2. Finally, for the 32ID ensemble set we have two ensembles with $am_l = 0.001$ and 0.0042, and we measure on 180 and 148 configurations respectively (with 8 MD time units separation) and bin over 4 configurations. Note that the results for the 32I $m_l = 0.008$ ensemble and the 24I $m_l = 0.01$ ensemble are not included in the global fits due to the pion mass cut, but we include the results here for completeness.

Ens. set	(am_l, am_s)	$t_0^{1/2}/a$	w_0/a
32I	(0.004, 0.03)	1.7422(11)	2.0124(26)
	(0.004, 0.025)	1.7510(14)	2.0310(34)
32I	(0.006, 0.03)	1.7362(9)	1.9963(19)
	(0.006, 0.025)	1.7439(15)	2.0136(34)
32I	(0.008, 0.03)	1.7286(11)	1.9793(24)
	(0.008, 0.025)	1.7359(12)	1.9913(24)
24I	(0.005, 0.04)	1.3163(6)	1.4911(15)
	(0.005, 0.03225)	1.3237(12)	1.5071(22)
24I	(0.01, 0.04)	1.3050(7)	1.4653(14)
	(0.01, 0.03225)	1.3126(12)	1.4808(30)
32ID	(0.001, 0.045)	1.0268(3)	1.2178(7)
	(0.001, 0.04625)	1.0262(3)	1.2088(10)
32ID	(0.0042, 0.045)	1.0225(3)	1.2042(7)
	(0.0042, 0.04625)	1.0220(3)	1.2031(8)

TABLE XXIX. The Wilson flow scales in lattice units on the 32I, 24I and 32ID ensembles at the simulated strange quark mass and the reweighted mass closest to the physical value. The quark masses are given in bare lattice units.

On all three ensembles we use reweighting in the sea strange quark mass to constrain the mass dependence. The number of reweighting steps and the mass ranges used are given in the aforementioned papers. For the results presented in this section, we list only the simulated value and the closest reweighted value to the physical strange quark mass. The simulated strange quark masses are 0.03, 0.04 and 0.045 for the 32I, 24I and 32ID ensembles respectively, and the physical strange masses are as follows: $(am_s)^{32\mathrm{I}} = 0.0248(2)$, $(am_s)^{24\mathrm{I}} = 0.0322(2)$ and $(am_s)^{32\mathrm{ID}} = 0.0462(5)$. The values we obtain are given in Table XXIX.

2. Vector current renormalization

In Section III C we describe how the renormalization coefficient relating the domain wall local axial current to the physically-normalized Symanzik current can be determined via the quantity

Ens. set.	(am_l)	Z_V
24I	0.03	0.71611(8)
	0.02	0.71498(13)
	0.01	0.71409(20)
	0.005	0.71408(58)
	$-am_{\rm res}$	0.71273(26)
32I	0.008	0.74435(42)
	0.006	0.74387(55)
	0.004	0.74470(99)
	$-am_{\rm res}$	0.74404(181)

TABLE XXX. Z_V measured on the 24I and 32I ensembles, and the extrapolated value in the chiral limit.

$Z_V/Z_{\mathcal{V}}$, which relates the local vector current V_μ to the conserved 5D current $\mathcal{V}_\mu$. This quantity is used to renormalize the decay constants. In our earlier works [5, 6] we obtained Z_V by fitting directly to the ratio of two-point functions,

$$\frac{Z_V}{Z_{\mathcal{V}}} = \frac{\sum_{i=1}^{3}\sum_{\vec{x}}\langle \mathcal{V}_i^a(\vec{x},t)V_i(\vec{0},0)\rangle}{\sum_{i=1}^{3}\sum_{\vec{x}}\langle V_i^a(\vec{x},t)V_i(\vec{0},0)\rangle}. \tag{E1}$$

Since the lightest state that couples to the vector operator is the noisy ρ meson, for this work we instead determine the ratio for the 48I, 64I and 32Ifine ensembles via the three-point function, $\langle\pi|V_\mu|\pi\rangle$, as described in Section III C 2; this procedure gives a substantially more precise result than the above. In the global fits we attempt to describe the aforementioned ensembles, along with 32I and 24I ensemble sets, using the same continuum scaling trajectory. In order to guarantee consistent scaling behavior we must therefore recompute Z_V on the 32I and 24I ensemble sets using the new method. This is not necessary for the 32ID ensembles, which are described by a different scaling trajectory.

On the 24I ensemble set we measured on 147 and 153 configurations of the $am_l = 0.005$ and 0.01 ensembles respectively. We also included 85 measurements on the heavier $am_l = 0.02$ ensemble and 105 measurements on the $am_l = 0.03$ ensemble described in Ref. [35]. For the 32I ensembles we measure on 135, 152 and 120 configurations of the $am_l = 0.004$, 0.006 and 0.008 ensembles respectively. In Table XXX we list the measured values on each ensemble and extrapolated to the chiral limit.

Appendix F: Non-perturbative renormalization

In order to determine the renormalization coefficients for the quark masses and B_K, we use what is now the standard framework for our collaboration: the Rome-Southampton non-perturbative renormalization schemes [39] with momentum sources, twisted boundary conditions and non-exceptional kinematics [67–70]. This setup has already been described in several previous publications [5, 69, 71, 72], and results in tiny statistical errors, infra-red contamination suppression, and consistent removal of a^2 discretization effects in the vertex functions.

A key aspect of the RI-MOM approach is that any other, potentially regularization dependent, scheme may be easily converted into the RI-MOM scheme using momentum-space scattering amplitudes determined (either perturbatively or non-perturbatively) solely within that other scheme. This makes RI schemes a very useful intermediate scheme for converting between lattice calculations and $\overline{\text{MS}}$.

The amputated vertex functions Π_O of the operators of interest O (in this paper O represent flavour non-singlet bilinear and four-quark fermion operators) are computed on Landau gauge-fixed configurations, for which we use the timeslice by timeslice FASD algorithm [28]). We use non-exceptional 'symmetric' momentum configurations, defined by the condition

$$p_1^2 = p_2^2 = q^2 \,, \tag{F1}$$

where, for bilinear vertices, p_1 and p_2 are the incoming and outgoing quark momenta respectively, and for the four-quark vertices used to compute Z_{B_K} the quark momenta are assigned as follows: $d(p_1)\bar{s}(-p_2) \to \bar{d}(-p_1)s(p_2)$. In the above, $q = p_1 - p_2$ is the momentum transfer.

In contrast to the symmetric scheme, the original RI-MOM scheme defined in Ref. [39], which we do not include here, corresponds to the zero-momentum transfer kinematics, i.e. $p_1 = p_2$, and suffers from enhanced non-perturbative effects at high energies arising from low-momentum loop effects; in particular the effects of the dynamical chiral symmetry breaking are greatly enhanced [6].

