Spring 5-15-2018

Lattice Scales from Gradient Flow and Chiral Analysis on the MILC Collaboration's HISQ Ensembles

Nathan Joseph Brown
Washington University in St. Louis

Recommended Citation
Brown, Nathan Joseph, "Lattice Scales from Gradient Flow and Chiral Analysis on the MILC Collaboration's HISQ Ensembles" (2018). *Arts & Sciences Electronic Theses and Dissertations*. 1514.
https://openscholarship.wustl.edu/art_sci_etds/1514

Washington University in St. Louis

Department of Physics

Dissertation Examination Committee:
Claude Bernard, Chair
Bhupal Dev
Renato Feres
Martin Israel
Michael Ogilvie
Demetrios Sarantites

Lattice Scales from Gradient Flow and Chiral Analysis on the MILC Collaboration's HISQ Ensembles

by

Nathan Brown

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2018
Saint Louis, Missouri

Contents

List of Tables

List of Figures

Acknowledgments

First and foremost, I would like to acknowledge the tremendous amount of effort and patience Prof. Claude Bernard has put into my development. Without his steady head, vast depth and breadth of knowledge, and insights, an already long dissertation project could have taken much longer. I would also like to especially thank him for the incredibly helpful and consistent writing reviews, even when I was not as consistent.

Next, I am grateful to the faculty and fellow graduate students that helped me gain an understanding of the broader physics community and learn to have fun while learning. I would like to especially point out the steady members of my committee, Prof. Michael Ogilvie, Prof. Carl Bender, and Prof. Mark Alford. I have had the pleasure of sitting in all of their classes as well as their guidance on a long road.

Outside of the department I would like to acknowledge the MILC Collaboration members, who welcomed me into their collaboration and often provided as much support, it not more, than local faculty. I always enjoyed the dedicated time and atmosphere during the summer and winter workshops hosted by Prof. Carleton DeTar in Utah and Prof. Doug Toussaint in Arizona, respectively.

Finally, I would like to thank my wife Miranda and newborn daughter Nyssa. Miranda's love and patience, especially these last couple months, has been superhuman, and Nyssa's unconditional giggling kept me working hard long past exhaustion.

Nathan Brown

Washington University in Saint Louis
May 2018

ABSTRACT OF THE DISSERTATION

Lattice Scales from Gradient Flow and
Chiral Analysis on the MILC Collaboration's HISQ Ensembles
by
Nathan Brown
Doctor of Philosophy in Physics
Washington University in St. Louis, May 2018
Research Advisor: Professor Claude Bernard

The interactions of quarks and gluons form most of the visible matter around us. Yet, extracting precise predictions from the field theory describing them, Quantum Chromodynamics (QCD), is notoriously difficult. By simulating the QCD interaction on a Euclidean space time lattice, the field theory can be regularized non-perturbatively and familiar statistical techniques from classical statistical mechanics can be applied. Then, by systematically improving each component of the process, high precision results can be obtained. Some of the possible components to be improved include the discretization of the continuum action, the determination of the lattice scale(s), the generation of gauge field ensembles, and the interpolations to physical quark masses.

This work focuses on two primary analyses. Both start from the MILC collaboration's gauge field ensembles with $N_f = 2+1+1$ flavors of highly improved staggered quarks (HISQ). The first analysis focuses on high precision scale setting. Gradient flow is used to smooth the gauge configurations and determine the relative lattice spacings for continuum extrapolations. The relative scale is set through two theoretically motivated observables, $\sqrt{t_0}$ and w_0. The continuum value and mass dependence for each of these scales is extracted to aid with future scale setting. With the lattice spacing set, the second analysis uses staggered chiral perturbation theory to fit lattice data for pseudoscalar meson masses and decay constants. This allows simultaneous interpolation of results to physical quark masses and control of discretization artifacts introduced by the staggered formalism. After extrapolation of the results to the continuum, preliminary results for the pion and kaon decay constants and light quark masses are obtained.

Chapter 1

Introduction

The Standard Model of particle physics has enjoyed widespread acceptance as a description of the fundamental interactions. However, there exist several frontiers where the current version of the Standard Model has not explained experimental observations. In addition, the range of energy scales over which the model has been explored, both experimentally and theoretically, is limited. To test possible extensions of the Standard Model and efficiently explore uncharted territory, high precision techniques are required. This work contains high precision calculations of the properties of fundamental particles in the part of the Standard Model that describes the strong interactions of quarks and their bound states: Quantum Chromodynamics (QCD). In this chapter, the stage is set for these calculations with a review of the theoretical foundations of Lattice QCD and a presentation of the data sets used to perform the calculations.

1.1 Lattice QCD

Quantum field theories suffer from the problem of divergences, which arise from the infinite number of degrees of freedom of a field defined at every point in continuous space-time. In particular, there is no lower limit of the wavelength of modes of the field in Fourier-transform space, and thus no upper limit on the energy and momentum of the modes. QCD has the additional feature that the strength of the interaction between particles at long distances is so strong that the theory is inherently non-perturbative. Both the non-perturbative nature of QCD and the issue of divergences must be addressed in calculations of properties of bound states of quarks, such as protons and neutrons or, as considered in this work, pions and kaons. In order to regularize the divergences, continuous space-time can be replaced with a discrete, finite-volume lattice. Lattice regularization is appropriate for non-perturbative calculations because the theory becomes well defined without reference to a perturbative expansion. This makes the lattice an ideal approach for studying QCD.

The introduction of the lattice motivates several questions, such as how the particle fields of QCD should be represented, how the action should be modified, and how well the lattice approximates reality. The lattice also has inherent benefits, such as a finite number of degrees of freedom. This allows the lattice theory to be simulated by computer, making it possible to perform precise calculations and extrapolations of observable quantities. In this section the lattice is introduced to continuum QCD and some of the resulting questions and benefits are explored in greater depth.

1.1.1 Quantum Chromodynamics

The Standard Model of particle physics classifies all known elementary particles and provides a model of the three relevant interactions between these particles: strong, weak, and electromagnetic. (The fourth fundamental interaction, gravity, is extremely weak and generally irrelevant for elementary-particle interactions.) Each of these interactions is represented as a gauge-invariant quantum field theory, where the force between fermionic fields (half-integer-spin particles) is carried by a specific gauge boson field (integer-spin particle). Since each interaction is gauge invariant, with a unique symmetry group, the interactions are often identified by this group. Two of these interactions, electromagnetism and the weak force, are best understood when unified into the electroweak force with gauge group $U(1) \times SU(2)$. The remaining interaction, the strong force, is described by the $SU(3)$ gauge group. The associated field theory is called Quantum Chromodynamics (QCD).

The fermionic fields of QCD represent six *flavors* of quarks and antiquarks, which can be organized into three generations based on their masses and quantum numbers: up (u) and down (d), strange (s) and charm (c), and bottom (b) and top (t). Each flavor of quark carries three color charges ("red," "green," and "blue") which transform as a triplet under the gauge group. The gauge field of QCD represents eight massless species of gluons labeled by a combination of color and anticolor and in direct correspondence with the eight generators of $SU(3)$; the gauge fields thus transform in the adjoint representation of the gauge group. The gluons of QCD are the only massless gauge fields in the Standard Model that can directly interact with each other (*i.e.*, they possess color charge). The reasons why gluons have color charge and the consequences of their interaction is worth further exploration.

The QCD Lagrangian

The Lagrangian of QCD, and all gauge-invariant quantum field theories, can be derived by starting from the Lagrangian of a free fermionic field $\psi(x)$ [78].[1] Applying a local gauge transformation to the kinetic energy term will change the Lagrangian because the derivative, which connects the field at different locations, does not transform simply under gauge transformations. To account for this, a 4-component gauge field $A_\mu(x)$ ($\mu = 0, 1, 2, 3$) is introduced to define a covariant derivative that does transform simply. Specifically, let $V(x) \in SU(3)$ specify a local gauge transformation where

$$\begin{aligned} \psi(x) &\rightarrow V(x)\psi(x), && (1.1) \\ A_\mu(x) &\rightarrow V(x)\left(A_\mu(x) + \frac{i}{g}\partial_\mu\right)V^\dagger(x), && (1.2) \end{aligned}$$

and $g \in \mathbb{R}$ is a gauge coupling constant. One can define the covariant derivative $\not{D}$,

$$\not{D} = \gamma_\mu(\partial_\mu - igA_\mu), \tag{1.3}$$

where γ^μ are the 4 Dirac matrices in Euclidean space and the repeated index μ is implicitly summed over. Then the gauge transformation of $\not{D}\psi$ is identical to the transformation of ψ.

$$\not{D}\psi(x) \rightarrow V(x)\not{D}\psi(x). \tag{1.4}$$

---

[1] Unless specifically noted, I will be working exclusively in "Euclidean space," where the t of Minkowski space is replaced with it so that the metric is Euclidean. See Sec. 1.1.4 for further details.

With the covariant derivative in place, the fermion Lagrangian L_{fermion} with mass m is now invariant under the local gauge transformation.

$$L_{\text{fermion}}\left[\psi, m, A, g\right] = \bar{\psi}\left(\not{D} - m\right)\psi. \tag{1.5}$$

The full QCD Lagrangian is then formed by adding L_{fermion} for each flavor of quark and introducing a kinetic energy term L_{gauge} for A_μ that is also locally gauge invariant. [2]

$$L_{\text{QCD}} = \left(\sum_{i=1}^{N_f} L_{\text{fermion}}\left[\psi_i, m_i, A, g\right]\right) + L_{\text{gauge}} \tag{1.6}$$

The simplest kinetic energy term for A_μ can be constructed from the square of the gauge field strength,

$$\begin{aligned} F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu - ig\left[A_\mu, A_\nu\right] && (1.7)\\ L_{\text{gauge}} &= -\frac{1}{2}\text{Tr}\left[F_{\mu\nu}F_{\mu\nu}\right]. && (1.8) \end{aligned}$$

Under a gauge transformation, $F_{\mu\nu}(x)$ transforms covariantly:

$$F_{\mu\nu}(x) \to V(x)F_{\mu\nu}(x)V^\dagger(x), \tag{1.9}$$

from which it follows that L_{gauge} is gauge invariant.

The equations of motion come from varying the Lagrangian with respect to A_μ (and its derivatives); therefore, cubic and quartic terms in A_μ or its derivatives in the Lagrangian lead

---

[2] The different fermions in Eq. (1.6) have independent masses; however, the bare coupling g of each fermion to the gauge field is the same. This is because the coupling constant g can be defined independently of the fermions through the self-interaction of the gauge field.

to non-linear equations of motion, which are interpreted as self-coupling. One can therefore see that the coupling of gauge fields to one another arises from the non-zero commutator in $F_{\mu\nu}(x)$. In other words, the non-zero coupling of gauge fields with one another is a direct consequence of the non-Abelian structure of the gauge group.

The Running Coupling Constant

One of the most important consequences of gluons self-interacting is the effect on the energy-scale dependence of the strong coupling constant g. This effect can be inferred through the beta function $\beta(g) = \mu \frac{dg}{d\mu}$ of QCD, where μ is an energy-scale. At one loop in perturbation theory, the beta function for $SU(N)$ gauge field theory is [78]

$$\beta(g) = -\frac{b_0 g^3}{(4\pi)^2}, \qquad b_0 = \frac{11}{3}N - \frac{2}{3}N_f. \tag{1.10}$$

The term in b_0 proportional to N arises from loops involving interactions between the gauge fields while the term proportional to N_f comes from fermionic interactions. In a $U(1)$ gauge field theory where gauge fields cannot interact, the term proportional to N does not appear. The presence of this gauge-field interaction term in theories like QCD is important because it changes the sign of the beta function.

Since b_0 is positive for $SU(N)$ gauge groups with sufficiently small numbers of fermions, such as QCD where $N = 3$ and $N_f = 6$, the beta function is negative. In such a theory, the running coupling $\alpha_s(\mu) = g^2(\mu)/4\pi$ will be small at large energies, and it will be large at small energies. This effect is so important that the two limiting cases are given names: asymptotic freedom and confinement.

Asymptotic freedom refers to interactions at large energy (or small separations) where the coupling is sufficiently small and perturbation theory in α_s can be applied. In the this regime, quarks are effectively free. Confinement refers to the opposite end of the spectrum, where quarks at large separations (or small energy) possess a very large coupling. As a consequence, the potential energy of the interaction between quarks increases with separation. This is opposite to the usual expectation developed from classical theories of electromagnetism and gravity where interactions at large distances can be safely ignored. As the separation between quarks increases, there will eventually be enough potential energy to create quark-antiquark pairs. These new quark pairs bind with the original quarks to form bound states. As a result, quarks are never found free in everyday circumstances where the temperatures are too low to produce significant energy; they are confined to bound states.

Chiral Symmetry [78, 71]

Aside from local gauge symmetry, there are several other important symmetries that can be inferred from the continuum QCD Lagrangian in Eq. (1.6). This section focuses on the vector and larger chiral symmetries, where flavors of right and left handed light quarks are independently mixed. In the limit of massless quarks the chiral symmetry is exact, but it can be applied to theories with light quark masses. In Secs. 1.1.5 and 1.2.2 this symmetry is shown to have several important implications for tuning lattice simulations to physical quark masses.

Since vector and chiral symmetries may mix flavors, the Lagrangian of Eq. (1.6) is best expressed in matrix form. In the case of N fermionic fields denoted by field ψ_i with mass m_i for $i = 1, ...N$, the free fermion Lagrangian is

$$L_{\text{fermion,N}} = \bar{\psi}(\not{D} - M)\psi \tag{1.11}$$

where $\psi = (\psi_1, \psi_2, ..., \psi_N)$, $\not{D} = I_N \not{D}$, and M is a diagonal matrix with $M_{ii} = m_i$. In the case of degenerate masses, this Lagrangian clearly has a new symmetry under rotations $U \in U(N)$ of the N flavors of fermions.

$$\psi \to U\psi, \qquad \bar{\psi} \to \bar{\psi}U^\dagger \tag{1.12}$$

This is referred to as the vector symmetry $U(N)$. The vector symmetry is seen to be a subgroup of the larger chiral symmetry group after the original field ψ is decomposed into its left and right-handed chiral projections, $\psi = \psi_L + \psi_R$.

$$\begin{aligned}
\psi_R &= \frac{1}{2}(1 + \gamma_5)\psi, \qquad \bar{\psi}_R = \bar{\psi}\frac{1}{2}(1 - \gamma_5) && (1.13)\\
\psi_L &= \frac{1}{2}(1 - \gamma_5)\psi, \qquad \bar{\psi}_L = \bar{\psi}\frac{1}{2}(1 + \gamma_5) && (1.14)
\end{aligned}$$

Notice, the left (right) handed projector for particles is the right (left) handed projector for anti-particles. In terms of these projections, the Lagrangian for N massive quarks becomes

$$L_{\mathrm{fermion,N}} = \bar{\psi}_L \not{D} \psi_L + \bar{\psi}_R \not{D} \psi_R - \bar{\psi}_L M \psi_R - \bar{\psi}_R M \psi_L. \tag{1.15}$$

The two projections decouple in the kinetic term due to the presence of γ_μ in $\not{D}$. This allows the $U(N)$ symmetry of massless (special case of degenerate) fermions to be expanded to $U(N)_L \times U(N)_R$ where the left and right handed chiral projections transform separately.

$$\begin{aligned} \psi_L &\rightarrow U_L \psi_L, \qquad \bar{\psi}_L \rightarrow \bar{\psi}_L U_L^\dagger \quad (1.16) \\ \psi_R &\rightarrow U_R \psi_R, \qquad \bar{\psi}_R \rightarrow \bar{\psi}_R U_R^\dagger \quad (1.17) \end{aligned}$$

While the Lagrangian possesses chiral symmetry, the measure $[d\bar{\psi}][d\psi]$ does not respect the complete chiral symmetry. Specifically, the axial symmetry $U(1)_A$ where the singlet left and right handed portions are treated differently $U_L = U_R^\dagger = \exp(i\theta) I_N$ is lost due to this chiral anomaly. The remaining chiral symmetry is $SU(N)_L \times SU(N)_R \times U(1)_V$.

1.1.2 Lattice Regularization

Most continuum quantum field theories exhibit divergences due to the infinite degrees of freedom available to virtual particles traveling around internal loops. Specifically, every loop in perturbation theory that contains a particle with momentum p unconstrained by momentum conservation contributes a factor of $\int d^4p$. This integral will diverge unless enough internal particles that depend on the momentum p are present to make the contributions from large momentum quickly diminish. To deal with these divergences and perform calculations that are finite at each step for observables the theory is regularized by introducing a maximum energy scale (a "cutoff"), which effectively keeps the number of degrees of freedom finite. There are several methods for introducing this energy scale, such as placing an upper bound on the momenta, decreasing the dimension of spacetime, or replacing continuous spacetime with a discretized version. After any of these methods are used, the energy scale introduced by regularization is removed from physical observables by renormalizing all the parameters and fields of the theory in terms of observable quantities.