We compute projected, amputated vertex functions of the form

$$\Lambda_O^{bare}(\mu, a) = \mathscr{P}\{\Pi_O(q^2, a)\}_{\mu^2 = q^2} \,. \tag{F2}$$

Precise definitions of the projectors $\mathscr{P}$ depend on the choice of operator, the kinematics, and the choice of scheme. In practice the Green's functions are first computed at finite values of the quark

mass and then extrapolated to the chiral limit; this quark mass dependence is however very mild for the non-exceptional schemes considered here and we omit it below for the purpose of clarity.
The renormalization factors are defined by imposing

$$\frac{Z_O}{Z_q^{n/2}}(\mu,a) \times \Lambda_O^{bare}(\mu,a) = \Lambda_O^{tree}\,. \tag{F3}$$

where Z_q is the quark wave function renormalization factor, and n the number of fermion fields in O. A second, separate condition is required in order to extract Z_q. Note that the right-hand side of the above depends on the choice of projector.
In order to simplify the equations, we introduce the following notation:

$$\bar{\Lambda}_O = \Lambda_O^{bare} \times (\Lambda_O^{tree})^{-1}\,. \tag{F4}$$

projection scheme and for each ensemble, as a function of the external momenta.
In this work we are only interested in quantities that renormalize multiplicatively, such that the Z-factors and the Λs are simply scalars. For a general lattice action with non-zero chiral symmetry breaking, the four-quark operator responsible for $K-\bar{K}$ mixing in fact mixes with other operators, and Z_O and Λ_O^{bare} become matrix-valued [71]. However for our choice of action, the residual chiral symmetry breaking is negligible and only multiplicative renormalization is required.
Once a bare matrix element $\langle O\rangle^{bare}(a)$ of the operator O has been computed on a lattice with lattice spacing a, the Z-factor can be used to convert it into the corresponding MOM-scheme:

$$\langle O\rangle^{MOM}(\mu,a) = \left(\frac{Z_O}{Z_q^{n/2}}(\mu,a)\right)^{MOM} \times \langle O\rangle^{bare}(a)\,. \tag{F5}$$

In order to connect the lattice results to phenomenology, they have to be matched to a scheme suitable for a continuum computation, such as $\overline{\mathrm{MS}}$; this is performed using perturbation theory. The final equation reads:

$$\langle O\rangle^{\overline{\mathrm{MS}}}(\mu,a) = c^{\overline{\mathrm{MS}}\leftarrow MOM}(\mu) \times \langle O\rangle^{MOM}(\mu,a) \tag{F6}$$

This quantity has a well-defined continuum limit as any potential divergences are absorbed by the Z-factors.
We remind the reader that the Z-factors defined above are scheme dependent. The renormalization scheme is fixed by the choice of projectors and of kinematics; specifically, with the choice of symmetric kinematics given above, it depends on the projector used for the operator O and that

used to extract Z_q. For both the quark mass renormalization factor, Z_m, and the B_K renormalization factor, Z_{B_K} we use two SMOM schemes; for the former these are the RI-SMOM and the RI-SMOM$_{\gamma_\mu}$ [70] schemes, and for the latter the SMOM(γ^μ, γ^μ) and SMOM$(\not{q}, \not{q})$ [40] schemes. In the main analysis we use the difference between the $\overline{\text{MS}}$ results computed using these two intermediate schemes as an estimate of the systematic error associated with the truncation of the perturbative series used to compute the SMOM$\to$ $\overline{\text{MS}}$ matching factors.

a. Renormalization of the quark masses

Our determination of the quark masses from the global fits uses an intermediate scheme that is hadronically defined and explicitly dependent on our choice of lattice regulator. The renormalization factors from bare masses to this temporary hadronic scheme are denoted Z_l and Z_h for light and strange quarks respectively. For quark masses we can convert this temporary scheme to an SMOM scheme by determining the SMOM renormalization Z_m^{SMOM} in the usual way and then determining the continuum limit of the ratio $\frac{Z_m^{RI}}{Z_{l/h}}$, and from there to $\overline{\text{MS}}$ in the usual way. This is described in more detail in Section V C.

We first introduce the renormalization factor of the flavour non-singlet bilinears. We define Λ_S and Λ_P, the amputated and projected Green's functions of the scalar and pseudoscalar bilinear operators respectively, as

$$\Lambda_S = \text{tr}\left[\Pi_S \cdot I\right], \qquad \Lambda_P = \text{tr}\left[\Pi_P \cdot \gamma_5\right]. \tag{F7}$$

Similarly, for the local vector and axial currents we define:

$$\Lambda_{V,A} = \text{tr}\left[\Pi_{V_\mu,A_\mu} \cdot \Gamma^{(s)}_{V_\mu,A_\mu}\right]. \tag{F8}$$

where (s) denotes the choice of projector. Following Ref. [70], we define the γ_μ and the $\not{q}$-schemes (or projectors) in the following way:

$$\Gamma^{(\gamma_\mu)}_{V_\mu} = \gamma_\mu, \text{ and } \Gamma^{(\gamma_\mu)}_{A_\mu} = \gamma_\mu\gamma_5, \tag{F9}$$

and

$$\Gamma^{(\not{q})}_{V_\mu} = \not{q} q_\mu/q^2, \text{ and } \Gamma^{(\not{q})}_{A_\mu} = \not{q} q_\mu \gamma_5/q^2. \tag{F10}$$

For completeness, we also renormalize the tensor current. The vertex function is $\Pi_{\sigma_{\mu\nu}}$, where

$$\sigma_{\mu\nu} = \frac{i}{2}\left[\gamma_\mu, \gamma_\nu\right], \tag{F11}$$

and the amputated and projected vertex are

$$\Lambda_T = \mathrm{tr}\left[\Pi_{\sigma_{\mu\nu}} \cdot \Gamma^{(s)}_{\sigma_{\mu\nu}}\right] . \tag{F12}$$

For the projectors, we use

$$\Gamma^{(\gamma_\mu)}_{\sigma_{\mu\nu}} \;=\; \sigma_{\mu\nu}\,, \text{ and } \Gamma^{(\not{q})}_{\sigma_{\mu\nu}} \;=\; \sigma_{\nu\rho} q_\rho q_\mu / q^2 . \tag{F13}$$

The corresponding renormalization factors $Z_{S,V,T,A,P}/Z_q$ can then obtained by imposing Eq. (F3) with $n = 2$.

To obtain the renormalization factor of the quark mass, Z_m, we take the ratio of the vector and scalar bilinears in order to cancel the quark wave-function renormalization:

$$Z^{(s)}_m(\mu, a) = \frac{\bar{\Lambda}_S(\mu, a)}{Z_V(a) \times \bar{\Lambda}^{(s)}_V(\mu, a)} , \tag{F14}$$

where Z_V is computed hadronically via the procedure given in Section III C 2. In the previous equation, we have used the fact that $Z_m = 1/Z_S = 1/Z_P$ in the chiral limit. Similarly, we should expect $Z_A = Z_V$ up to some small corrections arising, for example, from the fact that we work at finite L_s, or due to infrared contaminations. In our estimate of the systematic errors, we have also replaced Λ_S by Λ_P and Λ_V by Λ_A in Equation (F14).

b. *Renormalization of the kaon bag parameter*

The renormalization factor Z_{B_K} is defined in a similar manner. The amputated Green's function of the relevant four-quark operator $\mathcal{O}_{VV+AA}$ describing $K - \bar{K}$ oscillations in the Standard Model is computed numerically with a certain choice of kinematics and projected onto its tree-level value. We normalize by the square of the average between the vector and axial bilinear:

$$Z^{(s_1,s_2)}_{B_K}(\mu, a) \times \frac{\bar{\Lambda}^{(s_1)}_{VV+AA}(\mu, a)}{\bar{\Lambda}^{(s_2)}_{AV}(\mu, a)^2} = 1 , \tag{F15}$$

where

$$\Lambda_{AV} = \frac{1}{2}(\Lambda_V + \Lambda_A) , \tag{F16}$$

such that the quark field renormalization cancels in the ratio. In practice we find that the difference between the vector and axial vertices are very small, hence choosing the average rather than simply Λ_V or Λ_A in the denominator, has no discernable effect.

In Eq. (F15), the superscripts s_1 and s_2 label the choice of projectors. We refer the reader to Refs. [5, 40] for the details on the implementation, including the explicit definitions of projectors.

1. Numerical details and discussion

a. *Quark mass renormalization*

For the quark mass renormalization we require only the values on the 32I and 24I ensembles, which together are sufficient to perform the continuum extrapolation of $Z_m/Z_{l/h}$. Here we discuss an update of the analysis performed in Ref. [5] using the newly-determined lattice spacings and a number of additional data points.

In the Rome-Southampton method, the projected vertex functions are first computed at finite quark mass before being extrapolated to the chiral limit. For each ensemble, we use unitary valence quark masses and extrapolate linearly in the quark mass. In the sea sector, the strange quark mass remains fixed to - or close to - its physical value. Since we do not observe any relevant quark mass dependence in our data, we neglect the systematic error associated with the fact that the sea strange quark mass is not extrapolated to zero.

We use partially-twisted boundary conditions to obtain momenta of the following form:

$$p_{in} = \frac{2\pi}{L}(-\tilde{m}, 0, \tilde{m}, 0)\,, \tag{F17}$$

$$p_{out} = \frac{2\pi}{L}(0, \tilde{m}, \tilde{m}, 0)\,, \tag{F18}$$

where $\tilde{m}$ combines the Fourier mode with the twist angle θ

$$\tilde{m} = m + \theta/2\,, \quad m \in \mathbb{N}\,. \tag{F19}$$

The fact that these momenta all point in the same direction up to hypercubic rotations means that they lie upon a common continuum scaling curve (i.e. their a^2 dependence is the same), allowing us to unambiguously take the continuum limit.

For the 24I lattice, in addition to the momenta listed in [5], we have generated additional points closer to the 3 GeV point at which we ultimately evaluate the Z-factors. More precisely, the twist angle θ is chosen to be $n \times 3/16$, with $n = 15, 16, \ldots 19$. The results can be found in the following section.

b. *Renormalization of B_K*

As B_K is a scheme dependent quantity, we must perform our global fits to renormalized data, and as a result we require values of the renormalization coefficients to be computed on all of the ensembles used in the analysis: the 32I, 24I, 48I, 64I, 32Ifine and 32ID. This differs from the quark mass determination, for which we used a hadronically defined intermediate scheme during the continuum extrapolation and converted to $\overline{\mathrm{MS}}$ *a posteriori*. In this appendix we present updated values of the 32I, 24I and 32ID Z_{B_K} results in Ref. [5], as well as new values for the 48I, 64I and 32Ifine.

For our new ensembles, we have considered only one value of the valence quark mass $m_l^{sea} = m_l^{val}$. Again, due to the modest chiral dependence previously observed for the non-exceptional schemes, we expect the associated systematic error to be negligible compare to the other sources of errors (in particular the perturbative matching).

As the 32ID ensemble is comparatively coarse, we renormalize at a lower scale $\mu_0 \sim 1.4$ GeV and use the non-perturbative continuum step-scaling factor $\sigma_{B_K}^{(s_1,s_2)}(\mu,\mu_0)$ to run to 3 GeV. This procedure is discussed in Ref. [5]. The step-scaling factor is obtained by performing a continuum extrapolation of the ratio

$$\sigma_{B_K}^{(s_1,s_2)}(\mu,\mu_0;a) = Z_{B_K}^{(s_1,s_2)}(\mu;a)/Z_{B_K}^{(s_1,s_2)}(\mu_0;a)\,, \tag{F20}$$

computed on the 32I and 24I lattices.

Since the values of the lattice spacings have been updated, the numbers quoted here differ slightly from our previous work. The strategy is the following: we use the same 32ID lattice renormalization coefficient, $Z_{B_K}^{(s_1,s_2)}(\mu_0,a_{32\mathrm{ID}})$, as used previously, but notice that the corresponding value of μ_0 obtained with the new lattice spacings is 1.4363 GeV rather than 1.426 GeV. As a result we must recompute the step-scaling factor. The results for $Z_{B_K}^{(s_1,s_2)}$ at μ_0 can be found in Table XLII and our updated results for $\sigma_{B_K}^{(s_1,s_2)}$ are reported in Table XLIII. For each scheme, the 32ID renormalization factor evaluated at $\mu = 3$ GeV is then simply given by

$$Z_{B_K}^{(s_1,s_2)}(\mu,a_{32\mathrm{ID}}) = \sigma^{(s_1,s_2)}(\mu,\mu_0) \times Z_{B_K}^{(s_1,s_2)}(\mu_0,a_{32\mathrm{ID}})\,. \tag{F21}$$

2. Numerical results

a. Bilinears and quark mass renormalization

The values for the amputated vertex functions $\bar{\Lambda}$ (normalized by the tree level value) at finite quark mass and in the chiral limit computed on the 24I ensemble are given in Tables XXXI and XXXII for the SMOM$_{\gamma^\mu}$ and SMOM schemes respectively. The corresponding numbers for the 32I ensembles are given in Tables XXXIII and XXXIV. Recall that we use only one choice of projector

$am = 0.01$					
q/GeV	2.911997	2.973955	3.035912	3.097870	3.159827
S	1.1492(3)	1.1455(3)	1.1422(3)	1.1390(2)	1.1360(2)
V	1.0530(2)	1.0537(2)	1.0543(2)	1.0550(2)	1.0557(2)
T	1.0225(2)	1.0244(2)	1.0263(2)	1.0281(2)	1.0299(2)
A	1.0527(2)	1.0534(2)	1.0541(2)	1.0548(2)	1.0556(2)
P	1.1520(3)	1.1480(3)	1.1444(3)	1.1409(2)	1.1377(2)
$am = 0.005$					
q/GeV	2.911997	2.973955	3.035912	3.097870	3.159827
S	1.1491(2)	1.1455(2)	1.1421(1)	1.1390(1)	1.1360(1)
V	1.0529(1)	1.0536(1)	1.0542(1)	1.0549(1)	1.0556(1)
T	1.0225(2)	1.0244(2)	1.0262(1)	1.0281(1)	1.0299(1)
A	1.0528(1)	1.0534(1)	1.0541(1)	1.0548(1)	1.0556(1)
P	1.1517(2)	1.1478(2)	1.1441(2)	1.1407(2)	1.1375(2)
$am = -am_{\mathrm{res}}$					
q/GeV	2.911997	2.973955	3.035912	3.097870	3.159827
S	1.1491(7)	1.1455(6)	1.1421(6)	1.1389(5)	1.1359(5)
V	1.0527(5)	1.0534(4)	1.0540(4)	1.0547(4)	1.0555(4)
T	1.0223(5)	1.0243(5)	1.0261(5)	1.0280(5)	1.0299(5)
A	1.0528(4)	1.0535(4)	1.0542(4)	1.0549(4)	1.0556(4)
P	1.1512(7)	1.1473(7)	1.1436(7)	1.1402(6)	1.1370(6)

TABLE XXXI. Projected, amputated vertex functions $\bar{\Lambda}$ for the vector, axial-vector and tensor operators in the SMOM$_{\gamma^\mu}$ scheme computed on the two 24I ensembles, and in the chiral limit, at scales close to the chosen renormalization scale of 3 GeV. In this table we also include the projected, amputated scalar and pseudoscalar vertices.

for the scalar and pseudoscalar vertices, specifically those given in Eq. (F7). The results for these vertices computed on the 24I and 32I ensembles are included in Tables XXXI and XXXIII respectively.