In QCD, a non-perturbative method of regularizing the integrals is required for low energy scales due to the large coupling at these scales; lattice regularization meets this criterion. In lattice field theories, continuous space-time is replaced with a lattice of points separated by a finite spacing a. In general, this spacing can be different for each dimension but all further discussion will be restricted to cases where a is uniform across dimensions. The Fourier transform of any function on the lattice will be periodic in momentum space with a period of $2\pi/a$. Therefore, the momentum of particles is restricted to the interval $(-\pi/a, \pi/a]$ which effectively introduces a momentum cutoff of π/a on all fields. When lattice theory is simulated on a computer, the theory needs to be restricted to a finite volume $V = L^4$ ($L = Na$) as well. In this case, there is also an infra-red momentum cutoff with magnitude $2\pi/Na$. The effects of both cutoffs must eventually be removed by taking $a \to 0$ and $L \to \infty$. Controlling the errors introduced by extrapolating a to 0 requires a systematic study of how the parameter a is introduced into the path integral and what affects it may have on any symmetries of the continuum theory.

1.1.3 Lattice Action and Fields [63]

A naive modification of the continuum field theory when placed on the lattice would be to place all fields on the sites of the lattice and replace all derivatives with appropriate finite differences. However, this will not work because, as shown in the derivation of the continuum QCD Lagrangian in Sec. 1.1.1, the gauge fields form the covariant derivative needed to maintain the gauge invariance of the Lagrangian. Since the gauge fields are modifications to the derivative which, when replaced with a finite difference, links adjacent lattice sites, the gauge fields must be placed on the links between sites. With this modification, the Lagrangian of Lattice QCD is derived using a similar procedure to the derivation of any other field theory Lagrangian.

Gauge Fields on the Lattice

The gauge fields sit on all links of the lattice. The link variable $U_\mu(x) \in SU(3)$ associated with the gauge field at a specific link is identified by the location x of a site on one end of the link and the direction μ to the connected site $x + \hat{\mu}a$. To construct the lattice version of L_{gauge} a combination of link variables that is invariant under a local gauge transformation $V(x)$ must be derived.

Under a local gauge transformation, the link variables transform covariantly as

$$U_\mu(x) \to V(x) U_\mu(x) V^\dagger(x + a\hat{\mu}). \tag{1.18}$$

Since the gauge transformations at different sites are independent of one another, a gauge invariant product of links must form a closed loop. The simplest such loop is a 1×1 square in two dimensions, called the plaquette $U_{x;\mu,\nu}$. The gauge action is then formed from the sum of the trace over all plaquettes

$$\begin{aligned} S_g &= \frac{\beta}{3}\sum_{x,\mu,\nu}^{\nu<\mu} \mathrm{ReTr}\left[1 - U_{x;\mu,\nu}\right] && (1.19) \\ U_{x;\mu,\nu} &= U_\mu(x)U_\nu(x+\hat{\mu}a)U_\mu^\dagger(x+\hat{\nu}a)U_\nu^\dagger(x) && (1.20) \end{aligned}$$

where β is a parameter related to the coupling constant g.

To express β in terms of the continuum gauge field coupling g, the link variables are defined from the gauge fields and required to have a lattice action corresponding to the continuum action in the limit $a \to 0$. Defining the link variable in terms of the path-ordered exponential of the gauge fields along the corresponding link,

$$U_\mu(x) = P\exp\left\{iagA'_\mu(x)\right\}, \qquad A'_\mu(x) = \frac{1}{a}\int_x^{x+\hat{\mu}a} dy_\nu A_\nu(y) \tag{1.21}$$

yields a leading order action

$$S_{\mathrm{gauge}} = -\frac{\beta g^2}{12}\int d^4x \mathrm{Tr}\left[F_{\mu\nu}(x)F_{\mu\nu}(x)\right] + O(a^2). \tag{1.22}$$

By setting $\beta = 6/g^2$, this action reproduces the continuum action (see Eq. (1.8) as $a \to 0$. With the link variable and an associated lattice gauge action that matches continuum QCD, the lattice fermion action can now be defined using an appropriate substitute for the covariant derivative.

Fermion Fields on the Lattice

To create a lattice action for the fermionic field, a discretization must be chosen for the covariant derivative appearing in the fermion portion of the Lagrangian in continuum gauge field theory (Eq. (1.5)). In a free fermion field theory one could start by replacing the partial derivative at site x with a finite difference, preferably a symmetric difference in order to keep the operator Hermitian,

$$\partial_\mu \psi(x) \rightarrow \frac{1}{2a}\left[\psi(x+a\hat{\mu}) - \psi(x-a\hat{\mu})\right]. \tag{1.23}$$

However, this action is not locally gauge invariant. The construction of the discretized covariant derivative is similar, but requires the inclusion of link variables, which act as parallel transporters of the field ψ along the links

$$D_\mu \psi(x) \rightarrow \nabla_\mu \psi(x) = \frac{1}{2a}\left[U_\mu(x)\psi(x+a\hat{\mu}) - U_\mu^\dagger(x-a\hat{\mu})\psi(x-a\hat{\mu})\right]. \tag{1.24}$$

Substituting this covariant derivative into Eq. (1.5) and replacing integrals over space-time with sums yields the fermion lattice action (in Euclidean spacetime, see Sec. 1.1.4)

$$S_{\text{fermion}} = \sum_x a^4 \bar{\psi}(x) \left(\sum_\mu \gamma_\mu \nabla_\mu + m \right) \psi(x) \; . \tag{1.25}$$

While locally gauge invariant, the ‘naive’ action of Eq. (1.25) does not match the desired continuum theory due to an increase in the number of fermions of the theory. In fact, for each dimension of space-time, the process of replacing the first order Dirac equation with a suitable finite difference equation doubles the number of solutions. In other words, in 4-dimensional space-time, there are 16 solutions of the of the Dirac equation, rather than just one. This doubling is directly inferred from the poles of the fermion propagator in momentum space. Since doubling occurs even in the free theory, consider the inverse propagator for free fermions in momentum space derived from Eq. (1.25) with $U_\mu(x) = 1$ for all links.

$$aS^{-1}(ap) = am + i\sum_{\mu} \gamma_\mu \sin(ap_\mu) \tag{1.26}$$

One can quickly infer that poles for a *massless* fermion occur at the sixteen corners of the momentum space hyper-cube, where $ap_\mu = 0$ or π for each μ. In Feynman diagrams, the fifteen extraneous fermions contribute as virtual particles in loops, and thus affect all physical observables even if they do not appear in external lines. Therefore, to correctly match the continuum theory, which has only one particle per fermionic field, the extraneous doublers must be removed. Various modifications can be made to the naive fermion action or path integral to partially or completely remove the doublers. This thesis focuses on the staggered formalism, which is explained further in Sec. 1.2.2.

1.1.4 The Path Integral Formulation [78, 32]

The path integral formulation of quantum field theory is the main starting point for calculations of observable phenomena. In Minkowski spacetime for a scalar field $\phi(x)$, the partition function is defined as

$$Z[J] = \int [d\phi] e^{i[S(\phi)+J\phi]} \tag{1.27}$$

where S is the associated action of the theory under investigation, J represents an external source, and the integral is over all possible configurations of the field ϕ in space-time. The expectation value of an observable property $O[\phi(x)]$ is then calculated through the variation of the partition function with respect to the external source.

$$\begin{aligned} \langle O[\phi(x)]\rangle &= \frac{1}{Z} O\left(\frac{\delta}{i\delta J(x)}\right) Z[J]\bigg|_{J=0} && (1.28)\\ &= \frac{1}{Z[0]}\int [d\phi] O(\phi) e^{iS(\phi)} && (1.29)\end{aligned}$$

To perform calculations on the lattice, importance sampling is used to numerically estimate the integral. However, importance sampling of the above integral is impossible due to the sign problem; the phase $\exp(iS)$ oscillates rapidly and produces drastic cancellations between different parts of the integration region. To get around this problem, a Wick rotation, $t \to it$ and $S \to iS$, is used to move from Minkowski spacetime to Euclidean spacetime.

$$Z[J] = \int [d\phi] e^{-[S(\phi)+J\phi]}, \qquad \langle O[\phi(x)]\rangle = \frac{1}{Z}\int [d\phi] O(\phi) e^{-S(\phi)} \tag{1.30}$$

The partition function of QCD on the Euclidean lattice is defined as

$$Z = \int \Pi_x [d\psi(x) d\bar{\psi}(x)] \Pi_{x,\mu} [dU_\mu(x)] e^{-\left[S_{\text{lat}}(\psi,\bar{\psi},U)+\bar{\eta}\psi+\bar{\psi}\eta+JU\ \right]} \tag{1.31}$$

where S_{lat} is an appropriately discretized action; η, $\bar{\eta}$, and J are external sources of the quarks and gluons; and the path integral is over the quark and antiquark fields ψ, $\bar{\psi}$ at each site and link variables U on each link. Specifically, $[d\psi(x) d\bar{\psi}(x)]$ is the measure over all flavors (and tastes, see Sec. 1.2.2) of the Grassmann variables representing the quarks and $[dU_\mu(x)]$ is the Haar measure for $SU(3)$. As in the case of a scalar field, observables may be calculated by evaluating specific functional derivatives of the partition function with respect to the sources. But, before evaluating these derivatives the partition function can be partially simplified by using the specific form of the action for lattice QCD.

After choosing a formulation of lattice fermions that removes the doublers in the continuum limit, the action of lattice QCD with external sources can be expressed as

$$S_{\text{lat}} = \bar{\psi} M \psi + \bar{\eta}\psi + \bar{\psi}\eta + S_g \tag{1.32}$$

where the quarks' degrees of freedom have been written in matrix form and S_g absorbs all terms not involving the fermion fields. For example, if the naive fermions discussed in Sec. 1.1.3 are chosen, then the element of M connecting position x to y is

$$M_{x,y} = \left(\sum_\mu \gamma_\mu \nabla_\mu + m \right) \delta_{x,y}. \tag{1.33}$$

The important point for this discussion is that the action is quadratic in the field ψ, which means the integrand in the partition function is Gaussian in ψ and can be analytically integrated. Performing the integration yields a determinant factor as well as a propagator connecting the quark's external sources due to the leftover from completing the square

$$Z = \int \Pi_{x,\mu} [dU_\mu(x)] \det[M] e^{\bar{\eta} M^{-1} \eta - S_g}. \tag{1.34}$$

Observable quantities can now be calculated from a simpler integral

$$\begin{aligned}\langle O(\psi,\bar{\psi},U)\rangle &= \frac{1}{Z}\int \Pi_{x,\mu}[dU_\mu(x)]O(-\frac{\delta}{\delta\bar{\eta}},\frac{\delta}{\delta\eta},-\frac{\delta}{\delta J})\det[M]e^{\bar{\eta}M^{-1}\eta-S_g}\Big|_{\eta,\bar{\eta}=0} &(1.35)\\ &= \frac{1}{Z[0]}\int \Pi_{x,\mu}[dU_\mu(x)]O'(M^{-1},U)\det[M]e^{-S_g}. &(1.36)\end{aligned}$$

For the final expression, O is redefined as O' in order to implicitly show that only observables with pairings of ψ with $\bar{\psi}$ would yield non-zero derivatives, inserting factors of the propagator M^{-1} in the process.

Now that the oscillating phase factor from Minkowski space has been replaced with a positive definite exponential and the resulting integral is simplified, importance sampling is used to estimate the observable. Specifically, a subset of size N drawn from all possible gauge field configurations is created where the probability of drawing a specific configuration from the subset approaches the probability distribution of the full set of configurations in the limit $N \to \infty$. The required probability distribution $W_{\text{eq}}[U]$ can be read off directly from Eq. (1.36)

$$P[U] = \frac{1}{Z[0]}\det[M]e^{-S_{\text{g}}(U)}, \tag{1.37}$$

where $\det[M]$ clearly must be real and positive semi-definite. Finally, the expectation value of the observable is approximated by the average over the subset of gauge configurations

$$\langle O\rangle \approx \frac{1}{N}\sum_{\{U_i\}} O'(M^{-1},U_i). \tag{1.38}$$

In summary, to perform a calculation on the lattice two major steps must be taken: choosing an action to approximate the continuum and generating an *important* subset of the possible gauge configurations for that action. After these steps, an observable of interest is chosen, measured, and averaged over the ensemble of gauge configurations. However, the calculation is not done at this point. The lingering effects of lattice regularization must be removed. Also, the chosen action may have a set of bare couplings (such as the quark masses) that do not exactly correspond to reality. To reach a calculation of the observable that can be compared to experimental observations, several gauge ensembles must be generated, and the results must be extrapolated to realistic conditions.

1.1.5 Physical Observables

To compare lattice computations to experimental observations all parameters that are not set to realistic values must be extrapolated or interpolated to match reality. Such parameters almost always include the quark masses, the volume of spacetime, and the lattice spacing. To perform these fits you generally need two things: a data set consisting of a reasonable range of the independent variable(s) and a model for the dependence of the lattice computation on the variable(s). The first requirement is further complicated by the fact that none of the dimensionful independent variables are known ahead of performing the simulation. To pin down the values of the lattice spacings, quark masses, etc. the overall scale must first be set.

Scale Setting

Based on careful inspection of the full lattice action defined in Eq. (1.19) and Eq. (1.25), the parameters for generating an ensemble of gauge field configurations are the coupling β

(or g), the *dimensionless* quark masses am, and the number of lattice points along each dimension $N = L/a$. Notice, all of these parameters are dimensionless. As a consequence, the only observables that can be calculated directly from an ensemble are also dimensionless (e.g. aM_π or a ratio like M_K/M_π). To extract a number that has dimension, a reference scale must be set for the ensemble using a different observable (whose value is taken from experiment or from a prior lattice computation extrapolated to the continuum).

One might be tempted to break apart one or more of these dimensionless quantities and attempt to treat the lattice spacing a as an additional input parameter. The problem with this is that a is not an independent variable; a is a function of g Furthermore, the relationship between g and a is too complicated to evaluate with the required amount of precision ahead of simulations. Instead, a relationship between a and each simulation value of g is determined by extrapolating a dimensionless quantity to zero coupling, then removing the reference scale.

To clarify this procedure, consider a mass independent scale setting scheme where the scale derived in the chiral limit ($am \to 0$) applies at non-zero masses as well. Let m_x be some observable of interest with mass dimension and let m_r be a reference scale with mass dimension and known continuum value. Computations of these observables on each ensemble result in the dimensionless quantities $M_x(g) = am_x$ and $M_r(g) = am_r$. The continuum value m_x is then computed from the continuum extrapolation of the ratio of the ensemble results [77].

$$m_x = m_r \lim_{g \to 0} \frac{M_x(g)}{M_r(g)}, \tag{1.39}$$

where $g \to 0$ is equivalent to $a \to 0$ by asymptotic freedom. This leads to a natural definition of $a(g)$ and $m_x(a)$ away from the continuum that depends on the choice of reference scale

$$a(g) = \frac{M_r(g)}{m_r}, \qquad m_x(a) = \frac{M_x[g(a)]}{a}, \qquad m_x = \lim_{a \to 0} m_x(a). \tag{1.40}$$

This definition of a sets the value of $m_r(a) = m_r$ to be constant. So, if the values of a determined from two reference scales with different g dependence are compared, they will differ by discretization corrections.

With a reference scale chosen and measured, any dimensionful or dimensionless quantities can now be computed on the ensemble. All that is left is to perform the extrapolations to realistic conditions.

Extrapolations/Interpolations

As previously mentioned, there are three primary extrapolations or interpolations that take place in lattice QCD simulations. The lattice spacing must be extrapolated to zero to match the continuum, the effects of simulating on a finite volume must be removed while taking into account any boundary conditions, and for each flavor of simulation quark mass, results must be interpolated or extrapolated to 'physical' quark mass values that reproduce a collection of known hadron masses (one hadron for each unique quark mass).

The most basic model for continuum extrapolations would be a Taylor series in powers of a^2.[3] If improvements are made to the lattice action the coefficient of the leading term can

---

[3] Odd powers of a are not allowed when there is sufficient chiral symmetry, which is the case in all of this work.

be reduced or the entire leading term can be replaced with combinations like $\alpha_s a^2$ where $\alpha_s = g_s^2/(4\pi)$ and g_s is the renormalized coupling constant. The number of powers included in the Taylor series depends on the range of a included in the data set(s) and is often allowed to vary between multiple fits, which contributes some systematic error to the continuum results.

For the finite volume corrections, the simplest methodology is to simulate in large enough volumes that the difference between infinite spacetime and the simulations is negligible. To verify that the difference is small enough, two different volumes can be simulated for a relatively cheap ensemble (e.g. larger lattice spacing) and compared. In some cases, such as where chiral perturbation theory is applicable, the precise functional form of the finite volume corrections is known and can be accounted for while performing the other extrapolations. Removing the finite volume corrections and re-performing the other extrapolations then yields a measure of the impact from finite volume corrections.

As in the previous two cases, light quark masses can be interpolated to the physical values by a naive approach as well as a more sophisticated approach. This work focuses on the use of chiral perturbation theory to provide a well defined expansion in powers of the quark masses m_q up to next-to next-to leading order (NNLO), $O(m_q^2)$, then adds on higher order terms as necessary to reach quark masses closer to the physical strange mass.

Quarks that are heavier than strange may be treated in several ways. In the MILC HISQ ensembles [12], the next-heaviest quark, the charm quark, is included directly in simulations as a dynamical quark flavor. Results computed with the dynamical charm quark mass close to physical can be re-tuned at leading order based on the observable's dependence on the effective energy scale of QCD Λ_{QCD} (see Sec.2.3.2). Higher order effects, if required, can be accounted through heavy quark effective theory and enter as an expansion in powers of

$1/m_c$. For simulations involving still-heavier quarks or that don't treat the charm quark as dynamic, the quarks may be integrated out of the theory by scaling Λ_{QCD}.

Chiral Perturbation Theory [71]

Chiral perturbation theory (χPT) is one example of an effective theory that can be used to control extrapolations. χPT is an effective theory that describes QCD in the low energy regime. In this regime, the effective theory is built as an expansion in the momentum p of the lightest mesons. When the mass m_q is zero, the symmetry is exact, but even in the case of non-zero quark masses, χPT can be written simultaneously as an expansion in powers of the light quark masses and momenta. Once the effective Lagrangian is constructed based on chiral and other symmetries, the functional form used for extrapolations is formed by expanding both the Lagrangian and desired operator to the required order in p^2 and m_q.

While the original continuum Lagrangian is expressed in terms of the quark fields, the chiral effective theory is expressed in terms of the light hadrons since they are the lowest mass excitations of QCD.

The presence of eight hadrons significantly lighter than the rest suggests the full chiral symmetry for the three light quarks $SU(3)_L \times SU(3)_R \times U(1)_V$ is spontaneously broken down to the $U(3)_V$. This spontaneous chiral symmetry breaking would usually produce eight massless Goldstone bosons, one for each degree of freedom lost in the restriction of the chiral group to its vector subgroup. However, the explicit chiral symmetry breaking of the light quark mass terms gives each of these Goldstone bosons a non-zero mass through perturbative corrections. Even with non-zero masses, these eight hadrons represent the lowest mass excitations of QCD and are therefore the fundamental degrees of freedom of the effective continuum Lagrangian.

The original three light-mass quark fields u, d, and s can be packaged into a new bosonic field $\Phi = \sum_{a=1}^{8} \phi_a T_a$, where T_a refers to the eight generators of the broken symmetry group $SU(3)$. This field can also be re-expressed based on the quark content of the eight Goldstone bosons.

$$\Phi = \begin{pmatrix} \frac{\pi^0}{\sqrt{2}} + \frac{\eta}{\sqrt{6}} & \pi^+ & K^+ \\ \pi^- & -\frac{\pi^0}{\sqrt{2}} + \frac{\eta}{\sqrt{6}} & K^0 \\ K^- & \bar{K}^0 & -\frac{2\eta}{\sqrt{6}} \end{pmatrix} \sim \begin{pmatrix} u \\ d \\ s \end{pmatrix} \begin{pmatrix} \bar{u} & \bar{d} & \bar{s} \end{pmatrix} \tag{1.41}$$

A unitary field Σ is then created as the exponential of the hermitian, traceless field Φ.

$$\Sigma = e^{i2\Phi/f} \tag{1.42}$$

Under the full chiral transformation, Σ is transforms as

$$\Sigma \to U_L \Sigma U_R^\dagger, \qquad \Sigma^\dagger \to U_R \Sigma^\dagger U_L^\dagger. \tag{1.43}$$

Aside from the new field Σ, the explicitly chiral-breaking mass term must also be incorporated into the effective theory. Spurion analysis is used to systematically account for how the mass breaks the chiral symmetry. The constant mass matrix M is temporarily elevated as a field and allowed to transform to preserve the chiral symmetry,

$$M \to U_L M U_R^\dagger, \qquad M^\dagger \to U_R M^\dagger U_L^\dagger \tag{1.44}$$

Using the above transformation properties, the effective leading order chiral Lagrangian can be built. The terms must satisfy all symmetries of the full theory, most noticeably chiral symmetry. Gauge invariance is trivial at the level of the chiral theory because the meson fields Φ and Σ are themselves gauge invariant. Since Σ is unitary, the leading order term involving only Σ must contain derivatives. By Lorentz (or 4-dimensional Euclidean) invariance, the minimum number of derivatives needed is two. With M included, a combination with Σ would also respect the symmetry. This yields the following leading order Lagrangian.

$$L_{\mathrm{LO}} = \frac{f^2}{8}\mathrm{Tr}\left(\partial_\mu \Sigma \partial_\mu \Sigma^\dagger\right) - \frac{\mu f^2}{4}\mathrm{Tr}\left(M\Sigma^\dagger + M^\dagger \Sigma\right) \tag{1.45}$$

Substituting in Eq. (1.41) and expanding to tree level yields the meson masses at leading order in terms of the quark masses. In the case where $m_u = m_d \to m_l$, the masses are

$$\begin{aligned} m_\pi^2 &= 2\mu m_l \\ m_K^2 &= \mu(m_l + m_s) \\ m_\eta^2 &= \frac{1}{3}\mu(2m_l + 4m_s). \end{aligned} \tag{1.46}$$

and the Gell-Mann-Okubo relation can be quickly verified.

$$m_\eta^2 = \frac{1}{3}\left(4m_K^2 - m_\pi^2\right) \tag{1.47}$$

To reach higher precision results, the effective Lagrangian must be expanded to higher orders and loops must be included in the expansions of observables. To keep the power counting straight, notice that the squared meson mass m_B^2 is of approximately the same order as a typical quark mass m_q from the tree level expressions. So, factors of on-shell meson momentum $p^2 \sim m_B^2$ arising from the partial derivative operations will correspond to roughly the same order as quark masses entering through M. Next order terms have 4 derivatives, 2 derivatives and 1 factor of M, or 2 factors of M.

1.2 Gauge Ensembles

One of the key requirements for measuring anything on the lattice is an ensemble of gauge configurations representative of the distribution in Eq. (1.37). The generation of this ensemble is nontrivial due to the large dimension of the path integral that needs to be sampled from and the unknown normalization constant $Z[0]$. Section 1.2.1 is an introduction to the process of ensemble generation, and gives an overview of Markov Chains and the Metropolis Algorithm. Then, the more complicated generation algorithms employed for the MILC ensembles are summarized in Sec. 1.2.3.

To compare to physical results, an assortment of ensembles must be generated so that extrapolations can be performed over the lattice spacing, quark masses, and volume. Furthermore, varios considerations need to be taken into account when choosing the action and boundary conditions so that extrapolations can be made with more precise models and controlled expansions. Improvements to the fermion action, with specific attention on the staggered quark formalism, are covered in Sec. 1.2.2. Then, the specific range of ensemble parameters and fermion improvements employed in the generation of the MILC ensembles are outlined in Sec. 1.2.3.

1.2.1 Ensemble Generation [63]

An ensemble of gauge configurations is generally created as a sequence, called a Markov chain, where each individual gauge configuration in the sequence is generated from the previous member. The process associated with transforming one gauge configuration into another is characterized by a transition probability whose elements $P[U', U]$ determine the likelihood of transitioning from the gauge configuration U to U'. By accounting for all the possible ways of transitioning into a new configuration U', the transition probability can also be interpreted as a transformation of the ensemble probability distribution $W[U]$

$$W[U'] = \int [dU] P[U', U] W[U] \equiv PW, \tag{1.48}$$

where the latter form is expressed as matrix multiplication. For the generated sequence of gauge configurations to correspond to a sample from the desired probability density W_{eq} specified in Eq. (1.37), the transition probability must satisfy a few conditions. To reach the desired distribution from any initial configuration, repeated application of the transition probability must move the distribution towards W_{eq}. This can be summarized as a limiting procedure,

$$\lim_{n \to \infty} P^n W = W_{\text{eq}} \tag{1.49}$$

Once the equilibrium distribution is reached, the sequence must continue to sample from the same distribution. In other words, the equilibrium distribution must be an eigenvector of the transition probability with eigenvalue 1 (due to normalization).

$$PW_{\text{eq}} = W_{\text{eq}} \tag{1.50}$$

As long as the equilibrium distribution satisfying these two properties is unique for a given transition probability, the Markov chain initialized from an arbitrary gauge configuration will eventually converge to the required equilibrium distribution.

While these properties are straightforward to understand, it is difficult to immediately check if a transition probability satisfies them. Instead, a handful of simpler conditions are usually imposed which are sufficient to guarantee the above conditions. Chief among these is strong ergodicity

$$P[U', U] > 0. \tag{1.51}$$

In practice, strong ergodicity may not be satisfied by a single update; however, $P[U', U]$ can be redefined to include any finite number of transitions. In this sense, strong ergodicity is satisfied when any gauge configuration is reachable from any other after a finite number of transitions. The second simpler condition that is often employed is detailed balance.

$$P[U', U]W_{\text{eq}}[U] = P[U, U']W_{\text{eq}}[U'] \tag{1.52}$$

Detailed balance requires that during any transition at equilibrium, the probability that is redistributed from configuration U to U' is balanced by the probability redistributed from the reverse process, U' to U. When applied together, strong ergodicity and detailed balance are sufficient to derive Eqs. (1.49) and (1.50). Detailed balance is sufficient here but it is not necessary; in practical applications most algorithms do not satisfy detailed balance.

The equilibrium density is shown to be an eigenvector by directly applying detailed balance to the left hand side of the eigenvalue equation, where the eigenvalue 1 results from the normalization of the transition probability.

$$PW_{\text{eq}} \equiv \int [dU] P[U', U] W_{\text{eq}}[U] = \left(\int [dU] P[U, U'] \right) W_{\text{eq}}[U'] = W_{\text{eq}} \tag{1.53}$$

Deriving the limiting condition in Eq. (1.49) is more involved, but is a direct consequence of the transition probability being positive definite, e.g. strong ergodicity. The Perron-Frobenius theorem states that there must exist a unique maximum eigenvalue and corresponding eigenvector for positive matrices. By proving the eigenvalue 1 associated with the equilibrium density is the largest eigenvalue for the transition matrix, this theorem shows the equilibrium distribution is a unique eigenvector and all other eigenvalues must be less than one. As a consequence, all contributions to an initial density that are orthogonal to the equilibrium density decay away with successive transitions.

Showing $\lambda = 1$ is the maximum possible eigenvalue is a straightforward result after integrating the magnitude of each side of the eigenvalue equation $PW = \lambda W$. Both integrations result in an overall normalization constant C from the potential eigenvector W.

$$\begin{aligned}
|\lambda|\,C &= |\lambda| \int [dU']\,|W[U']| \\
&= \int [dU']\,|\lambda W[U']| \\
&= \int [dU'] \left| \int [dU] P[U',U] W[U] \right| \\
&\leq \int [dU'] \int [dU] P[U',U]\,|W[U]| \\
&= \int [dU]\,|W[U]| \\
&\equiv C
\end{aligned}$$

Cancelling this factor yields the desired result of $|\lambda| \leq 1$.

Now that the simpler conditions of strong ergodicity and detailed balance are understood to be sufficient, the main question still remains: how can a transition probability satisfying these conditions be constructed. Many algorithms exist that meet these conditions. In this section we'll focus on Metropolis algorithms.

The Metropolis algorithm is best summarized as an accept-reject procedure. The first step is to randomly modify the individual links of a gauge configuration U to produce a potential next configuration U'. Then, the modified gauge configuration is either accepted or rejected based on the relative weight of each configuration in the equilibrium probability density $r = W_{\mathrm{eq}}[U']/W_{\mathrm{eq}}[U]$. Specifically, if the new gauge configuration has a lower action (higher probability) then the gauge configuration is accepted. If the new gauge configuration has a higher action, then it is randomly accepted with probability r.

$$P[U', U] \propto \min(1, r) \tag{1.54}$$

Detailed balance can be verified by simple inspection of each case ($r > 1$ and $r < 1$). Strong ergodicity depends primarily on how new gauge links are proposed. Often it is impractical to update all the links simultaneously, due to the high likelihood that $r << 1$. Instead, an individual update might only change one or a smaller subset of links at a time. In such a case ergodicity is restored by repeatedly applying the Metropolis step until every link within the lattice has been considered for update.

One additional feature of the Metropolis algorithm that makes it more advantageous for lattice field theory computations is its independence from any normalization constant. In the case of lattice field theory, the overall normalization constant $Z[0]$ is not calculable, so its necessary to have an algorithm that avoids it. Some alternatives to the Metropolis algorithm, like heat bath, still require a partial normalization to be calculated.

1.2.2 Improving Lattice Fermions

The naive fermion action in Eq. (1.25) will result in a doubling of the number of fermions for each dimension on the Euclidean lattice. The doublers are not directly a problem on the lattice, but the extra fermions must be removed in the continuum limit where any remaining doublers would effect results through internal loops. Wilson fermions attempt to remove the effects of doublers from loops in the continuum by givig them mass os $O(1/a)$ on the lattice, so they decouple int the continuum limit. However, this breaks chiral symmetry and results in additive mass renormalization. Staggered fermions provide an alternative formulation that respects a subset of the original chiral symmetry while addressing the doubler problem in the continuum.

Staggered Fermions

Staggered fermions were first invented by J. Kogut, L. Susskind, and T. Banks [51, 7, 79]. The basic approach of staggered fermions is to decouple the spin components of naive fermions at the lattice sites and redistribute them at adjacent sites. In the process, the original sixty-four naive fermion components are first reduced to sixteen components (four spin by four taste), then finally four components by using the fourth root procedure to remove the unwanted tastes from the continuum loops.

The first reduction in the number of fermion components starts by spin diagonalizing the fermion action at each lattice site [10]. A local transformation $\Gamma(x/a)$ is applied to the fields within the Dirac space

$$\psi(x) \to \Gamma(x/a)\chi(x), \qquad \Gamma(x/a) = \gamma_1^{(x_1/a)}\gamma_2^{(x_2/a)}\gamma_3^{(x_3/a)}\gamma_4^{(x_4/a)}. \tag{1.55}$$

Based on this definition, $\Gamma(x/a)$ has the property that it diagonalizes all of the gamma matrices appearing in the action alongside the covariant derivative. In this case, the leftover diagonal matrix is the identity times a factor of -1 for each gamma matrix that has to be commuted through γ_μ.

$$\eta_{x,\mu} \equiv \Gamma^\dagger(x/a)\gamma_\mu\Gamma(x/a+\hat{\mu}) = (-1)^{(x_1/a+\ldots+x_{\mu-1}/a)} \tag{1.56}$$

Because the resulting action is diagonal, the spin components decouple at each site and it is sufficient to keep only one component. The resulting action is known as the staggered fermion action for one flavor.

$$S_{\mathrm{KS}} = \sum_x a^4 \bar{\chi}(x) \left(\sum_\mu \eta_{x,\mu} \nabla_\mu + m \right) \chi(x) \tag{1.57}$$

At this point the degrees of freedom associated with the sixteen components of the remaining four doublers are spread out over the lattice sites as the one-component fermion field $\chi(x)$. To untangle these degrees of freedom, the remaining components need to be grouped together across the lattice sites. Because the remaining spin dependence in $\eta_{x,\mu}$ depends only on whether x is an even or odd numbered site along each dimension, combining adjacent sites into a unit representing all the possible combinations of even/odd indexes between dimensions is appropriate. This unit corresponds to the unit hypercube.

To translate the staggered action in Eq. (1.57) into a form that sums over the hypercubes of the lattice, each lattice site x is replaced by the combination of the origin of the containing hypercube y and the offset to one of the corners of a hypercube τ [63].

$$x = 2y + a\tau, \qquad \tau_\mu = 0, 1 \tag{1.58}$$

Note that, the separation between the origins of adjacent hypercubes is $a' = 2a$ to avoid including any site in multiple hypercubes. To complete the translation of sites to hypercubes the sum over sites is replaced with a sum over the corners of each hypercube and the dependence of η only on τ is made explicit.

$$S_{\text{KS}} = \frac{1}{16}\sum_y \sum_\tau (a')^4 \bar{\chi}(2y + a\tau)\left(\sum_\mu \eta_{\tau,\mu}\nabla_\mu + m\right)\chi(2y + a\tau) \tag{1.59}$$

Finally, the fields $\chi(2y+a\tau)$ at each corner of the hypercubes are combined into 2^4 component fields $q_{\alpha i}(y)$, labeled by two indices: spin $(1 \leq \alpha \leq 4)$ and taste $(1 \leq i \leq 4)$ [10, 63].

$$q_{\alpha i}(y) = \frac{1}{8}\sum_\tau \Gamma_{\alpha i}(\tau) U(y,\tau)\chi(2y + a\tau) \tag{1.60}$$

$$\bar{q}_{\alpha i}(y) = \frac{1}{8}\sum_\tau \bar{\chi}(2y + a\tau) U^\dagger(y,\tau)\Gamma^\dagger_{\alpha i}(\tau) \tag{1.61}$$

Here, $U(y,\tau)$ is a fixed product of links from the origin of each hypercube to the respective corners, and it is used to transport the field $\chi(2y + a\tau)$ to the origin y so that the kinetic energy term stays locally gauge invariant. The factors of Γ determine both the spin and taste transformation properties of q and $\bar{q}$. As a consequence, the products of $\bar{q}$ with q summed over all taste and spin indices are reducible to a single sum over the corners of the hypercube.