In Table XXXV we present $\bar{\Lambda}$ interpolated to 3 GeV using a polynomial ansatz in the momenta. For the 24I lattice, since we have a very fine resolution, we take the five momenta quoted in the

$am = 0.01$					
q/GeV	2.911997	2.973955	3.035912	3.097870	3.159827
V	1.1159(4)	1.1160(3)	1.1163(3)	1.1166(3)	1.1171(3)
T	1.0225(2)	1.0244(2)	1.0263(2)	1.0281(2)	1.0299(2)
A	1.1156(3)	1.1158(3)	1.1160(3)	1.1164(3)	1.1169(3)
$am = 0.005$					
q/GeV	2.911997	2.973955	3.035912	3.097870	3.159827
V	1.1158(3)	1.1159(3)	1.1162(2)	1.1165(2)	1.1169(2)
T	1.0225(2)	1.0244(2)	1.0262(2)	1.0281(1)	1.0299(1)
A	1.1156(3)	1.1158(2)	1.1160(2)	1.1163(2)	1.1167(2)
$am = -am_{\mathrm{res}}$					
q/GeV	2.911997	2.973955	3.035912	3.097870	3.159827
V	1.1156(9)	1.1158(9)	1.1159(8)	1.1161(8)	1.1164(8)
T	1.0224(5)	1.0243(5)	1.0262(5)	1.0280(5)	1.0299(5)
A	1.1156(9)	1.1158(8)	1.1160(8)	1.1162(7)	1.1165(7)

TABLE XXXII. Projected, amputated vertex functions $\bar{\Lambda}$ in the SMOM scheme computed on the two 24I ensembles, and in the chiral limit, at scales close to the chosen renormalization scale of 3 GeV.

tables. For the 32I results we use $q \sim 2.77, 3.10$ and 3.43 GeV in the interpolation.

We show the values of the quark mass renormalization in Table XXXVI. Using Table XXXV we can gauge the size of the systematic error on Z_m by comparing the S and P vertices and the A and V vertices. We observe that the differences between the vector and axial vector vertices are very small, and can therefore be neglected. The differences between the scalar and pseudoscalar vertices are slightly larger, but these correspond to only 0.01% changes if used in the computation of Z_m, and can therefore be ignored. As discussed above, the systematic error associated with not taking the chiral extrapolation of the sea strange quark mass can also be ignored. Note that the uncertainties on the lattice spacings are incorporated in these quantities in the main analysis.

$am = 0.008$							
q/GeV	1.186382	1.581953	2.067155	2.397900	2.769988	3.100733	3.431478
S	1.5760(95)	1.4124(23)	1.2881(7)	1.2346(4)	1.1920(2)	1.1648(2)	1.1446(1)
V	1.0568(13)	1.0425(4)	1.0376(1)	1.0368(1)	1.0374(1)	1.0387(0)	1.0405(0)
T	0.9072(10)	0.9403(3)	0.9668(1)	0.9796(1)	0.9915(1)	1.0005(0)	1.0083(0)
A	1.0357(9)	1.0369(4)	1.0364(1)	1.0362(1)	1.0371(1)	1.0385(0)	1.0404(0)
P	1.8453(92)	1.4853(22)	1.3065(9)	1.2425(5)	1.1956(3)	1.1665(2)	1.1457(2)
$am = 0.006$							
q/GeV	1.186382	1.581953	2.067155	2.397900	2.769988	3.100733	3.431478
S	1.5818(54)	1.4178(29)	1.2906(10)	1.2358(6)	1.1930(3)	1.1656(2)	1.1451(1)
V	1.0544(7)	1.0413(4)	1.0376(2)	1.0370(2)	1.0376(1)	1.0388(1)	1.0406(1)
T	0.9081(5)	0.9402(3)	0.9669(3)	0.9798(2)	0.9917(1)	1.0006(1)	1.0084(1)
A	1.0357(8)	1.0374(3)	1.0367(2)	1.0366(2)	1.0374(1)	1.0387(1)	1.0405(1)
P	1.8124(66)	1.4745(23)	1.3048(6)	1.2419(6)	1.1956(3)	1.1669(2)	1.1458(2)
$am = 0.004$							
q/GeV	1.186382	1.581953	2.067155	2.397900	2.769988	3.100733	3.431478
S	1.5697(61)	1.4163(24)	1.2915(6)	1.2363(3)	1.1927(2)	1.1653(1)	1.1448(1)
V	1.0542(10)	1.0418(2)	1.0373(2)	1.0368(1)	1.0374(1)	1.0387(1)	1.0405(1)
T	0.9078(8)	0.9404(2)	0.9668(2)	0.9798(1)	0.9917(1)	1.0005(1)	1.0083(1)
A	1.0396(8)	1.0383(2)	1.0366(2)	1.0365(1)	1.0373(1)	1.0386(1)	1.0404(1)
P	1.8346(113)	1.4761(21)	1.3050(9)	1.2414(4)	1.1948(3)	1.1663(1)	1.1455(1)
$am = -am_{\mathrm{res}}$							
q/GeV	1.186382	1.581953	2.067155	2.397900	2.769988	3.100733	3.431478
S	1.5605(173)	1.4219(57)	1.2955(14)	1.2384(7)	1.1934(4)	1.1659(3)	1.1453(3)
V	1.0510(25)	1.0413(6)	1.0369(3)	1.0369(3)	1.0375(2)	1.0388(2)	1.0405(1)
T	0.9085(18)	0.9406(6)	0.9668(4)	0.9799(3)	0.9918(2)	1.0006(2)	1.0084(1)
A	1.0438(18)	1.0400(5)	1.0368(4)	1.0368(3)	1.0376(2)	1.0388(1)	1.0405(1)
P	1.7969(254)	1.4637(53)	1.3030(22)	1.2400(11)	1.1940(7)	1.1659(4)	1.1452(3)

TABLE XXXIII. Projected, amputated vertex functions $\bar{\Lambda}$ for the vector, axial-vector and tensor operators in the SMOM$_{\gamma^\mu}$ scheme computed on the three 32I ensembles, and in the chiral limit, at scales close to the chosen renormalization scale of 3 GeV. In this table we also include the projected, amputated scalar and pseudoscalar vertices.

$am = 0.008$							
q/GeV	1.186382	1.581953	2.067155	2.397900	2.769988	3.100733	3.431478
V	1.1756(30)	1.1387(12)	1.1145(4)	1.1047(2)	1.0979(2)	1.0945(1)	1.0928(1)
T	0.9077(10)	0.9406(3)	0.9668(1)	0.9796(1)	0.9916(1)	1.0005(0)	1.0083(0)
A	1.1659(27)	1.1359(12)	1.1137(4)	1.1043(2)	1.0977(2)	1.0943(1)	1.0927(1)
$am = 0.006$							
q/GeV	1.186382	1.581953	2.067155	2.397900	2.769988	3.100733	3.431478
V	1.1735(15)	1.1367(10)	1.1147(6)	1.1049(4)	1.0981(3)	1.0946(3)	1.0929(3)
T	0.9089(6)	0.9405(3)	0.9669(3)	0.9798(2)	0.9917(1)	1.0006(1)	1.0084(1)
A	1.1661(15)	1.1347(10)	1.1142(6)	1.1046(4)	1.0979(3)	1.0945(3)	1.0928(3)
$am = 0.004$							
q/GeV	1.186382	1.581953	2.067155	2.397900	2.769988	3.100733	3.431478
V	1.1760(19)	1.1377(8)	1.1138(4)	1.1045(2)	1.0979(2)	1.0944(1)	1.0927(1)
T	0.9086(7)	0.9408(2)	0.9667(2)	0.9798(1)	0.9917(1)	1.0005(1)	1.0083(1)
A	1.1713(20)	1.1365(8)	1.1134(4)	1.1043(2)	1.0978(2)	1.0943(1)	1.0926(1)
$am = -am_{\text{res}}$							
q/GeV	1.186382	1.581953	2.067155	2.397900	2.769988	3.100733	3.431478
V	1.1769(54)	1.1368(23)	1.1131(9)	1.1042(5)	1.0979(4)	1.0944(3)	1.0926(3)
T	0.9097(17)	0.9411(6)	0.9667(3)	0.9799(3)	0.9918(2)	1.0006(2)	1.0084(1)
A	1.1777(53)	1.1372(23)	1.1132(9)	1.1043(5)	1.0980(4)	1.0944(3)	1.0926(3)

TABLE XXXIV. Projected, amputated vertex functions $\bar{\Lambda}$ in the SMOM scheme computed on the three 32I ensembles, and in the chiral limit, at scales close to the chosen renormalization scale of 3 GeV.

Lattice	Scheme	S	V	T	A	P
24I	γ^μ	1.1441(6)	1.0536(4)	1.0251(5)	1.0538(4)	1.1457(7)
	$\not{q}$	-	1.1158(8)	1.0251(5)	1.1159(8)	-
32I	γ^μ	1.1736(4)	1.0383(2)	0.9981(2)	1.0383(2)	1.1737(5)
	$\not{q}$	-	1.0957(3)	0.9978(2)	1.0958(3)	-

TABLE XXXV. The bilinear amputated, projected vertex functions $\bar{\Lambda}$ interpolated to 3 GeV. Note that these errors do not include the lattice spacing uncertainty.

	24I	32I
γ_μ	1.523(1)	1.519(4)
$\not{q}$	1.439(1)	1.440(4)

TABLE XXXVI. Quark mass renormalization factors $Z_m^{(s)}(3\text{ GeV}, \mathrm{a})$ computed on the 24I and 32I lattices at 3 GeV in the two SMOM-schemes. Note that these errors do not include the lattice spacing uncertainty.

b. Renormalization of B_K

We quote the results for the projected vertex function $\bar{\Lambda}_{VV+AA}$ at finite masses and in the chiral limit for various momenta on the 32I and 24I ensembles in Tables XXXVII and XXXVIII. The corresponding values each computed at a single quark mass on the 48I, 64I and 32Ifine ensembles are given in Tables XXXIX, XL and XLI respectively.

To obtain the final results we construct the ratio given in Equation (92) at finite quark masses for a few momenta surround the desired scale, either $\mu_0 = 1.4363$ GeV or $\mu = 3$ GeV, take the chiral limit and then perform the interpolation with a polynomial ansatz. Similarly to the quark mass case, the procedure is very robust and does not depend on the order we perform these operations, nor on the details of the interpolation. The final results for Z_{B_K} on the various ensembles are given in Table XLII, and the continuum step-scaling factors used to run the 32ID renormalization factor to 3 GeV are quoted in Table XLIII.

As with the quark mass renormalization the only significant source of systematic error on these results arises from the perturbative matching to $\overline{\text{MS}}$, which we estimate using the full difference between our final predictions for B_K determined via the two intermediate SMOM schemes. As above, we incorporate the uncertainties on the lattice spacings into our renormalization factors in the main analysis.

Appendix G: Random number generator

After all the data presented in this paper was generated, it was found that the U(1) noise generated from the freshly initialized random number generator (RNG) in CPS [73] is vulnerable to correlations, such that the expectation value of $\left\|\sum_x \sum_{j=1}^{N} e^{-i\theta(x)_j}\right\|^2/V$ deviates from N. This correlation is not observed when U(1) noise is replaced with gaussian noise, for which the accept/reject procedure used in generating the gaussian random numbers appears to eliminate the

(γ^μ,γ^μ) scheme, lowest momenta				
q/GeV	1.172282	1.563201	1.858201	1.920141
$am=0.01$	1.1453(14)	1.1617(9)	1.1702(5)	1.1722(5)
$am=0.005$	1.1458(16)	1.1600(8)	1.1688(4)	1.1708(4)
$am=-am_{\text{res}}$	1.1466(48)	1.1574(26)	1.1665(14)	1.1685(13)
(γ^μ,γ^μ) scheme, highest momenta				
q/GeV	2.973122	3.035062	3.097002	3.158942
$am=0.01$	1.2116(5)	1.2145(5)	1.2174(5)	1.2205(5)
$am=0.005$	1.2113(3)	1.2142(3)	1.2171(3)	1.2202(3)
$am=-am_{\text{res}}$	1.2108(13)	1.2137(13)	1.2167(14)	1.2197(14)
$(\not{q},\not{q})$ scheme, lowest momenta				
q/GeV	1.172282	1.563201	1.858201	1.920141
$am=0.01$	1.3017(25)	1.2921(14)	1.2851(10)	1.2846(10)
$am=0.005$	1.2996(23)	1.2876(15)	1.2825(8)	1.2821(8)
$am=-am_{\text{res}}$	1.2962(64)	1.2803(49)	1.2782(30)	1.2779(27)
$(\not{q},\not{q})$ scheme, highest momenta				
q/GeV	2.973122	3.035062	3.097002	3.158942
$am=0.01$	1.3012(7)	1.3037(7)	1.3064(8)	1.3092(8)
$am=0.005$	1.3008(5)	1.3033(5)	1.3059(4)	1.3087(4)
$am=-am_{\text{res}}$	1.3003(18)	1.3027(18)	1.3052(18)	1.3078(18)

TABLE XXXVII. Chiral extrapolation of $\bar{\Lambda}_{VV+AA}$ in both schemes on the 24I ensemble for the momentum points in the vicinity of the 1.4 GeV scale, and those in the vicinity of the 3 GeV matching scale.

observed correlation. We also confirmed that the U(1) noise generated from the CPS RNG for the thermalized gauge configurations on our previous ensembles do not show the correlation, due to the de-correlating effect of the gaussian RNG used for the pseudofermion fields.

To further test the robustness of gaussian random numbers generated from CPS RNG, we reproduced the 2+1 flavor DWF ensemble used in Ref. [74], with the RNG's replaced with the Mersenne Twister [75], implemented in C++11. Each 2^4 hypercube of lattice sites was initialized with randomized seeds. We confirmed that the plaquette agrees to within 1 standard devia-

(γ^μ, γ^μ) scheme							
q/GeV	1.186382	1.581953	2.067155	2.397900	2.769988	3.100733	3.431478
$am = 0.008$	1.0985(21)	1.1117(11)	1.1240(2)	1.1311(3)	1.1399(2)	1.1483(2)	1.1573(1)
$am = 0.006$	1.0991(20)	1.1111(6)	1.1246(5)	1.1318(4)	1.1404(3)	1.1487(3)	1.1577(3)
$am = 0.004$	1.1008(16)	1.1120(2)	1.1236(5)	1.1312(4)	1.1401(3)	1.1483(2)	1.1573(2)
$am = -am_{\text{res}}$	1.1034(43)	1.1129(11)	1.1237(13)	1.1316(9)	1.1404(7)	1.1485(6)	1.1574(5)
$(\not{q}, \not{q})$ scheme							
q/GeV	1.186382	1.581953	2.067155	2.397900	2.769988	3.100733	3.431478
$am = 0.008$	1.2473(40)	1.2352(21)	1.2262(5)	1.2233(3)	1.2238(3)	1.2266(2)	1.2316(2)
$am = 0.006$	1.2471(25)	1.2334(13)	1.2271(10)	1.2242(7)	1.2243(6)	1.2271(5)	1.2321(5)
$am = 0.004$	1.2515(28)	1.2345(13)	1.2253(9)	1.2231(4)	1.2239(4)	1.2266(3)	1.2316(2)
$am = -am_{\text{res}}$	1.2566(77)	1.2341(38)	1.2251(22)	1.2231(12)	1.2241(10)	1.2267(8)	1.2316(7)

TABLE XXXVIII. Chiral extrapolation of $\bar{\Lambda}_{VV+AA}$ in both schemes on the 32I ensemble for all simulated momenta.

tion: 0.588064(12) from 8460 MD units compared to 0.588052(9) from the configurations used in Ref. [74]. All the random numbers generated from CPS RNG for the work presented here were gaussian random numbers. The only exception are the Z(3) random numbers for Z_3 box source used for the Ω baryon in Section III E, which was generated independently from the CPS RNG.