The substitution of q into the action of Eq. (1.59) can be expressed as an expansion in powers of the lattice spacing a' about the continuum

$$S_{\rm KS} = (a')^4 \sum_y \bar{q}(y) \left\{ \sum_\mu (\gamma_\mu \otimes I) \nabla_\mu + m(I \otimes I) + O(a') \right\} q(y), \tag{1.62}$$

where the direct products are a mixture of spin and taste bases and the new first derivative operator is over the doubled lattice spacing a'

$$\nabla_\mu q(y) = \frac{1}{2a'} \left[U(y, 2\hat{\mu}) \psi(y + a'\hat{\mu}) - U^\dagger(y - a'\hat{\mu}, 2\hat{\mu}) \psi(x - a\hat{\mu}) \right]. \tag{1.63}$$

As an example of the kind of operators that appear in the $O(a')$ terms and above, consider the free staggered quark theory where $U_\mu(x) = 1$ [10, 63]. Here, no $O(a'^2)$ or higher terms appear and the remaining $O(a')$ correction explicitly breaks the taste-symmetry of the continuum terms.

$$S_{\rm KS} = (a')^4 \sum_y \bar{q}(y) \left\{ \sum_\mu (\gamma_\mu \otimes I) \nabla_\mu + m(I \otimes I) + a' \sum_\mu (\gamma_5 \otimes \xi_\mu \xi_5) \Delta_\mu \right\} q(y), \tag{1.64}$$

ξ_μ are matrices that act in taste space but are otherwise equivalent to the usual Dirac matrices γ_μ, and Δ_μ is a second derivative operator over the doubled lattice spacing.

$$\Delta_\mu q(y) = \frac{1}{a'^2} \left[U(y, 2\hat{\mu}) q(y + a'\hat{\mu}) - 2q(y) + U^\dagger(y - a'\hat{\mu}, 2\hat{\mu}) q(x - a\hat{\mu}) \right] \tag{1.65}$$

Since all of the potentially taste-breaking terms, like the above second derivative correction, appear at $O(a)$ or higher, taste symmetry is restored in the continuum at the price of four degenerate tastes of quarks. The apparent $O(a)$ corrections are misleading, however. In the interacting theory, it is possible to prove that $O(a)$ effects are artifacts of our particular choice of spin-taste definitions, and the true physical corrections are $O(a^2)$ [73, 56, 54]. In the non-interacting case, the corrections may be removed completely [46].

To finally remove all artifacts of the original doublers from lattice results, the effects of the remaining four degenerate continuum tastes must be reduced to that of a single species at the loop level [10]. To do this, recall that when fermions are integrated out of the continuum partition function they leave behind a factor of $\det(M)$ for each flavor of quark present in the gauge configuration distribution. Since the four tastes act identically to quark flavors and are degenerate in the continuum, they each eventually contribute the same $\det M$. By replacing $\det(M)$ with $\det^{1/4}(M)$ in the partition function, the four tastes are effectively reduced to one in the continuum. This procedure is known as the fourth-root procedure and the resulting lattice fermions are known as rooted, staggered quarks.

Staggered Symmetry [10]

Aside from the computational efficiency of staggered fermions, a primary reason staggered fermions are used in chiral analysis is because they retain a subgroup of the chiral symmetry of the continuum. This can be seen most clearly by considering the one component field χ. Start by defining a site as even or odd depending on whether $\sum_\mu x_\mu/a$ is even or odd. Since the derivative term connects the quark fields at adjacent sites D always switches site parity. In comparison, the mass term maintains parity since it connects fields at the same site. If he mass term is set to zero, the Lagrangian displays an exact even-odd $U(1)_e \times U(1)_o$ symmetry.

$$\begin{aligned} \chi(x) &\rightarrow \exp \begin{pmatrix} i\alpha_e, & x = \text{even} \\ i\alpha_o, & x = \text{odd} \end{pmatrix} \chi(x), && (1.66) \\ \bar{\chi(x)} &\rightarrow \bar{\chi}(x) \exp \begin{pmatrix} -i\alpha_o, & x = \text{even} \\ -i\alpha_e, & x = \text{odd} \end{pmatrix}, && (1.67) \end{aligned}$$

The division of the fermion components into even and odd sets that transform separately is reminiscent of the left and right handed projections transforming under separate flavor transformations. However, there is one important distinction on the lattice. The axial transformation $U(1)_\epsilon$ where $\alpha_e = -\alpha_o \rightarrow \alpha$ is not a taste singlet, and the symmetry is therefore not violated by the anomaly. This is most evident in the spin-taste basis

$$q(y) \rightarrow \exp\left[i\alpha(\gamma_5 \otimes \xi_5)\right] q(y), \qquad \bar{q}(y) \rightarrow \bar{q(y)} \exp\left[i\alpha(\gamma_5 \otimes \xi_5)\right]. \tag{1.68}$$

Note that while the $U(1)_\epsilon$ symmetry is an axial symmetry, it is not the taste-singlet axial symmetry $U(1)_A$ which is still broken by the anomaly. Instead, each of the $U(1)_\epsilon$ symmetries are residual subgroups of the larger chiral symmetry group applied on the lattice.

Because the even-odd symmetry is an exact symmetry in the massless limit and it is separately satisfied for each flavor, there can be no additive mass renormalization terms. When combined with other lattice symmetries that prevent mass-like terms from forming by connecting components of the staggered field on adjacent sites (such as shift symmetries), this ensures that the mass renormalization is only multiplicative.

The Asqtad Improvement [10]

To reach high precision a systematic improvement program must be followed to reduce all lattice artifacts of a given order. This often starts with an improvement of the action used to generate gauge ensembles. Here I will explain in detail the $O(a^2)$ tadpole improvement procedure (called asqtad) for staggered quarks, first developed in [30, 55, 52, 67, 68, 23] While additional improvements are made to the final MILC ensembles analyzed in this work, the asqtad improvement provides an illustrative example of the systematic approach required to reach higher precision without resorting to brute force and the associated computational cost.

The asqtad improvement procedure addresses several different sources of discretization artifacts. First and foremost, there are leading convergence issues introduced by the discretization of fermions on the lattice. These appear at one loop due to so-called *tadpole* diagrams, where internal gluons with high momenta compete with the explicit a^2 dependence from the vertex of the loop. To counteract the effect of tadpole diagrams, a mean field approach is taken which effectively rescales all links in the action by a constant called the tadpole factor μ_0. After this, the taste-breaking lattice artifacts of the staggered formalism must be addressed. Taste breaking effects in the interacting theory are due to the exchange of gluons carrying total momentum with one or more components near π/a. To address the additional exchange of gluons with transverse momentum components near π/a or 0 between two staggered fields, the gauge links can be fattened by adding alternative paths between adjacent sites of the lattice. However, fattening the links causes $O(a^2)$ corrections to be added to the low-momentum coupling of the gluon with the quarks. To cancel those corrections an additional "staple" term is added, called the "Lepage term." To remove $O(a^2)$ artifacts in the first derivative term in Eq. (1.64), the so called "Naik/three hop term" is added to the first order derivative. All of these improvements taken together represent the asqtad improvement procedure for staggered fermions.

The tadpole improvement can be quickly derived by expanding the gauge links in terms of the continuum field, and evaluating the resulting lattice perturbation theory in terms of an expansion in both a and g. Expanding Eq. (1.21) about $a = 0$ and inserting into the interaction $\bar{\psi}U\psi$ yields the expected continuum interaction term, followed by a first order lattice artifact representing a vertex between two quarks and two gluons.

$$U_\mu(x) = P\left\{1 + iagA'_\mu(x) - \frac{a^2g^2}{2}A'_\mu(x)^2 + \dots\right\}, \tag{1.69}$$

While the expansion explicitly pulls out a factor of a^2, contractions of gluons at this vertex also yield factors of the lattice spacing from ultraviolet divergences. Simple power counting with one gluon loop leads a factor of $p^2 \propto 1/a^2$ which cancels the explicit a^2 dependence. To improve the rate of convergence to the continuum, the high momentum contributions of $A_\mu(x)$ within the loop must be integrated out, leaving only lower momentum dynamical gauge fields. This is equivalent to replacing the standard link with a scaled link that correctly reproduces the mean value of a link μ_0 at first order.

$$U'_\mu(x) = \mu_0 U_\mu(x), \qquad \left\langle U'_\mu(x)\right\rangle \approx \left\langle \mu_0\left(1 + iagA'_\mu(x)\right)\right\rangle = \mu_0 \tag{1.70}$$

To leave the action invariant, the overall coefficient $1/\mu_0^4$ of the plaquette is absorbed into the coupling $g^2 \to g'^2 = g^2/\mu_0^4$. Finally, the value of μ_0 can be determined either by directly computing the mean link after fixing the gauge or, more commonly, by measuring through the fourth root of the plaquette without gauge fixing.

$$\mu_0 = \left\langle \frac{1}{3}\mathrm{Tr}U_{x;\mu\nu} \right\rangle^{1/4} \tag{1.71}$$

With the tadpole improvement implemented, the next issue to address is the additional lattice artifacts introduced by the staggered formalism. These artifacts introduce interactions with explicit a dependence, and they contribute to tree level masses of mesons. Therefore, it is especially important to diminish the taste artifacts because they control part of the continuum extrapolation as well as the interpolation to physical quark masses.

The dominant contribution of the taste artifacts comes from the exchange of gluons with momentum components of $p = 0, \pi/a$, except when all components are near 0; that is the physical case. This interaction is provided by the kinetic energy term $\bar{\psi}\gamma_\mu\nabla_\mu\psi$, and could therefore by corrected by altering the gauge links within the covariant derivative. This alteration must maintain the local gauge invariance of the action, remove the interaction term when the components of the momentum of the exchanged gluon are near 0 or π/a (but not all near 0), and vanish in the continuum limit. The addition of various orders[4] of the lattice Laplacian of the gauge link fits these requirements.

---

[4] the Lepage term $\Delta_\nu^{2l}U_\mu(x)$ is explained in further detail later in this section

$$\begin{aligned}\Delta^l_\nu U_\mu(x) &\equiv \frac{1}{a^2}\left[U_\nu(x)U_\mu(x+a\hat{\nu})U^\dagger_\nu(x+a\hat{\mu})\right. && (1.72)\\ &+ \left. U^\dagger_\nu(x-a\hat{\nu})U_\mu(x-a\hat{\nu})U_\nu(x-a\hat{\nu}+a\hat{\mu}) - 2U_\mu(x)\right] \\ \Delta^{2l}U_\mu(x) &\rightarrow U_\mu(x) + \frac{a^2}{4}\sum_{\nu\neq\mu}\Delta^l_\nu U_\mu(x) + \frac{a^4}{32}\sum_{\rho\neq\nu\neq\mu}\Delta^l_\rho\Delta^l_\nu U_\mu(x) && (1.73)\\ &+ \frac{a^6}{384}\sum_{\sigma\neq\rho\neq\nu\neq\mu}\Delta^l_\sigma\Delta^l_\rho\Delta^l_\nu U_\mu(x) - \frac{a^2}{4}\sum_{\nu\neq\mu}\Delta^{2l}_\nu U_\mu(x)\end{aligned}$$

The lattice Laplacian essentially fattens the link it acts on by adding 1 × 1 products of links in a direction ν transverse to the original link direction μ. Since everything it adds is a product of links connecting the same original two lattice sites, it maintains the same gauge covariance as the original link.

To see that the interaction vanishes for momentum near the corners of the Brillouin zone, start by considering any momenta with the longitudinal momentum component at π/a. In this case, the forward and backward terms of the covariant derivative in the μ direction add to give a factor of $cos(ap_\mu/2)$, which vanishes regardless of what the other components of momentum are. For the cases where the longitudinal momentum is 0, we can expand $U\mu(x)$ to first order in g to find the equivalent transformation for $A_\mu(x)$, and then transform to momentum space. Doing this explicitly up to the first order Laplacian yields

$$A_\mu(p) \rightarrow A_\mu(p) + \frac{1}{4}\sum_{\nu\neq\mu}\left[2A_\mu(p)\left(\cos(ap_\nu)-1\right) + 4\sin\left(\frac{ap_\mu}{2}\right)\sin\left(\frac{ap_\nu}{2}\right)A_\nu(p)\right] \quad (1.74)$$

For the case of a single non-zero, transverse component $p = (0, p_\nu, 0, 0)$ with $p_\nu = \pi/a$, the first term in the sum precisely cancels the original gauge field. For concreteness ν is set equal to 2 in the preceding example without loss of generality. Similarly, for two and three transverse components of π/a, the second and third power Laplacian terms are needed to complete the cancellation of the original field, respectively.

The Lepage term $\Delta_\nu^{2l} U_\mu(x)$ is added to cancel $O(a^2)$ modifications of the low-energy quark-gluon interaction introduced by the Laplacian terms. The Lepage term is essentially a modification of the lattice Laplacian with the 1×1 staple extended to a 2×1 staple in the ν direction. Both the Lepage term and the lattice Laplacian terms satisfy

$$\Delta_\nu^l U_\mu(x) \approx \Delta_\nu^{2l} U_\mu(x) \approx a D_\nu F_{\nu,\mu} + O(a^2) \tag{1.75}$$

The coefficient is chosen to cancel the $O(a^2)$ corrections introduced by the various Laplacian terms. Furthermore, the Lepage term vanishes at all of the corners of the Brillouin zone so that the corrections introduced by the Laplacian terms are unaffected.

Finally, to complete the asqtad improvement, the discretization artifacts originating from the free staggered fermion propagator are canceled. This is achieved by adding the Naik term to the covariant derivative.

$$\begin{aligned} \nabla_\mu \chi(x) \quad &\rightarrow \quad \nabla_\mu \chi(x) - \frac{a^2}{6} \nabla_\mu^3 \chi(x) \qquad\qquad (1.76) \\ &= \quad \frac{9}{8} \nabla_\mu \chi(x) - \frac{1}{48a} \left[U_\mu(x) U_\mu(x + a\hat{\mu}) U_\mu(x + 2a\hat{\mu}) \chi(x + 3a\hat{\mu}) \right. \\ &\quad\quad \left. - U_\mu^\dagger(x - a\hat{\mu}) U_\mu^\dagger(x - 2a\hat{\mu}) U_\mu^\dagger(x - 3a\hat{\mu}) \chi(x - 3a\hat{\mu}) \right] \end{aligned}$$

As evident above, when the Naik term is expressed as an action on the single component staggered field $\chi(x)$, it produces links between sites that are three of the original lattice spacings apart. This is required to ensure the couplings between different tastes remains the same.

1.2.3 Overview of MILC Ensembles [12]

At this point, sufficient theoretical background has been provided to present and understand the MILC Collaboration's gauge ensembles. This includes the justifications for parameter ranges such as quark masses and lattice spacings, the various improvements to the action, and how the ensembles are generated.

Tables 1.1 and 1.2 provide an overview of the parameters specifying each ensemble in this work. The only parameters that can be estimated before the complete simulation are the lattice dimensions, ratio fo quark masses, and number of generated configurations. More precise values of the lattice spacing and masses determined from a combination of gradient flow and staggered chiral perturbation theory analysis are results of the latter sections.

There are four lattice spacings, ranging from $a \approx 0.15\,\mathrm{fm}$ down to $a \approx 0.06\,\mathrm{fm}$. Finer MILC gauge ensembles at $a \approx 0.42\,\mathrm{fm}$ and $0.03\,\mathrm{fm}$ are currently being generated, but are not included in this analysis. However, a table with preliminary scale values calculated on these in-progress ensembles is included in Appendix A. All of the ensembles included in this work have $N_f = 2+1+1$ flavors of staggered quarks, where the up and down quark masses are set equal to m_l and the charm quark mass is tuned close to its physical value. There is one ensemble at $a \approx 0.12$fm where the up and down quark masses are are not equal ($m'_{l1} = 0.1m_s$,

$m'_{l2} = 0.25m_s$, $m'_s = 0.45m_s$) in order to probe for isospin breaking effects; however, that ensemble has not yet been included in this preliminary analysis.

The ensembles can be broadly divided into two subsets depending on whether the simulation strange quark mass is approximately set equal to or lower than the physical strange quark mass. The subset of ensembles with nearly-physical strange-quark masses cover all four lattice spacings, and contains three values of the ratio of the simulation light quark mass to the physical strange quark mass m'_l/m_s: 1/5, 1/10, and 1/27 which is approximately the ratio for the physical light quark mass. The subset of ensembles with lighter-than-physical strange-quark masses are all at $a \approx 0.12$fm and the values of m'_l and m'_s are primarily selected to help understand and control strange-quark mass dependence and the extrapolation to the chiral limit (all quark masses vanishing).

In addition to the extra lighter-than-physical strange-quark mass ensembles at $a \approx 0.12$fm, there are two additional volumes simulated for the $m'_l/m'_s = 1/10$, nearly-physical strange-quark mass ensemble. The three volumes with these parameters provide a check that finite volume effects, which typically are proportional to $\exp(-M_\pi L)$, are under control.

The action for the MILC $N_f = 2 + 1 + 1$ gauge ensembles is composed of a one-loop Symanzik improved gauge action and the highly improved staggered quark (HISQ) action. The Symanzik improved gauge action is achieved through the addition smeared plaquettes to the Wilson action, specifically 2×1 and $1 \times 1 \times 1$ loops. The coefficients of these additions are tuned perturbatively at one loop order in the interacting theory and the links are then tadpole improved. This reduces the discretization errors from the gauge action to $O(\alpha_s^2 a^2, a^4)$.

β	$\approx a$ (fm)	m_l'/m_s'	$N_s^3 \times N_t$	$M_\pi L$	M_π (MeV)	$N_{\rm lats}$
5.80	0.15	1/5	$16^3 \times 48$	3.8	314	1020
5.80	0.15	1/10	$24^3 \times 48$	4.0	214	1000
5.80	0.15	1/27	$32^3 \times 48$	3.2	130	1000
6.00	0.12	1/5	$24^3 \times 64$	4.5	299	1040
6.00	0.12	1/10	$24^3 \times 64$	3.2	221	1020
6.00	0.12	1/10	$32^3 \times 64$	4.3	216	1000
6.00	0.12	1/10	$40^3 \times 64$	5.4	214	1028
6.00	0.12	1/27	$48^3 \times 64$	3.9	133	1000
6.30	0.09	1/5	$32^3 \times 96$	4.5	301	1011
6.30	0.09	1/10	$48^3 \times 96$	4.7	215	1000
6.30	0.09	1/27	$64^3 \times 96$	3.7	130	1031
6.72	0.06	1/5	$48^3 \times 144$	4.5	304	1016
6.72	0.06	1/10	$64^3 \times 144$	4.3	224	1166
6.72	0.06	1/27	$96^3 \times 192$	3.7	135	583

Table 1.1: HISQ gauge configuration ensembles with strange and charm quark masses set at or very close to their physical values. The columns from first to last show the coupling $\beta = 10/g^2$ for HISQ, approximate lattice spacing, the ratio of the simulation masses of the light quark to the strange quark, the lattice dimensions, the product of the Goldstone pion mass and the spatial extent of the lattice, the Goldstone pion mass in MeV, and the number of equilibrated gauge configurations. The pion masses have been converted to physical units using $F_{\rm p4s}$ [13].

The HISQ action is an extension of the asqtad improved quark action, designed to further reduce taste violations at $O(\alpha_s^2 a^2)$. While the asqtad action addressed the free staggered fermion propagator and taste violations due to the exchange of a single gluon, taste can still be transformed if the quark field interacts with multiple gluons. The HISQ action adds additional smearings of the gauge links in the covariant derivative in order to reduce, but not eliminate, artifacts of multiple gluon exchanges. Let the asqtad link transformation outlined in Eq. (1.73) be denoted by $\mathcal{F}^{f7L}$ but with the coefficient of the Lepage term smearing doubled. Also, let the same transformation without the 1×2 Lepage term be denoted by $\mathcal{F}^{f7}$. Then the HISQ transformation can be expressed as a series of steps,

m_l'/m_s	m_s'/m_s	$N_s^3 \times N_t$	N_{lats}
0.10	0.10	$32^3 \times 64$	1020
0.10	0.25	$32^3 \times 64$	1020
0.10	0.45	$32^3 \times 64$	1020
0.10	0.60	$32^3 \times 64$	1020
0.25	0.25	$24^3 \times 64$	1020
0.20	0.60	$24^3 \times 64$	1020
0.175	0.45	$32^3 \times 64$	1020

Table 1.2: HISQ gauge configuration ensembles with lighter-than-physical strange quark masses. All ensembles have a lattice spacing of $a \approx 0.12$fm and charm-quark mass as close to its physical value. The columns in order show the approximate ratio of the simulation light quark masses to the physical strange quark mass, the approximate ratio of the simulation strange-quark mass to the physical strange-quark mass, the lattice dimensions, and the number of equilibrated configurations.

$$\begin{aligned} U_\mu(x) &\rightarrow X_\mu(x) = \mathcal{F}^{f7L} W_\mu(x) && (1.77) \\ & \quad\; W_\mu(x) = \mathcal{U}\mathcal{F}^{f7} U_\mu(x) \\ \nabla_\mu[U]\chi(x) &\rightarrow \left(\nabla_\mu[X] - \frac{a^2}{6}(1+\epsilon)(\nabla_\mu)^3[W]\right)\chi(x) && (1.78) \end{aligned}$$

where $\mathcal{U}$ is a projection operator onto $U(3)$ or $SU(3)$ that bounds the smeared links and the last equation is the modified Naik term. In the second equation, square brackets are used to denote which form of the links are used in the operator and ϵ is introduced as a correction to compensate for $O[(am)^4]$ and $O[\alpha_s(am)^2]$ errors [44, 10].

While the HISQ action reduces taste violations and makes it possible to simulate quark masses near charm dynamically, it does not systematically eliminate all taste-violating, two-gluon exchange interactions. Therefore, the orders at which taste-violating and overall discretization errors appear are left unchanged; the coefficients are however significantly reduced. The non-interacting one-loop Symanzik improved gauge action is $O(\alpha_s^2 a^2)$ due to perturbative treatment of the coefficients to one-loop (including dynamical quark loops). The generic discretization error of the tree-level asqtad action is $O(\alpha_s a^2)$, and the removal of single gluon taste-exchange interactions reduces the taste violation errors to $O(\alpha_s^2 a^2)$.

Two algorithms were employed to generate the MILC Gauge ensembles used in this work, both based on a form of hybrid molecular dynamics (HMD) algorithm. In all HMD algorithms a fictitious momentum field is introduced so that the gauge fields can be evolved along a classical trajectory of the Hamiltonian in fictitious time. While the gauge fields evolve, the momentum field is periodically refreshed by a Gaussian heatbath (a Markov Chain algorithm satisfying detailed balance Eq. (1.52), like the Metropolis algorithm). The primary difficulty with performing this evolution is evaluating the contribution to the force on the gauge fields from the fermion determinant. Most enhancements to the basic HMD algorithm in lattice QCD focus on improvements to this calculation. In particular, both the rational hybrid molecular dynamics (RHMD) and rational hybrid monte carlo (RHMC) algorithms are employed in this work to balance the need for sufficient precision of the fermion force with increased computational complexity.

The RHMD algorithm [49] starts by replacing the fourth root of the determinant of the fermion matrix M in the path integral with a Gaussian integral over a new pseudo-fermionic field ϕ whose action is $\bar{\phi} M^{-1/4} \phi$. Then, a rational function is used to approximate $M^{-1/4}$. This rational approximation is the primary practical advantage of the RHMD algorithm, but because of that approximation, as well as inevitable numerical errors in the molecular dynamics updates, it is not an exact algorithm. The RHMC algorithm [49, 35, 37, 36] is a further extension of the RHMD algorithm which applies a Metropolis accept or reject step at the end of each trajectory. The addition of the accept or reject step makes the RHMC algorithm exact with the consequence of a slower evolution, since a new gauge field configuration is not reached with every trajectory.

For the finer lattices ($a \approx 0.09$fm and $m_l/m_s = 1/27$, and all $a \approx 0.06$fm) where the evaluation of the fermion force is well behaved and the imprecision of RHMD is not a significant factor in any observables of interest, the RHMD algorithm is employed. On coarser lattices ($a \approx 0.09$fm and $m_l/m_s \geq 1/10$, and all $a \geq 0.12$fm), RHMC is used. For two ensembles in between ($a \approx 0.09$fm and $m_l/m_s = 1/27$, $a \approx 0.06$fm and $m_l/m_s = 1/10$), both algorithms were employed in order to ensure that the deviations of RHMD from RHMC were not significant compared to statistical error.

Chapter 2

Gradient Flow and Scale Setting [17]

2.1 Introduction

Scale setting holds central importance in lattice QCD for two reasons. First, the continuum extrapolation of any quantity, dimensionful or dimensionless, requires a precise determination of the relative scale between ensembles with different bare couplings. Second, the precision to which one may determine a dimensionful quantity in physical units is limited by the precision of the scale in physical units (the *absolute scale*). Because scale setting limits the precision of so many calculations, it is important to identify quantities with the highest level of precision to set the scale.

To make progress towards this goal a thorough understanding of the restrictions on quantities that may be used for scale setting is required. In principle, any dimensionful quantity that is finite in the continuum limit may be employed. The relative scale may be set by calculating a dimensionful quantity and comparing its value in lattice units at different lattice spacings for the same quark masses. For absolute scale setting, one needs to compare the quantity in lattice units to the physical value. If the quantity is experimentally accessible the comparison to the physical value is straightforward. For a quantity that is inaccessible to experiment, its physical value in the continuum is inferred by comparison to an experimental quantity. In other words, an experimental quantity may be used directly for relative and absolute scale setting, but a quantity that is inaccessible to experiment requires the lattice measurement of a second, experimentally accessible quantity for absolute scale setting. The use of a non-experimental quantity for scale setting may still be worthwhile if it can be determined on the lattice with small statistical and systematic errors for relatively small computational cost. This is due to the large gain in control over continuum extrapolations at the cost of a small decrease in the precision of absolute scales. This has led to the consideration of theoretically-motivated, but not experimentally measurable, quantities such as r_0 and r_1 [76, 21], F_{p4s} [12], and, more recently, $\sqrt{t_0}$ [58] and w_0 [31] from gradient flow [65, 59].

The ideal scale-setting quantity has small statistical and systematic errors. However, since systematic errors arise from a variety of sources, such as discretization effects, dependence on the simulation (possibly unphysical) quark masses, finite volume effects, and excited states, it is difficult to reduce all error sources simultaneously. For example, the scales r_0 and r_1 are computed from asymptotic fits in time t to the heavy quark potential $V(r)$ with quark separation r, such that $r^2 dV/dr = 1.65$ or 1, for $r = r_0$ and r_1, respectively [76, 21]. The statistical errors in $V(r)$ are generally small, but they grow with t/a and may become a problem at small lattice spacings where larger values of t/a are needed to reduce systematic errors from excited states [12]. As another example, consider F_{p4s}, the fictitious pseudoscalar decay constant with degenerate valence quarks of mass $m_{\mathrm{v}} = 0.4 m_s$ and physical sea-quark masses [12]. The value of the valence-quark mass is chosen to be heavy enough to make it not too expensive to compute the correlators, but light enough for chiral perturbation theory to apply. However, F_{p4s} has strong dependence on the valence-quark mass. Thus, relatively small errors in determining am_s, the physical value of the strange-quark mass in lattice units, may lead to significant errors in aF_{p4s} through the value of the valence mass, $am_{\mathrm{v}} = 0.4am_s$. Further, the required asymptotic fits to correlators are difficult to automate and usually require significant human intervention.

Gradient flow [65, 59] has received considerable attention [77, 57, 70, 3] over the past few years because it is a theoretically-grounded smoothing operation that is simple to implement and can be used to obtain precisely determined scales. The basis for scale setting with gradient flow is the determination of the flow time for which a dimensionless, precise, and easily-computable quantity is smoothed to a predefined value. The original quantity proposed by Lüscher, t_0, is defined through the gauge field energy density [58]. Most modifications focus on reducing discretization errors in the same underlying flow or observable [77, 31, 47, 43]. All of these scales can be easily computed to a statistical precision of 0.1% or less, and have small quark-mass and discretization dependence. Finite-volume effects, the only remaining source of systematic error for relative scale setting, may also be kept very small.

Here, we present our computation of the gradient-flow scales $\sqrt{t_0}/a$ and w_0/a on the MILC, (2+1+1)-flavor, highly improved staggered quark (HISQ) ensembles [11, 12]. The HISQ configurations used in this analysis cover lattice spacings from $a \approx 0.15$ to 0.06 fm and include ensembles with physical, or heavier than physical, light-quark masses, and physical, or lighter than physical, strange-quark mass. The charm-quark mass is kept near its physical value. We perform a continuum extrapolation and interpolation to physical quark masses of w_0F_{p4s} and $\sqrt{t_0}F_{p4s}$ to determine the two scales in physical units, using our previous determination of F_{p4s} in physical units [13]. We find $\sqrt{t_0} = 0.1416(^{+8}_{-5})$ fm and $w_0 = 0.1717(^{+12}_{-11})$ fm, where statistical and all systematic errors have been added in quadrature.

We start with a review of the relevant theoretical details, including the gradient-flow equation in Sec. 2.2.1, definitions of the scales t_0 and w_0 in Sec. 2.2.2, chiral perturbation theory for flow quantities in Sec. 2.2.2, and lattice-spacing dependence in Sec. 2.2.2. The computational setup is described in Sec. 2.3.1. We discuss the raw lattice results in Sec. 2.3.2, include a brief comparison of the results for different ensemble-generation algorithms in Sec. 2.3.2, and estimate the integrated autocorrelation lengths in Sec. 2.3.2. Adjustments for charm-quark mass mistuning are performed in Sec. 2.3.2, and a simple extrapolation to the continuum of the results on the physical-mass ensembles is presented in Sec. 2.3.2. Section 2.3.3 then describes the quark-mass interpolation and continuum extrapolation. We present our results for w_0 and $\sqrt{t_0}$ in physical units in Sec. 2.4.1, and include comparisons with our earlier preliminary results. The continuum mass dependence of w_0 is deduced from our fits in Sec. 2.4.2 and used to compare the scales determined from w_0 to those determined from F_{p4s} in Ref. [13]; knowing the continuum mass dependence will be useful in determining the scales of new ensembles. Section 2.5 compares our results to those of other collaborations, and tabulates the precision of various methods for relative scale-setting.

Preliminary versions of this analysis have been described in Refs.[16] and [15].

2.2 Theoretical Framework

This section summarizes the theoretical details of gradient flow from Refs. [65, 59, 58, 31, 43, 8] that are relevant to the scale-setting analysis in later sections.

2.2.1 Diffusion Equation

Gradient flow [65, 59] is a smoothing of the original gauge fields A towards stationary points of the action S. The new, smoothed gauge fields $B(t)$ are functions of the 'flow time' t and are updated according to the diffusion-like equation below, where g_0 is the bare coupling.

$$\begin{aligned} \frac{dB_\mu}{dt} &= -g_0^2 \frac{\partial S}{\partial B_\mu} = D_\nu G_{\nu\mu} \,, \quad B_\mu(0) = A_\mu \,, \\ D_\nu X &= \partial_\nu X + [B_\nu,\, X] \ , \quad G_{\nu\mu} = \partial_\nu B_\mu - \partial_\mu B_\nu + [B_\nu,\, B_\mu] \ . \end{aligned} \tag{2.1}$$

On the lattice, the Yang Mills action is replaced by an appropriate discretized version. The gauge links $V(t)$ are updated in time according to

$$\frac{dV(t)_{i,\mu}}{dt} = -g_0^2 \frac{\partial S(V)}{\partial V_{i,\mu}} V_{i,\mu} \,, \quad V_{i,\mu}(0) = U_{i,\mu} \tag{2.2}$$

The change of $V(t)$ with flow time explicitly follows the steepest descent of the action with respect to the gauge field, with an additional factor of $V_{i,\mu}$ in the lattice formulation to ensure gauge covariance. For more details on the SU(3)-valued derivative, see the Appendix of Ref. [58].

As the flow time t increases, the gauge fields diffuse and short-distance lattice artifacts are removed. After modifying the flow equation with a flow-time-dependent gauge transformation of the field one can explicitly see the suppression of high momenta in the leading-order perturbative expansion of the gauge field in powers of the coupling g_0 [58]:

$$B_\mu(x,t) \approx \frac{1}{(4\pi t)^2} \int d^4y A_\mu(y) e^{-(x-y)^2/(4t)} \ , \quad \tilde{B}_\mu(p,t) \approx \tilde{A}_\mu(p) e^{-tp^2} \ . \tag{2.3}$$

The expansion also shows that the kernel in position space smooths over a sphere of root-mean-square radius $\sqrt{8t}$, which implies that discretization effects will be highly suppressed for flow times $t \gg a^2/8$.