(γ^μ,γ^μ) scheme					
q/GeV	2.72125	2.88132	2.96136	3.04139	3.20147
VV+AA	1.20472(14)	1.21216(8)	1.21604(8)	1.21996(8)	1.22827(6)
V	1.05201(5)	1.05371(3)	1.05463(4)	1.05557(3)	1.05753(1)
A	1.05196(3)	1.05368(2)	1.05458(4)	1.05553(3)	1.05745(4)
$(\not{q},\not{q})$ scheme					
q/GeV	2.72125	2.88132	2.96136	3.04139	3.20147
VV+AA	1.29658(31)	1.30250(14)	1.30598(10)	1.30955(10)	1.31773(25)
V	1.11640(15)	1.11660(7)	1.11697(4)	1.11749(5)	1.11902(15)
A	1.11633(13)	1.11659(5)	1.11695(4)	1.11747(5)	1.11902(14)

TABLE XXXIX. Vertex functions of the four-quark operators $\bar{\Lambda}_{VV+AA}$ and the bilinears $\bar{\Lambda}_V$ and $\bar{\Lambda}_A$ needed for Z_{B_K}, computed in both schemes on the 48I ensemble with $am = 0.00078$.

(γ^μ,γ^μ) scheme					
q/GeV	2.7823	2.94596	3.0278	3.10963	3.27329
VV+AA	1.13936(9)	1.14363(10)	1.14575(6)	1.14798(6)	1.15261(4)
V	1.03721(4)	1.03783(3)	1.03818(2)	1.03859(2)	1.03949(2)
A	1.03715(2)	1.03780(3)	1.03815(2)	1.03856(2)	1.03949(1)
$(\not{q},\not{q})$ scheme					
q/GeV	2.7823	2.94596	3.0278	3.10963	3.27329
VV+AA	1.22136(20)	1.22299(22)	1.22387(12)	1.22501(11)	1.22760(12)
V	1.09622(9)	1.09451(11)	1.09379(6)	1.09323(5)	1.09239(6)
A	1.09619(9)	1.09449(11)	1.09377(6)	1.09321(5)	1.09238(7)

TABLE XL. Vertex functions of the four-quark operators $\bar{\Lambda}_{VV+AA}$ and the bilinears $\bar{\Lambda}_V$ and $\bar{\Lambda}_A$ needed for Z_{B_K}, computed in both schemes on the 64I ensemble with $am = 0.000678$.

(γ^μ,γ^μ) scheme						
q/GeV	2.18326	2.61991	3.05656	3.49322	3.92987	4.36652
VV+AA	1.08793(30)	1.09602(53)	1.10356(30)	1.11137(28)	1.11909(12)	1.12751(11)
V	1.03160(12)	1.03005(15)	1.02984(8)	1.03049(8)	1.03162(5)	1.03327(4)
A	1.03076(11)	1.02976(17)	1.02972(8)	1.03042(8)	1.03159(5)	1.03324(5)
$(\not{q},\not{q})$ scheme						
q/GeV	2.18326	2.61991	3.05656	3.49322	3.92987	4.36652
VV+AA	1.18143(36)	1.17928(112)	1.17863(65)	1.18050(63)	1.18350(28)	1.18855(25)
V	1.10353(55)	1.09220(57)	1.08465(34)	1.08013(31)	1.07727(14)	1.07605(12)
A	1.10308(55)	1.09196(57)	1.08453(34)	1.08005(31)	1.07723(14)	1.07601(12)

TABLE XLI. Vertex functions of the four-quark operators $\bar{\Lambda}_{VV+AA}$ and the bilinears $\bar{\Lambda}_V$ and $\bar{\Lambda}_A$ needed for Z_{B_K}, computed in both schemes on the 32Ifine ensemble with $am = 0.0047$.

$Z_{B_K}^{(s_1,s_2)}(3\ \mathrm{GeV},\mathrm{a})$						
	24I	32I	48I	64I	32Ifine	32ID
(γ_μ,γ_μ)	0.9161(5)	0.9409(2)	0.91397(3)	0.94106(2)	0.9617(1)	-
$(\not{q},\not{q})$	0.9568(2)	0.9787(1)	0.954452(4)	0.978152(2)	0.9995(1)	-
$Z_{B_K}^{(s_1,s_2)}(1.4363\ \mathrm{GeV},\mathrm{a})$						
(γ_μ,γ_μ)	0.9546(10)	0.9809(93)	-	-	-	0.9210(8)
$(\not{q},\not{q})$	1.0488(16)	1.0638(20)	-	-	-	0.9992(11)

TABLE XLII. B_K renormalization factors $Z_{B_K}^{(s_1,s_2)}$ computed on the various ensembles. For the 32I, 24I we quote values at both 1.4363 GeV and 3 GeV, which are used to compute the step-scaling factor. For the coarse 32ID ensemble we only quote the value at the lower scale, and for the 48I, 64I and 32Ifine we do not quote the values at the lower scale as they are not needed for our analysis. These values do not include the effect of the uncertainty on the lattice spacing in their errors.

(γ_μ, γ_μ)	0.9573(21)
$(\not{q}, \not{q})$	0.9103(31)

TABLE XLIII. Continuum non-perturbative scale evolution $\sigma_{B_K}^{(s_1,s_2)}(\mu, \mu_0)$ extracted from the 24I and 32I lattices in two SMOM-schemes. As explained in the text, we choose $\mu_0 = 1.4363$ GeV and $\mu = 3$ GeV. These values do not include the effect of the uncertainty on the lattice spacing in their errors.

Appendix H: Global fit forms

The ChPT forms and their associated finite-volume corrections were originally determined in Ref. [35] and the analytic forms in Refs. [6, 40]. We have subsequently [5, 6, 40] added additional terms describing the scaling behavior and the dependence of the quantities on the heavy sea and valence quark masses where appropriate. In this analysis we also introduce linear fit forms to describe the Wilson flow scales. For the convenience of the reader we collect these disparate formulae below.

The ChPT forms for the pseudoscalar mass and decay constant are [5, 6, 35]:

$$m_{xy}^2 = \frac{\chi_x+\chi_y}{2}\left[1+L^{m_\pi}(\chi_x,\chi_y,\chi_l)\right]+c_{m_\pi,m_h}\frac{m_x+m_y}{2}(m_h-m_h^{\rm phys})\,, \tag{H1}$$

$$f_{xy} = f\left[1+c_f^{\mathbf{A}}a^2+L^{f_\pi}(\chi_x,\chi_y,\chi_l)\right]+c_{f_\pi,m_h}(m_h-m_h^{\rm phys})\,. \tag{H2}$$