2.2.2 Gradient-flow Scales

The process of gradient flow introduces a dimensionful, independent variable, the flow time. Since all quantities calculated from smoothed gauge links will be functions of the flow time, one may define a scale by choosing a reference time at which a chosen dimensionless quantity reaches a predefined value. If the dimensionless quantity is also finite in the continuum limit, then the reference time scale will be independent of the lattice spacing up to discretization corrections in powers of a^2. One of the easiest, dimensionless quantities to calculate with only gauge fields is the average total energy within a smoothed volume $V \propto t^2$. This is equivalent to calculating the product of the energy density and squared flow time $t^2\langle E(t)\rangle$. Lüscher has shown that the energy density is finite to all orders (when expressed in terms of renormalized quantities) [60], so $t^2\langle E(t)\rangle$ is a suitable candidate for setting the scale. A fiducial point c is chosen, and the reference scale is defined to be the flow time t_0 where

$$t_0^2\langle E(t_0)\rangle = c \ . \tag{2.4}$$

The fiducial point should be chosen so that for simulated lattice spacings a and volumes $V = L^3T$ (with $T \geq L$), the reference timescale t_0 falls between $a \ll \sqrt{8t_0} \ll aL$. The value of $c = 0.3$ has been found, empirically, to satisfy this relation [58, 31]. A larger fiducial point of $c = 2/3$ has also been proposed in order to reduce discretization errors, at the expense of somewhat larger finite-volume effects [77].

The renormalized expansion of $\langle E(t)\rangle$ to second order in g shows $t^2\langle E(t)\rangle$ is approximately constant [58]. For small flow times this agrees with computational results, but for larger flow times (including the scale t_0) $t^2\langle E(t)\rangle$ is found empirically to be linear in t [58, 31]. The transition of $\langle E(t)\rangle$ from t^{-2} to t^{-1} dependence is nonperturbative. However, we expect discretization errors to enter primarily for small flow times, before the lattice details are smoothed away. In accordance with this expectation, empirical evidence suggests that discretization effects have little impact on the slope of $t^2\langle E(t)\rangle$ at larger times [31]. Assuming the property is general, an improvement to the scale t_0 is computed by considering the slope:

$$\left[t\frac{d}{dt}t^2\langle E(t)\rangle\right]_{t=w_0^2} = c \ , \tag{2.5}$$

where w_0 is the improved scale. Again, the value of the fiducial point $c = 0.3$ or $c = 2/3$ is chosen to avoid discretization and finite volume effects.

Chiral Perturbation Theory

Because both scales t_0 and w_0 are defined in terms of the energy density $\langle E(t)\rangle$, and the energy density is a local, gauge-invariant quantity, chiral perturbation theory can be applied to determine the quark-mass dependence of the scales. This is an advantage over some other scales, such as r_0 or r_1, for which no chiral perturbation theory expansion is available. The mapping of $\langle E(t)\rangle$ to the chiral effective theory has been carried out by Bär and Golterman in Ref. [8]. The expansion for $\sqrt{t_0}$ in the $N_f = 2+1$ case in terms of the pion and kaon mass is

$$\begin{aligned}\sqrt{t_0} &= \sqrt{t_{0,ch}}\left[1 + k_1\frac{2M_K^2+M_\pi^2}{(4\pi f)^2}\right.\\ &+ \frac{1}{(4\pi f)^2}\left((3k_2-k_1)M_\pi^2\mu_\pi + 4k_2M_K^2\mu_K + \frac{1}{3}k_1(M_\pi^2-4M_K^2)\mu_\eta + k_2M_\eta^2\mu_\eta\right)\\ &+ \left. k_4\frac{(2M_K^2+M_\pi^2)^2}{(4\pi f)^4} + k_5\frac{(M_K^2-M_\pi^2)^2}{(4\pi f)^4}\right] \ ,\end{aligned} \tag{2.6}$$

where $t_{0,ch}$ is the value of t_0 in the chiral limit, the chiral logarithms are represented with the shorthand $\mu_Q = (M_Q/4\pi f)^2 \log{(M_Q/\mu)^2}$, and the k_i are low energy constants (LECs) that depend on the flow time. Note that chiral logarithms enter only at next-to-next-to-leading order (NNLO). The scale w_0 has the same expansion form to NNLO, but with different coefficients k_i. This is because the flow-time dependence of $\langle E(t)\rangle$ appears only in the LECs, allowing the differences between Eqs. (2.4) and (2.5) to be absorbed into redefinitions of the LECs.

One can generalize Eq. (2.6) to staggered chiral perturbation theory in order to explicitly take into account discretization effects from staggered taste-symmetry violations. In this paper, however, we have used simple polynomial expansions to parametrize lattice-spacing effects. There are two reasons for this choice. First, the quark-mass dependence of the gradient-flow scales is already small, as will be evident in Sec. 2.4.2, and nontrivial staggered effects would come in only with the chiral logarithms, which are of NNLO. For HISQ quarks, such effects are very small. Second, the number of undetermined coefficients in staggered chiral perturbation theory expansions would be too large in comparison to the number of independent data points available for interpolations. Unlike analyses of pseudoscalar masses or decay constants, here we have no valence quarks whose masses could be varied to increase the size of the dataset.

Discretization Effects

In determining the scales t_0 and w_0, lattice artifacts enter in three places: the action used to generate the initial configurations, the action of the gradient flow, and the choice of observable. Because ensemble generation is expensive, the action chosen for generating the gauge configurations is fixed in practice. Therefore, we only consider improvements to the gradient flow and energy density.

Empirical results suggest partial improvements of the flow or the energy density can yield smaller $O(a^2)$ terms. By using the tree-level improved Symanzik action instead of the Wilson action in the flow, the BMW collaboration found smaller cutoff effects for both gradient-flow scales on their Wilson-clover ensembles with 2-HEX smearing (with scale set by M_Ω) [31]. Similarly, using the symmetric, cloverleaf definition of the field strength tensor $G_{\mu\nu}$ in $\langle E\rangle = G_{\mu\nu}G_{\mu\nu}/4$, instead of the simpler sum over the plaquettes, yielded cutoff effects in $\sqrt{t_0}/r_0$ that were five times smaller [58]. Of course, applying partial improvements at different steps is not guaranteed to produce smaller cutoff effects in the final result. Also, for each case, the lattice-spacing dependence of the gradient-flow scale cannot be cleanly separated in the numerical results from the dependence of the additional quantity used to set the scale in the extrapolation to the continuum.

A detailed examination of the discretization effects on gradient-flow scales has been recently carried out in Ref. [43]. The net lattice-spacing dependence from all three stages of the calculation (dynamical action, flow, and observable) is determined at tree level in the gauge coupling from a calculation of $\langle E(t)\rangle$ at finite lattice spacing. For the clover observable chosen in this study

$$F(t) \equiv \langle t^2 E(t)\rangle = \frac{3(N^2-1)g_0^2}{128\pi^2}\left(1 + C_2 a^2/t + O(a^4/t^2) + O(g_0^2)\right) , \tag{2.7}$$

$$C_2 = 2c_f + \frac{2}{3}c_g - \frac{1}{24} , \tag{2.8}$$

where the coefficient c_f describes the gradient-flow action, and c_g describes the original gauge action used to generate the ensembles [43]. For our choices of Symanzik one-loop-improved gauge action ($c_g = -1/12$ at tree level) and Symanzik tree-level gradient flow ($c_f = -1/12$), we have $C_2 = -19/72$. Unfortunately, our choices of actions and observable lead to larger tree-level discretization terms than from many other combinations of common choices of action for the flow and observable. For more detail see Table 1 in Ref. [43].

Utilizing the known a^2 dependence of $F(t)$, improved scales are defined in Ref. [43] by cancelling the tree-level contributions to $F(t)$ in the implicit definitions of t_0 and w_0.

$$\left[\frac{t^2\langle E(t)\rangle}{(1 + C_2(a^2/t) + C_4(a^2/t)^2 + ...)}\right]_{t=t_{0,\text{imp}}} = c\ . \tag{2.9}$$

$$\left[t\frac{d}{dt}\frac{t^2\langle E(t)\rangle}{(1 + C_2(a^2/t) + C_4(a^2/t)^2 + ...)}\right]_{t=w^2_{0,\text{imp}}} = c \tag{2.10}$$

We compute the improved scales t_{imp} and w_{imp} and compare to the a^2 dependence of the unimproved scales in Sec. 2.3.2. An additional theoretical handle on the comparison can be made by expanding the unimproved scales directly as a power series in a^2 and calculating the coefficients.

The lattice-spacing dependence of the gradient-flow scales are proportional to C_2 and depend on the continuum flow-time dependence of $F(t)$ and its derivatives $F'(t) = t\frac{d}{dt}F(t)$ and $F''(t) = t^2\frac{d^2}{dt^2}F(t)$ evaluated at the corresponding scale $t = t_{0,cont}$ or $t = w^2_{0,cont}$. The next-to-leading-order coefficients are given by

$$t_0 = t_{0,cont}\left(1 - T_2\frac{a^2}{t_{0,cont}}\right), \qquad T_2 = C_2\frac{F}{F'} \approx -0.3568(2)\ , \tag{2.11}$$

$$w_0^2 = w^2_{0,cont}\left(1 - W_2\frac{a^2}{w^2_{0,cont}}\right), \qquad W_2 = C_2\frac{F' - F}{F'' + F'} \approx 0.070(2)\ , \tag{2.12}$$

Note that the coefficients T_2 and W_2 are identical to those derived for the improved scale in Ref. [43]; however, the a^2 coefficients in the above expression are $-T_2$ and $-W_2$ because Eq. (2.11) relates the (unimproved) scales at finite lattice spacing to the continuum scales. The numerical evaluation of F, F', and F'' for the estimates of T_2 and W_2 has been performed on the $a \approx 0.06$ fm, physical quark-mass ensemble (see Table 2.1). No systematic errors are included in these estimates; however, the systematic errors are expected to be small. We may conclude that $|T_2| > |W_2|$. This result, combined with the fact that the ratio W_2/T_2 depends only on the continuum flow-time dependence of F, implies that w_0^2 will have smaller tree-level discretization errors than t_0 for any choice of dynamical action, flow, and discretization of $\langle E \rangle$.

2.3 Details of the Computation

We compute the scales $\sqrt{t_0}/a$, w_0/a, $\sqrt{t_{0,\mathrm{imp}}}/a$, and $w_{0,\mathrm{imp}}/a$ on the MILC $N_f = 2+1+1$ HISQ ensembles [12, 11]. Tables 2.1 and 2.2 list the parameters and relevant observables for ensembles with the strange sea-quark mass tuned near its physical value, and well below its physical value, respectively. Table 2.3 gives the values of aF_{p4s} at physical quark masses and associated lattice spacings, which are needed for continuum extrapolations. The lattice spacings are calculated with a mass-independent scale-setting scheme; the continuum value $F_{p4s} = 153.90(9)(^{+21}_{-28})$MeV is taken from Ref. [13], where f_π was used to set the absolute scale. Physical values of am_s at each lattice spacing [13] are also tabulated. Using the physical quark-mass ratio $m_c/m_s = 11.747(19)(^{+59}_{-43})$ [13], these values of am_s determine values of the physical charm-quark mass for each ensemble in lattice units, which in turn will be used to adjust for mistunings of the charm sea-quark mass in Sec. 2.3.2. Finally, Table 2.3 lists the

β	$\approx a(\mathrm{fm})$	m_l'/m_s'	am_s'	am_c'	$N_s^3 \times N_t$	aM_π	aM_K
5.80	0.15	1/5	0.0650	0.838	$16^3 \times 48$	0.23653(22)	0.40261(25)
5.80	0.15	1/10	0.0640	0.828	$24^3 \times 48$	0.16614(10)	0.38067(16)
5.80	0.15	1/27	0.0647	0.831	$32^3 \times 48$	0.10180(09)	0.37093(16)
6.00	0.12	1/5	0.0509	0.635	$24^3 \times 64$	0.18917(15)	0.32358(20)
6.00	0.12	1/10	0.0507	0.628	$32^3 \times 64$	0.13424(09)	0.30813(15)
6.00	0.12	1/10	0.0507	0.628	$40^3 \times 64$	0.13400(06)	0.30821(09)
6.00	0.12	1/27	0.0507	0.628	$48^3 \times 64$	0.08153(04)	0.29851(11)
6.30	0.09	1/5	0.0370	0.440	$32^3 \times 96$	0.14055(17)	0.24061(18)
6.30	0.09	1/10	0.0363	0.430	$48^3 \times 96$	0.09852(08)	0.22688(12)
6.30	0.09	1/27	0.0363	0.432	$64^3 \times 96$	0.57215(04)	0.21946(09)
6.72	0.06	1/5	0.0240	0.286	$48^3 \times 144$	0.09438(16)	0.16191(16)
6.72	0.06	1/10	0.0240	0.286	$64^3 \times 144$	0.06713(06)	0.15452(09)
6.72	0.06	1/27	0.0220	0.260	$96^3 \times 192$	0.03887(03)	0.14269(06)

Table 2.1: HISQ ensembles with near-physical strange sea-quark mass. The first three columns list the gauge coupling constant, the approximate lattice spacing, and the ratio of light-to-strange sea-quark mass. The fourth and fifth column list the strange and charm sea-quark mass, respectively. (Quark masses with primes indicate simulation values of the ensemble, whereas unprimed masses indicate physical values.) All but two ensembles can be uniquely identified by the second and third columns. To differentiate between the two $a \approx 0.12$ fm, $m_l'/m_s' = 1/10$ ensembles we use the dimensions of the lattice, $N_s^3 \times N_t$, given in column six. The last two columns give the taste-Goldstone pion and kaon masses in lattice units.

effective coupling constant α_s calculated from taste violations of the HISQ pions in Ref. [13]. The couplings are then scaled by a constant so that $\alpha_s = \alpha_V(q^* = 1.5/a)$ for $\beta = 5.8$, where α_V is determined from the plaquette [40, 12].

2.3.1 Computational Setup

We solve the gradient-flow differential equation numerically using the Runga-Kutta algorithm generalized to SU(3) matricies, as originally proposed by Lüscher [58]. The routine discretizes the flow time with a step size ϵ and computes the gauge configuration at a later flow time $t = n\epsilon$ by iterating from the initial gauge configuration. The total error of the integration

$\approx m'_l/m_s$	$\approx m'_s/m_s$	am'_c	$N_s^3 \times N_t$	aM_π	aM_K
0.10	0.10	0.628	$32^3 \times 64$	0.13181(10)	0.13181(10)
0.10	0.25	0.628	$32^3 \times 64$	0.13250(09)	0.17385(11)
0.10	0.45	0.628	$32^3 \times 64$	0.13275(10)	0.21719(12)
0.10	0.60	0.628	$32^3 \times 64$	0.13324(10)	0.24509(13)
0.175	0.45	0.628	$32^3 \times 64$	0.17491(10)	0.23199(12)
0.20	0.60	0.635	$24^3 \times 64$	0.18850(17)	0.26382(18)
0.25	0.25	0.640	$24^3 \times 64$	0.20903(19)	0.20903(19)

Table 2.2: HISQ ensembles with a lighter-than-physical strange sea-quark mass. All ensembles have gauge coupling constant $\beta = 6.00$ and lattice spacing $a \approx 0.12$ fm. The first two columns list the approximate values of the light sea-quark mass m'_l and strange sea-quark mass m'_s in units of the physical strange-quark mass m_s. All of the ensembles may be uniquely identified by these two columns. The remaining columns are equivalent to those in Table 2.1.

β	$\approx a$(fm)	am_s	aF_{p4s}	a(fm)	α_s
5.80	0.15	$0.06863(^{+53}_{-39})$	0.119376(36)	$0.15305(^{+57}_{-41})$	0.58801
6.00	0.12	$0.05304(^{+41}_{-30})$	0.095403(28)	$0.12232(^{+45}_{-33})$	0.53796
6.30	0.09	$0.03631(^{+29}_{-21})$	0.068570(19)	$0.08791(^{+33}_{-24})$	0.43356
6.72	0.06	$0.02182(^{+17}_{-13})$	0.044237(12)	$0.05672(^{+21}_{-16})$	0.29985

Table 2.3: Values of am_s, aF_{p4s}, a (in fm), and α_s adjusted to physical values of the quark masses, for various couplings β. All results are from the analysis presented in Ref. [13]. The first two columns list the gauge coupling and approximate lattice spacing. The next two columns list the strange mass and F_{p4s} in lattice units. The lattice spacing from $F_{p4s} = 153.90(9)(^{+21}_{-28})$ MeV, in a mass independent scheme, is listed in the fifth column. The final column tabulates the strong coupling constant α_s determined from the taste splittings (see text). For am_s and a, the errors are the sum in quadrature of statistical and systematic errors. Only statistical errors are shown for aF_{p4s}.

up to flow time t scales like ϵ^3. For all ensembles analyzed in this paper, we find that the scales $\sqrt{t_0}/a$ and w_0/a determined at a step size of $\epsilon = 0.07$ cannot be differentiated, within statistical errors, from those at $\epsilon = 0.03$. We therefore consider $\epsilon = 0.03$ to be a conservative step size, and employ it for all results presented below.

Both the Wilson and Symanzik tree-level action for the gradient flow have been implemented and are publicly available in the current release of the MILC code [62]. This computation uses the Symanzik tree-level improved action in the gradient flow and the symmetric, cloverleaf definition of the field-strength tensor $G_{\mu\nu}$ in $\langle E\rangle = G_{\mu\nu}G_{\mu\nu}/4$.

2.3.2 Measurements of Gradient-Flow Scales

Tables 2.4 and 2.5 show the results for $\sqrt{t_0}/a$, w_0/a, $\sqrt{t_{0,\mathrm{imp}}}/a$, and $w_{0,\mathrm{imp}}/a$ on the HISQ ensembles. The scales $\sqrt{t_{0,\mathrm{imp}}}/a$ and $w_{0,\mathrm{imp}}/a$ were improved to $O(a^8)$ at tree level using Eq. (2.9) and the coefficients calculated in Ref. [43] for Symanzik-Symanzik-Clover. For the ensembles with the smallest lattice volumes, all configurations are included in the computation. As the volumes and cost become larger, a fraction of the configurations are run. The configurations in each subset are spaced uniformly across the ensembles, with spacings chosen to help reduce autocorrelations. The total number of generated configurations, number of configurations in the gradient-flow calculation, and molecular-dynamics time separation between the included configurations are also tabulated for each ensemble in Tables 2.4 and 2.5.

The error shown with each scale is statistical. It is determined by performing a jackknife analysis over the included subset of configurations in each ensemble. The jackknife bin size is set to be at least twice the integrated autocorrelation length of the energy density, which is

$\approx a$(fm)	m_l'/m_s'	N_{sim}/N_{gen}	τ	$\sqrt{t_0}/a$	w_0/a	$\sqrt{t_{0,\text{imp}}}/a$	$w_{0,\text{imp}}/a$
0.15	1/5	1020/1020	5	1.1004(05)	1.1221(08)	0.9857(04)	1.1069(10)
0.15	1/10	1000/1000	5	1.1092(03)	1.1381(05)	0.9932(02)	1.1258(06)
0.15	1/27	999/1000	5	1.1136(02)	1.1468(04)	0.9969(02)	1.1361(04)
0.12	1/5	1040/1040	5	1.3124(06)	1.3835(10)	1.2003(05)	1.3870(11)
0.12	1/10 ($32^3 \times 64$)	999/1000	5	1.3228(04)	1.4047(09)	1.2100(04)	1.4096(09)
0.12	1/10 ($40^3 \times 64$)	1000/1028	5	1.3226(03)	1.4041(06)	1.2098(03)	1.4089(06)
0.12	1/27	34/999	140	1.3285(05)	1.4168(10)	1.2152(05)	1.4225(11)
0.09	1/5	102/1011	$50, 60$	1.7227(08)	1.8957(15)	1.6280(08)	1.9053(16)
0.09	1/10	119/1000	36	1.7376(05)	1.9299(12)	1.6423(05)	1.9406(12)
0.09	1/27	67/1031	$32, 48$	1.7435(05)	1.9470(13)	1.6478(05)	1.9583(13)
0.06	1/5	127/1016	48	2.5314(13)	2.8956(33)	2.4618(12)	2.9049(33)
0.06	1/10	38/1166	96	2.5510(14)	2.9478(31)	2.4810(14)	2.9582(30)
0.06	1/27	49/583	48	2.5833(07)	3.0119(19)	2.5133(07)	3.0223(19)

Table 2.4: Values of the gradient-flow scales on the physical strange-quark HISQ ensembles listed in Table 2.1. The first two columns are the approximate lattice spacing and ratio of light to strange sea-quark mass, with the lattice dimensions appended as needed to identify each ensemble uniquely. The next column shows the ratio of number of configurations included in the gradient-flow calculation to the number of configurations in the ensemble. The fourth column lists the molecular-dynamics time separation τ between configurations included in the gradient-flow calculation. Multiple values are listed for cases where independent streams of the same ensemble did not use the same τ.

m_l'/m_s	m_s'/m_s	N_{sim}/N_{gen}	τ	$\sqrt{t_0}/a$	w_0/a	$\sqrt{t_{0,\text{imp}}}/a$	$w_{0,\text{imp}}/a$
0.10	0.10	102/1020	20	1.3596(06)	1.4833(13)	1.2441(06)	1.4932(13)
0.10	0.25	204/1020	20	1.3528(04)	1.4676(10)	1.2378(04)	1.4764(10)
0.10	0.45	205/1020	20	1.3438(05)	1.4470(10)	1.2296(05)	1.4544(11)
0.10	0.60	107/1020	20	1.3384(08)	1.4351(16)	1.2247(07)	1.4418(17)
0.175	0.45	133/1020	20	1.3385(05)	1.4349(13)	1.2248(05)	1.4415(14)
0.20	0.60	255/1020	20	1.3297(06)	1.4170(12)	1.2166(06)	1.4225(12)
0.25	0.25	255/1020	20	1.3374(07)	1.4336(14)	1.2236(06)	1.4402(15)

Table 2.5: Values of the gradient-flow scales on the HISQ lighter-than-physical strange-quark ensembles listed in Table 2.2. The first two columns are identical to those in Table 2.2 and used to identify the ensembles. The latter six columns are equivalent to those in Table 2.4.

determined in Sec. 2.3.2. In many cases the bin size is larger than would be naively estimated by increasing the bin size until the statistical error plateaus, which is further evidence for the conservative nature of our estimates of autocorrelation lengths.

Considering the low cost and ease of computation, we originally intended to analyse all configurations from the HISQ ensembles. However, the desired statistical accuracy is often reached well before an entire ensemble is analyzed, and the cost, although low compared to configuration generation, is significant enough that analyzing all configurations would be an inefficient use of resources at present. If higher-precision scales are needed in the future, it would be straightforward to complete the analysis on the full ensembles.

Comparison of RHMC and RHMD

As discussed in Ref. [12], two generation algorithms were employed for the HISQ ensembles: rational hybrid Monte Carlo (RHMC) and molecular dynamics (RHMD). As a check of the consistency of these two algorithms, we compute the ratio of w_0 computed on RHMC-generated configurations divided by w_0 computed on RHMD-generated configurations for the same bare gauge coupling and quark masses. For $a \approx 0.09$fm, $m_l'/m_s' \approx 1/27$, the ratio is $w_0^{RHMC}/w_0^{RHMD} = 1.0009(12)$. For $a \approx 0.06$fm, $m_l'/m_s' \approx 1/10$, the ratio is $w_0^{RHMC}/w_0^{RHMD} = 1.0002(26)$. For some configuration streams the pattern of fluctuations of w_0/a with molecular-dynamics time is not sufficient to reliably estimate the mean and standard deviation over that single stream. However, in the particular cases used for calculating the ratio, this issue is not evident. Figure 2.1 shows the fluctuations of the relevant streams for each ratio. For all fluctuations of w_0/a on a single stream, the length of the fluctuation in molecular-dynamics time units is small compared to the entire molecular-dynamics time span of the stream.

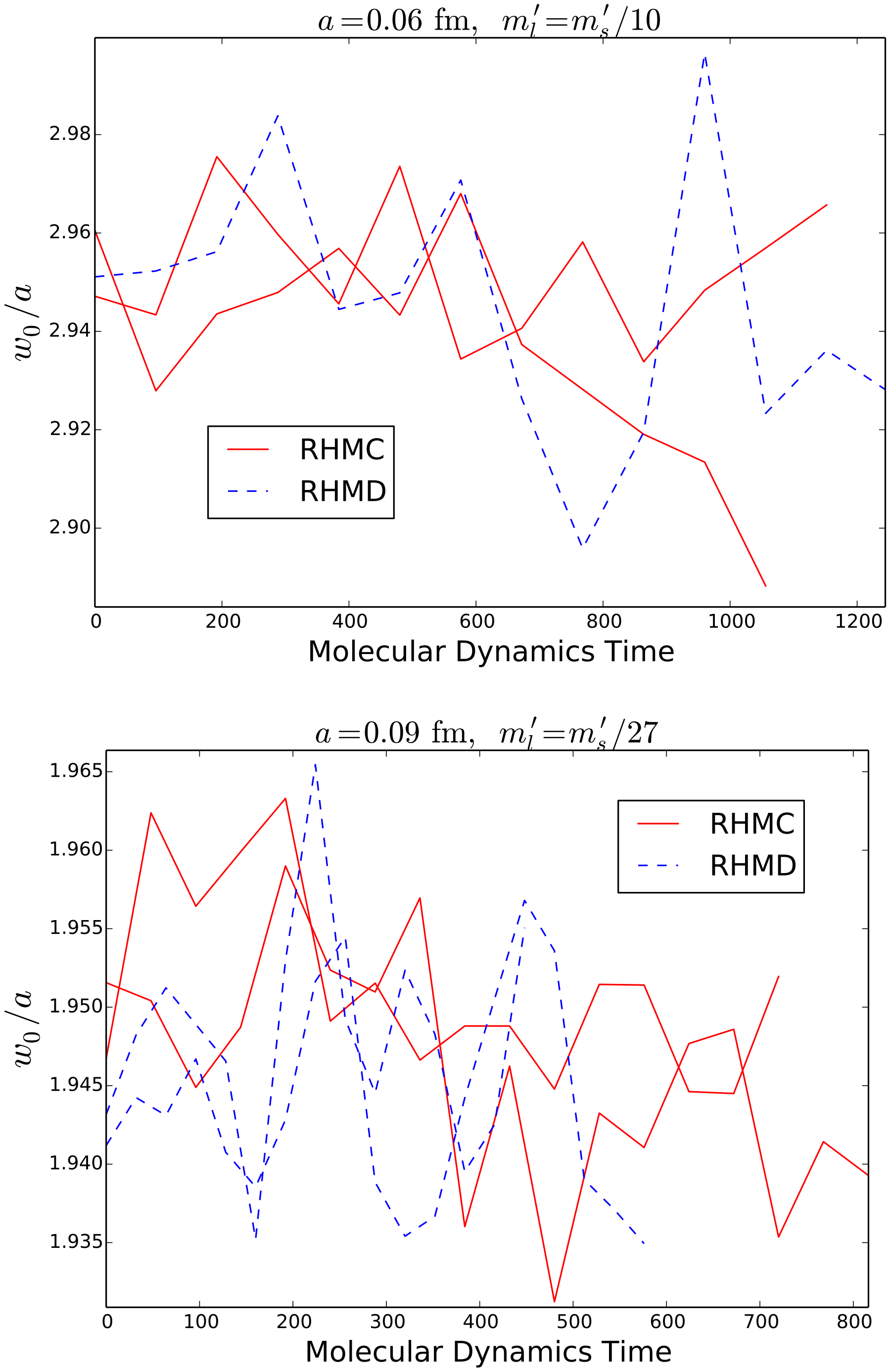

Figure 2.1: The scale w_0/a measured on individual configurations as a function of the simulation time in molecular-dynamics time units. Configuration streams generated with RHMC and RHMD are respresented by solid-red and dashed-blue lines, respectively.

Autocorrelation Lengths

We compute the autocorrelation function of $\langle E(t,\tau)\rangle$ as a function of the flow time t and the number, τ, of molecular-dynamics time units separating configurations. For the ensembles at $a \approx 0.15$ and 0.12 fm, where the full ensembles have been analyzed, we have a reliable estimate of the statistical error of the autocorrelation function for all values of τ. For the finer lattice spacings $a \approx 0.09$, 0.06 fm, estimating the autocorrelation functions for τ smaller than the separations listed in Tables 2.4 and 2.5 is impossible without calculating the gradient flow on more configurations. To address this, we have analyzed an additional 50 and 25 equilibrated configurations separated with $\tau = 24$ from the $m_l'/m_s = 1/10$, $a = 0.09$ and 0.06 fm ensembles, respectively. Most of these configurations are not included in the calculation of the gradient-flow scales; we keep the configurations used for computing the scales uniformly spread over each ensemble, with constant separation in τ.

Once the autocorrelation function of $\langle E(t)\rangle$ is computed, we integrate the function over the separation τ for each step in flow time t. To estimate the statistical error in the integrated autocorrelation length l, we first perform a jackknife analysis with a bin size b large enough that the statistical error is (approximately) unchanged with further increases in bin size. After determining l and estimating all errors, we take the upper bound as a conservative estimate of l and recalculate the statistical error with bin size $b \geq 2l$. If the new statistical error estimate leads to a new, higher upper bound l that does not satisfy $b \geq 2l$, then this procedure is repeated until the condition is met.

Since the autocorrelation function at fixed t exponentially decays to 0 as a function of τ, and the statistical noise increases with τ, an integration region is chosen to cut off noise from autocorrelations at larger τ. This produces an additional systematic error associated with the value of τ_{cut}, the cut-off in τ of the integration region. The systematic error from τ_{cut} is estimated by the difference between the largest and smallest integrated autocorrelation lengths for different 'reasonable' choices of the cutoff. A choice of the cutoff τ_{cut} is deemed reasonable if two criteria are met. First, τ_{cut} must be larger than or equal to the first separation $\tau_{\min}$ for which the autocorrelation function is within $2\sigma_{stat}$ of 0. This ensures the largest contributions to the integrated autocorrelation function are included in the integration. Second, the deviation between the integrated autocorrelation length on the full dataset and those on the first or second halves is no more than $3\sigma_{stat}$. This criterium is designed to eliminate situations where the number of data points are so few that the sample mean and sample standard deviation are not meaningful. Figure 2.2 illustrates the procedure for estimating the integration error on the $a\approx 0.12$ fm, $m_l/m_s = 0.1$ ensemble. To choose τ_{cut} for the central estimate of the integrated autocorrelation length, the integration region yielding the largest integrated autocorrelation length is chosen. With this choice our estimates cover the worst-case scenario, consistent with using the upper bound as a conservative estimate. However, we emphasize that our statistics are limited on the finer ensembles, and it is possible that there longer-range tails to the autocorrelation function that we are not including in the integrated autocorrelation length.

The integrated autocorrelation lengths with statistical and systematic error combined in quadrature are plotted in Fig. 2.3. Notice the autocorrelation length for $\langle E(t)\rangle$ appears to asymptotically increase for increasing flowtimes, as expected for a smoothing operation. The top of the range for our estimate of the integrated autocorrelation length at large flow times is 58 and 65 molecular dynamics time units for the $a \approx 0.09$fm and $a \approx 0.06$fm, $m_l' = m_s'/10$, physical strange-quark mass ensembles, respectively. In comparison, the integrated autocorrelation length of the topological charge appears to be roughly 40 and 300 molecular dynamics time units for the $a \approx 0.09$fm and $a \approx 0.06$fm, $m_l' = m_s'/5$, physical strange-quark

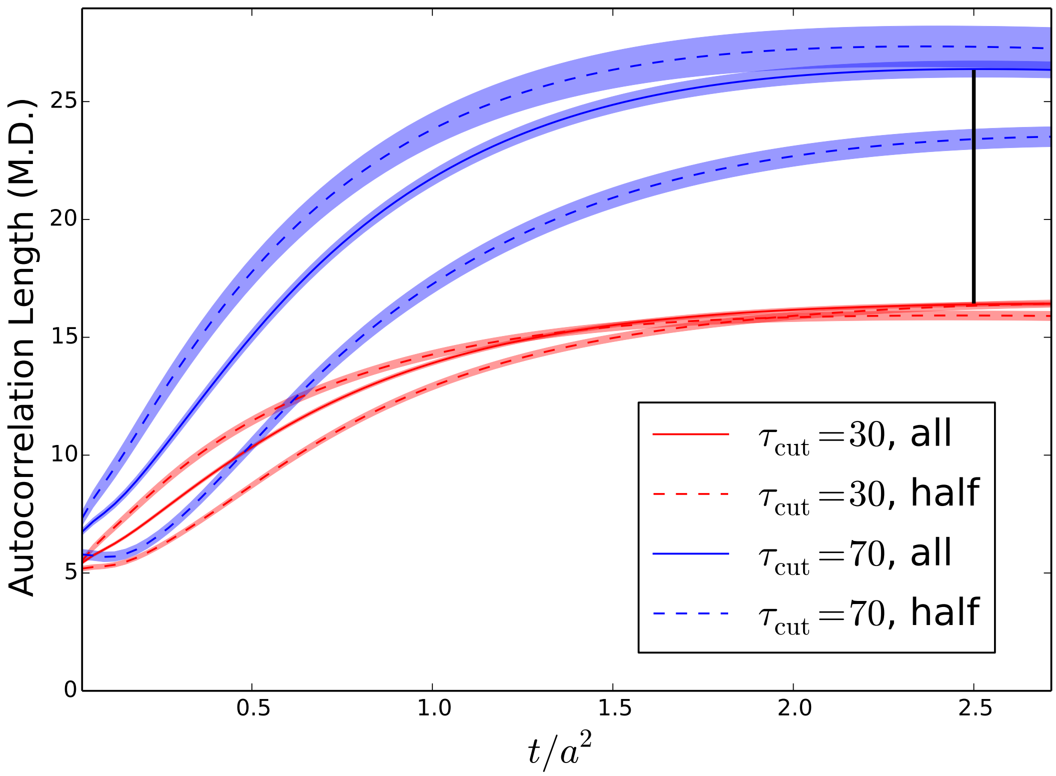

Figure 2.2: The jackknifed, integrated autocorrelation length versus flow time for different choices of the upper limit of the integration region, τ_{cut}. The data is from the $a \approx 0.12$ fm, $m_l'/m_s' = 0.1$ ensemble. The plots correspond to the largest and smallest 'reasonable' values of τ_{cut}, where the statistical errors are not completely uncontrolled. The autocorrelation lengths for the full ensemble (solid line) and for the two halves of the ensemble (dashed lines) are plotted for each value of τcut. The estimate of the systematic error associated with the the choice of τ_{cut} for large flow times is denoted by the black vertical line.