Here m_x and m_y are the (partially-quenched) valence light quark masses, m_l is the sea light quark mass and m_h the sea heavy quark mass. The quantity $\chi_x = 2Bm_x$, and the superscript **A** above the a^2 coefficient denotes the gauge action. We use the following notation for the gauge actions: **I** for the Iwasaki action and **ID** for the Iwasaki+DSDR. The logarithmic terms L^{m_π} are defined in Eq. B32 and B33 of Ref. [35] for non-degenerate and degenerate valence quark masses, respectively. Similarly, L^{f_π} are given in Eqs. B36 and B37 of the same document. For the kaon mass, decay constant and bag parameter we use the following forms [5, 6, 35, 40]:

$$m_{xy}^2 = m^{(K)}\left[1+\frac{\lambda_1\chi_l}{f^2}+\frac{\lambda_2\chi_x}{f^2}\right]+c_{m_K,m_y}(m_y-m_h^{\rm phys})+c_{m_K,m_h}(m_h-m_h^{\rm phys})\,, \tag{H3}$$

$$f_{xy} = f^{(K)}\left[1+c_{f_K,a}^{\mathbf{A}}a^2+\frac{\lambda_3\chi_l}{f^2}+\frac{\lambda_4\chi_x}{f^2}+L^{f_K}(\chi_x,\chi_l)\right]+c_{f_K,m_y}(m_y-m_h^{\rm phys})+c_{f_K,m_h}(m_h-m_h^{\rm phys})\,, \tag{H4}$$

$$B_{xy} = B_K^0\left[1+c_{B_K,a}^{\mathbf{A}}a^2+\frac{c_{B_K,m_l}\chi_l}{f^2}+\frac{c_{B_K,m_x}\chi_x}{f^2}-\frac{\chi_l}{32\pi^2f^2}\log\left(\frac{\chi_x}{\Lambda_\chi^2}\right)\right]+ c_{B_K,m_y}(m_y-m_h^{\rm phys})+c_{B_K,m_h}(m_h-m_h^{\rm phys})\,, \tag{H5}$$

where m_y and m_x are the heavy and light valence quark masses, respectively, and m_l and m_h are as above. Here the logarithmic term L^{f_K} is defined in Eq. B47 of Ref. [35]. For the Omega baryon mass we use

$$m_{vvv} = m^{(\Omega)}+c_{m_\Omega,l}m_l+c_{m_\Omega,v}(m_v-m_h^{\rm phys})+c_{m_\Omega,v}(m_h-m_h^{\rm phys})\,, \tag{H6}$$

where m_v is the valence heavy quark mass.

The analytic forms for the pseudoscalar mass and decay constant are [5, 6]

$$m_{xy}^2 = C_0^{m_\pi} + C_1(m_x + m_y)/2 + C_2 m_l + C_3(m_h - m_h^{\text{phys}}) \tag{H7}$$

$$f_{xy} = C_0^{f_\pi}(1 + C_a^{f_\pi,\mathbf{A}} a^2) + C_1(m_x + m_y)/2 + C_2 m_l + C_3(m_h - m_h^{\text{phys}}) \tag{H8}$$

where again, m_x and m_y are the valence light quark masses, and m_l and m_h are the sea light and heavy quark masses. For the kaon mass, decay constant and bag parameter [5, 6, 40],

$$m_{xy}^2 = C_0^{m_K} + C_1^{m_K} m_x + C_2^{m_K} m_l + C_3^{m_K}(m_y - m_h^{\text{phys}}) + C_4^{m_K}(m_h - m_h^{\text{phys}}), \tag{H9}$$

$$f_{xy} = C_0^{f_K}(1 + C_a^{f_K,\mathbf{A}} a^2) + C_1^{f_K} m_x + C_2^{f_K} m_l + C_3^{f_K}(m_y - m_h^{\text{phys}}) + C_4^{f_K}(m_h - m_h^{\text{phys}}), \tag{H10}$$

$$B_{xy} = C_0^{B_K}(1 + C_a^{B_K,\mathbf{A}} a^2) + C_1^{B_K} m_x + C_2^{B_K} m_l + C_3^{B_K}(m_y - m_h^{\text{phys}}) + C_4^{B_K}(m_h - m_h^{\text{phys}}), \tag{H11}$$

where, as before, m_y represents the heavy valence quark. Finally the analytic function for the Omega baryon mass is

$$m_{vvv} = C_0^{m_\Omega} + C_1^{m_\Omega} m_l + C_2^{m_\Omega}(m_v - m_h^{\text{phys}}) + C_3^{m_\Omega}(m_h - m_h^{\text{phys}}), \tag{H12}$$

where again m_v is the valence heavy quark mass. In general, the coefficients for these analytic functions are ordered as follows (skipping entries as appropriate): The valence light quark mass dependence; the sea light quark mass dependence; the valence heavy quark mass dependence; and the sea heavy quark mass dependence.

For this analysis we also define the following functions for the Wilson flow scales $t_0^{1/2}$ and w_0:

$$w_0 = c_{w_0,0}(1 + c_{w_0,a}^{\mathbf{A}} a^2) + c_{w_0,l} m_l + c_{w_0,h}(m_h - m_h^{\text{phys}}), \tag{H13}$$

$$\sqrt{t_0} = c_{\sqrt{t_0},0}(1 + c_{\sqrt{t_0},a}^{\mathbf{A}} a^2) + c_{\sqrt{t_0},l} m_l + c_{\sqrt{t_0},h}(m_h - m_h^{\text{phys}}). \tag{H14}$$

These fit functions are used for both the ChPTFV/ChPT and analytic ansätze.

Note that in the expressions above we do not show the a^2 coefficient for the pion, kaon and Omega baryon masses as they are fixed to zero by our choice of scaling trajectory (cf. Section V.A of Ref. [6]).

---

[1] R. A. Haring, M. Ohmacht, T. W. Fox, M. K. Gschwind, D. L. Satterfield, K. Sugavanam, P. W. Coteus, P. Heidelberger, M. A. Blumrich, R. W. Wisniewski, A. Gara, G. L.-T. Chiu, *et al.*, IEEE Micro **32**(2), 48 (2012), ISSN 0272-1732.

[2] P. A. Boyle, N. H. Christ, and C. Kim, IBM Journal of Research and Development pp. 13–13 (2013).

[3] P. A. Boyle, M. I. Buchoff, N. H. Christ, T. Izubuchi, C. Jung, T. C. Luu, R. D. Mawhinney, C. Schroeder, R. Soltz, P. Vranas, and J. Wasem, in *SC'13* (2013), pp. 4–4.

[4] P. A. Boyle, PoS **LATTICE2012**, 020 (2012).

[5] R. Arthur *et al.* (RBC and UKQCD collaborations), Phys.Rev. **D87**, 094514 (2013), 1208.4412.

[6] Y. Aoki *et al.* (RBC and UKQCD collaborations), Phys.Rev. **D83**, 074508 (2011), 1011.0892.

[7] T. Blum, T. Izubuchi, and E. Shintani, Phys.Rev. **D88**, 094503 (2013), 1208.4349.

[8] Y. Shamir, Nucl.Phys. **B406**, 90 (1993), hep-lat/9303005.

[9] V. Furman and Y. Shamir, Nucl.Phys. **B439**, 54 (1995), hep-lat/9405004.

[10] R. C. Brower, H. Neff, and K. Orginos, Nucl.Phys.Proc.Suppl. **140**, 686 (2005), hep-lat/0409118.

[11] R. C. Brower, H. Neff, and K. Orginos, Nucl.Phys.Proc.Suppl. **153**, 191 (2006), hep-lat/0511031.

[12] R. C. Brower, H. Neff, and K. Orginos (2012), 1206.5214.

[13] A. Borici, Nucl.Phys.Proc.Suppl. **83**, 771 (2000), hep-lat/9909057.

[14] A. Borici pp. 41–52 (1999), hep-lat/9912040.

[15] E. I. Zolotarev, Zap. Imp. Akad. Nauk. St. Petersburg **30**, no. 5 (1877), reprinted in his Collected works, Vol. 2, Izdat, Akad. Nauk SSSR, Moscow, 1932, p. 1-59.

[16] R. G. Edwards, U. M. Heller, and R. Narayanan, Nucl.Phys. **B540**, 457 (1999), hep-lat/9807017.

[17] J. van den Eshof, A. Frommer, T. Lippert, K. Schilling, and H. van der Vorst, Comput.Phys.Commun. **146**, 203 (2002), hep-lat/0202025.

[18] A. D. Kennedy, M. A. Clark, and P. J. Silva, PoS **LAT2009**, 021 (2009), 0910.2950.

[19] M. Albanese *et al.* (APE), Phys. Lett. **B192**, 163 (1987).

[20] P. de Forcrand, M. Garcia Perez, and I.-O. Stamatescu, Nucl.Phys. **B499**, 409 (1997), hep-lat/9701012.

[21] S. Schaefer, R. Sommer, and F. Virotta (ALPHA Collaboration), Nucl.Phys. **B845**, 93 (2011), 1009.5228.

[22] N. H. Christ, C. Dawson, T. Izubuchi, C. Jung, Q. Liu, *et al.*, Phys.Rev.Lett. **105**, 241601 (2010), 1002.2999.

[23] A. Stathopoulos and K. Orginos, SIAM J.Sci.Comput. **32**, 439 (2010), 0707.0131.

[24] Q. Liu, Columbia University PhD. Thesis (2012).

[25] D. Calvetti, L. Reichel, and D. C. Sorensen, ETNA **2**, 1 (1994).

[26] M. Luscher, JHEP **0707**, 081 (2007), 0706.2298.

[27] P. A. Boyle (2014), 1402.2585.

[28] R. J. Hudspith (2014), 1405.5812.

[29] T. Blum, P. Chen, N. H. Christ, C. Cristian, C. Dawson, *et al.*, Phys.Rev. **D69**, 074502 (2004), hep-lat/0007038.

[30] S. R. Sharpe (2007), 0706.0218.

[31] P. A. Boyle, PoS(LATTICE2014)087 (2014), 1411.5728.

[32] M. Luscher, JHEP **1008**, 071 (2010), 1006.4518.

[33] S. Borsanyi, S. Durr, Z. Fodor, C. Hoelbling, S. D. Katz, *et al.*, JHEP **1209**, 010 (2012), 1203.4469.

[34] A. Roessl, Nucl.Phys. **B555**, 507 (1999), hep-ph/9904230.

[35] C. Allton *et al.* (RBC and UKQCD collaborations), Phys.Rev. **D78**, 114509 (2008), 0804.0473.

[36] G. Colangelo, S. Durr, and C. Haefeli, Nucl.Phys. **B721**, 136 (2005), hep-lat/0503014.

[37] M. Luscher, Commun.Math.Phys. **104**, 177 (1986).

[38] K. A. Olive *et al.* (Particle Data Group), Chin.Phys. **C38**, 090001 (2014).

[39] G. Martinelli, C. Pittori, C. T. Sachrajda, M. Testa, and A. Vladikas, Nucl.Phys. **B445**, 81 (1995), hep-lat/9411010.

[40] Y. Aoki, R. Arthur, T. Blum, P. A. Boyle, D. Brommel, *et al.*, Phys.Rev. **D84**, 014503 (2011), 1012.4178.

[41] J. Frison *et al.*, PoS **LATTICE2014**, 285 (2014).

[42] S. Aoki, Y. Aoki, C. Bernard, T. Blum, G. Colangelo, *et al.* (2013), 1310.8555.

[43] E. Follana, C. T. H. Davies, G. P. Lepage, and J. Shigemitsu (HPQCD Collaboration, UKQCD Collaboration), Phys.Rev.Lett. **100**, 062002 (2008), 0706.1726.

[44] A. Bazavov *et al.* (MILC Collaboration), PoS **LATTICE2010**, 074 (2010), 1012.0868.

[45] S. Dürr, Z. Fodor, C. Hölbling, S. D. Katz, S. Krieg, *et al.*, Phys.Lett. **B705**, 477 (2011), 1106.3230.

[46] A. Bazavov *et al.* (MILC Collaboration), PoS **CD09**, 007 (2009), 0910.2966.

[47] A. Bazavov, C. Bernard, C. DeTar, X. Du, W. Freeman, *et al.*, PoS **LATTICE2010**, 083 (2010), 1011.1792.

[48] S. Durr, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, *et al.*, Phys.Lett. **B701**, 265 (2011), 1011.2403.

[49] S. Durr, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, *et al.*, JHEP **1108**, 148 (2011), 1011.2711.

[50] J. Laiho and R. S. Van de Water, PoS **LATTICE2011**, 293 (2011), 1112.4861.

[51] T. Bae, Y.-C. Jang, H. Jeong, J. Kim, J. Kim, *et al.*, PoS **LATTICE2013**, 476 (2013), 1310.7319.

[52] S. Borsanyi, S. Durr, Z. Fodor, S. Krieg, A. Schafer, *et al.*, Phys.Rev. **D88**, 014513 (2013), 1205.0788.

[53] S. R. Beane, W. Detmold, P. M. Junnarkar, T. C. Luu, K. Orginos, *et al.*, Phys.Rev. **D86**, 094509

(2012), 1108.1380.

[54] Y.-C. Chen and T.-W. Chiu (2014), 1403.1683.

[55] T. Blum, T. Doi, M. Hayakawa, T. Izubuchi, and N. Yamada, Phys.Rev. **D76**, 114508 (2007), 0708.0484.

[56] T. Blum, R. Zhou, T. Doi, M. Hayakawa, T. Izubuchi, *et al.*, Phys.Rev. **D82**, 094508 (2010), 1006.1311.

[57] T. Ishikawa, T. Blum, M. Hayakawa, T. Izubuchi, C. Jung, *et al.*, Phys.Rev.Lett. **109**, 072002 (2012), 1202.6018.

[58] N. Tantalo (2013), 1311.2797.

[59] S. Borsanyi, S. Durr, Z. Fodor, J. Frison, C. Hoelbling, *et al.*, Phys.Rev.Lett. **111**, 252001 (2013), 1306.2287.

[60] S. Borsanyi, S. Durr, Z. Fodor, C. Hoelbling, S. Katz, *et al.* (2014), 1406.4088.

[61] Z. Davoudi and M. J. Savage, Phys.Rev. **D90**, 054503 (2014), 1402.6741.

[62] C. Sachrajda, PoS **LATTICE2014** (2014).

[63] G. McGlynn and R. D. Mawhinney, PoS **Lattice2013**, 027 (2014), 1311.3695.

[64] P. A. Boyle, Comput.Phys.Commun. **180**, 2739 (2009).

[65] V. Maillart and F. Niedermayer (2008), 0807.0030.

[66] Y. Kikukawa and A. Yamada, Nucl.Phys. **B547**, 413 (1999), hep-lat/9808026.

[67] M. Gockeler, R. Horsley, H. Oelrich, H. Perlt, D. Petters, *et al.*, Nucl.Phys. **B544**, 699 (1999), hep-lat/9807044.

[68] Y. Aoki, P. A. Boyle, N. H. Christ, C. Dawson, M. A. Donnellan, *et al.*, Phys.Rev. **D78**, 054510 (2008), 0712.1061.

[69] R. Arthur and P. A. Boyle (RBC and UKQCD collaborations), Phys.Rev. **D83**, 114511 (2011), 1006.0422.

[70] C. Sturm, Y. Aoki, N. Christ, T. Izubuchi, C. T. C. Sachrajda, *et al.*, Phys.Rev. **D80**, 014501 (2009), 0901.2599.

[71] R. Arthur, P. A. Boyle, N. Garron, C. Kelly, and A. T. Lytle (RBC and UKQCD collaborations), Phys.Rev. **D85**, 014501 (2012), 1109.1223.

[72] P. A. Boyle, N. Garron, and A. T. Lytle (RBC and UKQCD collaborations), PoS **LATTICE2011**, 227 (2011), 1112.0537.

[73] C. Jung, PoS **LATTICE2013**, 417 (2013).

[74] T. Blum, P. A. Boyle, N. H. Christ, N. Garron, E. Goode, *et al.*, Phys.Rev. **D84**, 114503 (2011), 1106.2714.

[75] M. Matsumoto and T. Nishimura, ACM Trans. Model. Comput. Simul. **8**(1), 3 (1998), ISSN 1049-3301, URL `http://doi.acm.org/10.1145/272991.272995`.