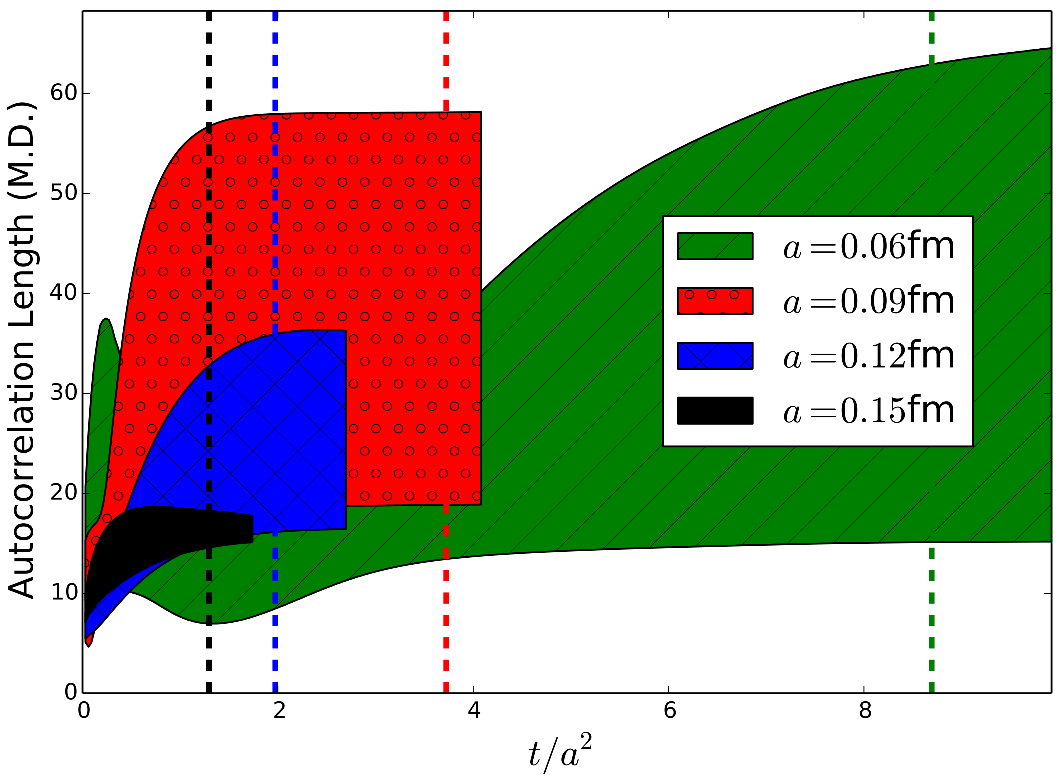

Figure 2.3: The integrated autocorrelation length (in molecular-dynamics time units) as a function of flow time for ensembles with $m_l'/m_s' = 0.1$ and different lattice spacings. The thickness of the colored regions show the full range of the $1\,\sigma$ errors, obtained by adding, in quadrature, the statistical error and systematic variation with τ_{cut}. Dashed vertical lines denote the flow times that determine w_0 on each ensemble where the color of the line matches the color of the shaded region. The range of flow times that determine $\sqrt{t_0}$ is similar.

mass ensembles [12]. This suggests the autocorrelation length for $\langle E(t) \rangle$ at large flowtimes is comparable to or smaller than the autocorrelation length of the topological charge.

Charm-Quark Mass Mistuning

Mistunings of the charm-quark mass on our ensembles vary between 1% and 11%. It is therefore important to account for the leading-order corrections in the charm-quark mass to the quantities we consider. Given any low-energy quantity Q that is proportional to a power p of $\Lambda^{(3)}_{\text{QCD}}$ in the effective three-flavor low-energy theory, the leading order heavy-quark mass dependence can be determined using the relation between $\Lambda^{(3)}_{\text{QCD}}$ and $\Lambda^{(4)}_{\text{QCD}}$ of the four-flavor theory [25, 26].

$$\Lambda^{(3)}_{\text{QCD}} = \Lambda^{(4)}_{\text{QCD}} \left(\frac{m_c}{\Lambda^{(4)}_{\text{QCD}}} \right)^{2/27} \qquad \frac{\partial \Lambda^{(3)}_{\text{QCD}}}{\partial m_c} = \frac{2}{27} \frac{\Lambda^{(3)}_{\text{QCD}}}{m_c} \tag{2.13}$$

For a pedagogical discussion see Ref. [61]. Equation (2.13) neglects discretization errors and physical $1/m_c$ corrections, and assumes that the lattice scale-setting procedure is independent of the heavy-quark mass. Given $Q = k(\Lambda^{(3)}_{\text{QCD}})^p$, where k and p are independent of m_c, the partial derivative of Q with respect to m_c at leading order is then

$$\frac{\partial Q}{\partial m_c} = kp \left(\Lambda^{(3)}_{\text{QCD}} \right)^{p-1} \frac{\partial \Lambda^{(3)}_{\text{QCD}}}{\partial m_c} = \frac{2p}{27} \frac{Q}{m_c} \ . \tag{2.14}$$

For a dimensionless ratio, such as F_{p4s}/w_0^{-1}, where both dimensional quantities share the same power p, the leading-order dependence on m_c will cancel. Higher-order corrections are negligible, even at the level of precision of this work. However, dependence on m_c will remain for dimensionless ratios where the powers are not identical because there is leading dependence on the light-quark masses. For example, M_π/F_{p4s}, where $M_\pi \propto (m_l \Lambda^{(3)}_{\text{QCD}})^{(1/2)}$, must be adjusted by $(-2/54)(M_\pi/F_{p4s})(\delta m_c/m_c)$ where δm_c is the mistuning of the charm-quark mass.

In this work, we must adjust six quantities: aM_π, aM_K, $\sqrt{t_0}/a$, w_0/a, $\sqrt{t_{0,\text{imp}}}/a$, and $w_{0,\text{imp}}/a$. When performing the continuum extrapolations and physical-mass interpolation, all six quantities will be scaled by aF_{p4s}. The products of aF_{p4s} and the gradient-flow scales would normally not need to be corrected to leading order in $1/m_c$. However, for the values

≈ a(fm)	m_l'/m_s'	aM_π	aM_K	$\sqrt{t_0}/a$	w_0/a
0.15	1/5	0.23619(22)[−34]	0.40204(25)[−57]	1.1035(05)[031]	1.1253(08)[032]
0.15	1/10	0.16598(10)[−16]	0.38030(16)[−37]	1.1113(03)[022]	1.1403(05)[022]
0.15	1/27	0.10169(09)[−11]	0.37051(16)[−41]	1.1161(02)[025]	1.1494(04)[026]
0.12	1/5	0.18904(15)[−13]	0.32335(20)[−22]	1.3142(06)[018]	1.3854(10)[019]
0.12	1/10 ($32^3\times 64$)	0.13420(09)[−04]	0.30804(15)[−09]	1.3236(04)[007]	1.4055(09)[008]
0.12	1/10 ($40^3\times 64$)	0.13396(06)[−04]	0.30812(09)[−09]	1.3233(03)[007]	1.4049(06)[008]
0.12	1/27	0.08151(04)[−02]	0.29843(11)[−08]	1.3292(05)[008]	1.4176(10)[008]
0.09	1/5	0.14039(17)[−16]	0.24033(18)[−27]	1.7267(08)[039]	1.9000(15)[043]
0.09	1/10	0.09849(08)[−03]	0.22681(12)[−07]	1.7386(05)[010]	1.9311(12)[011]
0.09	1/27	0.05719(04)[−03]	0.21936(09)[−10]	1.7451(05)[016]	1.9488(13)[018]
0.06	1/5	0.09400(16)[−38]	0.16125(16)[−65]	2.5518(13)[205]	2.9190(33)[234]
0.06	1/10	0.06686(06)[−27]	0.15390(09)[−62]	2.5716(14)[206]	2.9716(31)[238]
0.06	1/27	0.03885(03)[−02]	0.14262(06)[−07]	2.5860(07)[027]	3.0150(19)[031]

Table 2.6: Results for adjusted meson masses and gradient-flow scales, on the physical strange-quark ensembles listed in Tables 2.1 and 2.4. The adjustment corrects for charm-mass mistunings, as explained in the text. The first two columns are the approximate lattice spacing and ratio of light-to-strange sea-quark mass, with the lattice dimensions appended as needed to uniquely identify each ensemble. The remaining four columns list the masses aM_π, aM_K and gradient-flow scales $\sqrt{t_0}$, w_0 with associated statistical error in parentheses and change from the data before charm-quark mass adjustment in square brackets. The adjustments for $\sqrt{t_{0,\mathrm{imp}}}$ and $w_{0,\mathrm{imp}}$ are similar.

of aF_{p4s} taken from Ref. [13], an interpolation of aF_{p4s} to physical masses (including charm) at a fixed lattice spacing has already been performed. To correctly adjust the ratios, we define the lattice spacing to have no charm-quark mass dependence and directly adjust the remaining quantities. The values of aM_π, aM_K, $\sqrt{t_0}/a$, and w_0/a after correction for charm-quark-mass mistuning are listed in Tables 2.6 and 2.7. The adjustments for $\sqrt{t_{0,\mathrm{imp}}}/a$ and $w_{0,\mathrm{imp}}/a$ are similar to those for $\sqrt{t_0}/a$ and w_0/a, and are therefore not shown.

Simple Continuum Extrapolation

A simple continuum extrapolation can be quickly performed by including only the physical quark-mass ensembles. With just these ensembles, light-quark and strange-quark mass mistuning effects cannot be accounted for, and the statistical error will be larger than from a

m_l'/m_s	m_s'/m_s	aM_π	aM_K	$\sqrt{t_0}/a$	w_0/a
0.10	0.10	0.13177(10)[−04]	0.13177(10)[−04]	1.3603(06)[08]	1.4842(13)[08]
0.10	0.25	0.13247(09)[−04]	0.17380(11)[−05]	1.3535(04)[08]	1.4685(10)[08]
0.10	0.45	0.13271(10)[−04]	0.21713(12)[−06]	1.3446(05)[08]	1.4478(10)[08]
0.10	0.60	0.13320(10)[−04]	0.24502(13)[−07]	1.3392(08)[08]	1.4359(16)[08]
0.175	0.45	0.17487(10)[−05]	0.23192(12)[−07]	1.3393(05)[08]	1.4357(13)[08]
0.20	0.60	0.18837(17)[−13]	0.26364(18)[−18]	1.3316(06)[18]	1.4189(12)[20]
0.25	0.25	0.20883(19)[−21]	0.20883(19)[−21]	1.3400(07)[26]	1.4364(14)[28]

Table 2.7: Results for adjusted meson masses and gradient-flow scales, on the lighter-than-physical strange-quark ensembles listed in Tables 2.2 and 2.5. The adjustment corrects for charm-mass mistunings, as explained in the text. The first two columns are identical to those inTable 2.2 and are used to identify the ensembles. The latter four columns are equivalent to those in Table 2.6. The adjustments for $\sqrt{t_{0,\text{imp}}}$ and $w_{0,\text{imp}}$ are similar to the original scales and are not listed here.

fit to the complete dataset. Nevertheless, this extrapolation is useful because it provides a check on the final value from the more complicated fits and highlights the degree of improvement in discretization errors of w_0 over $\sqrt{t_0}$, as well as the scales $\sqrt{t_{0,\text{imp}}}$ and $w_{0,\text{imp}}$ over the originals $\sqrt{t_0}$ and w_0.

To perform the continuum extrapolation we multiply by the values of aF_{p4s} listed in Table 2.3 to create a dimensionless quantity that is finite in the continuum limit. We choose aF_{p4s} to keep the statistical errors smaller than what they would be from an experimentally accessible quantity such as f_π. To convert the final result to physical units, however, we must use $F_{p4s} = 153.90(09)(^{+21}_{-28})$ MeV, which was computed with the scale set by af_π. The advantage of using aF_{p4s} to set the intermediate scale is that it yields smaller relative scale errors from different ensembles, and thus aids in the extrapolation to the continuum.

Plots of $\sqrt{t_0}F_{p4s}$ and w_0F_{p4s} as a function of a^2 are shown in Fig. 2.4. The discretization improvement of w_0 over $\sqrt{t_0}$ is immediately evident in the differences between the coarsest and finest ensembles. In addition, the plot shows that the a^2 dependence is not trivial for w_0. This is not unexpected because we are using a highly improved configuration action (which

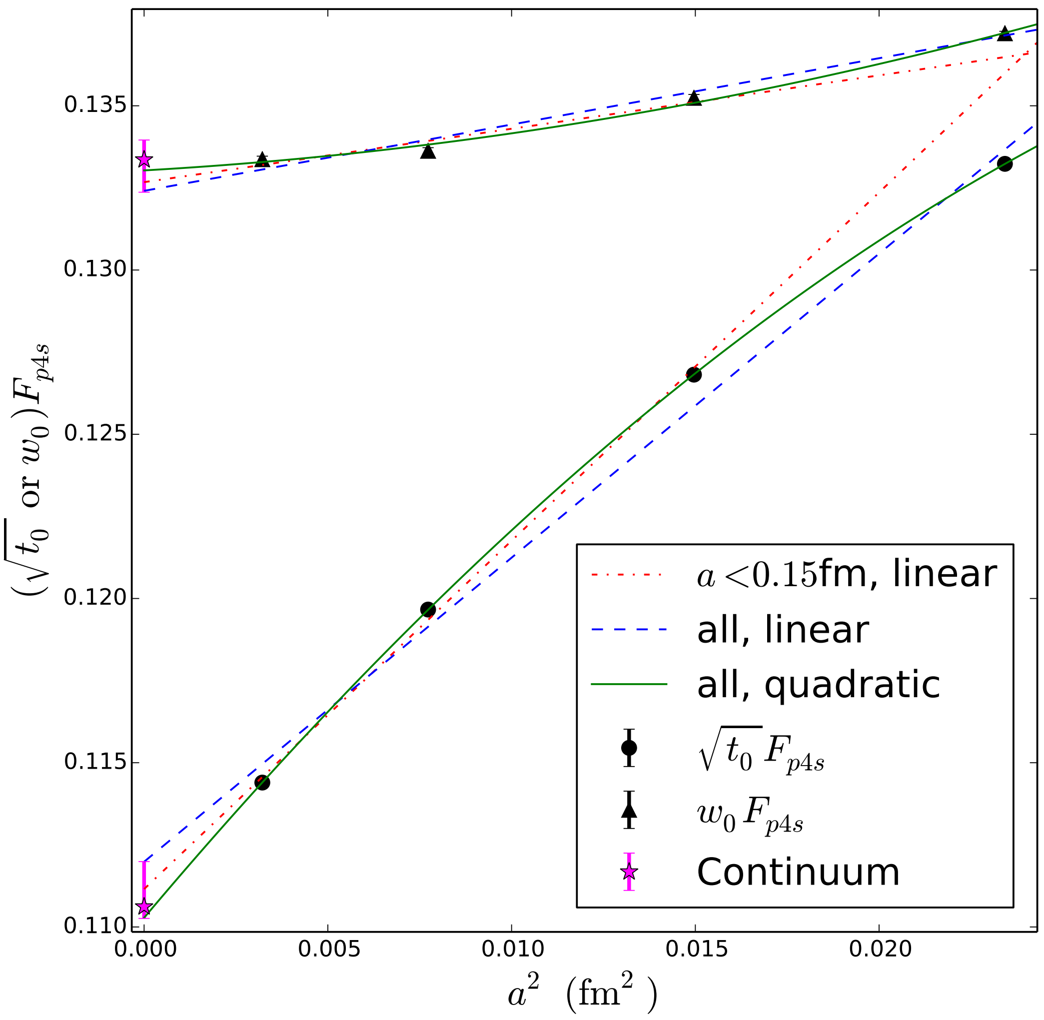

Figure 2.4: Simple continuum extrapolations for $\sqrt{t_0}F_{p4s}$ and w_0F_{p4s} over physical quark-mass ensembles only. Statistical error bars are present, but they are nearly invisible on this scale. Three fits to each dataset are shown. The red, dot-dashed line is a linear fit in a^2 to the three finer ensembles ($a < 0.15$ fm), the blue dashed line is a linear fit in a^2 to all four ensembles, and the green solid line is a quadratic fit in a^2 to all four ensembles. The continuum extrapolation points, calculated from $\sqrt{t_{0,\text{imp}}}$ and $w_{0,\text{imp}}$, are shown in magenta with error bars representing the sum of statistical and systematic uncertainties in quadrature.

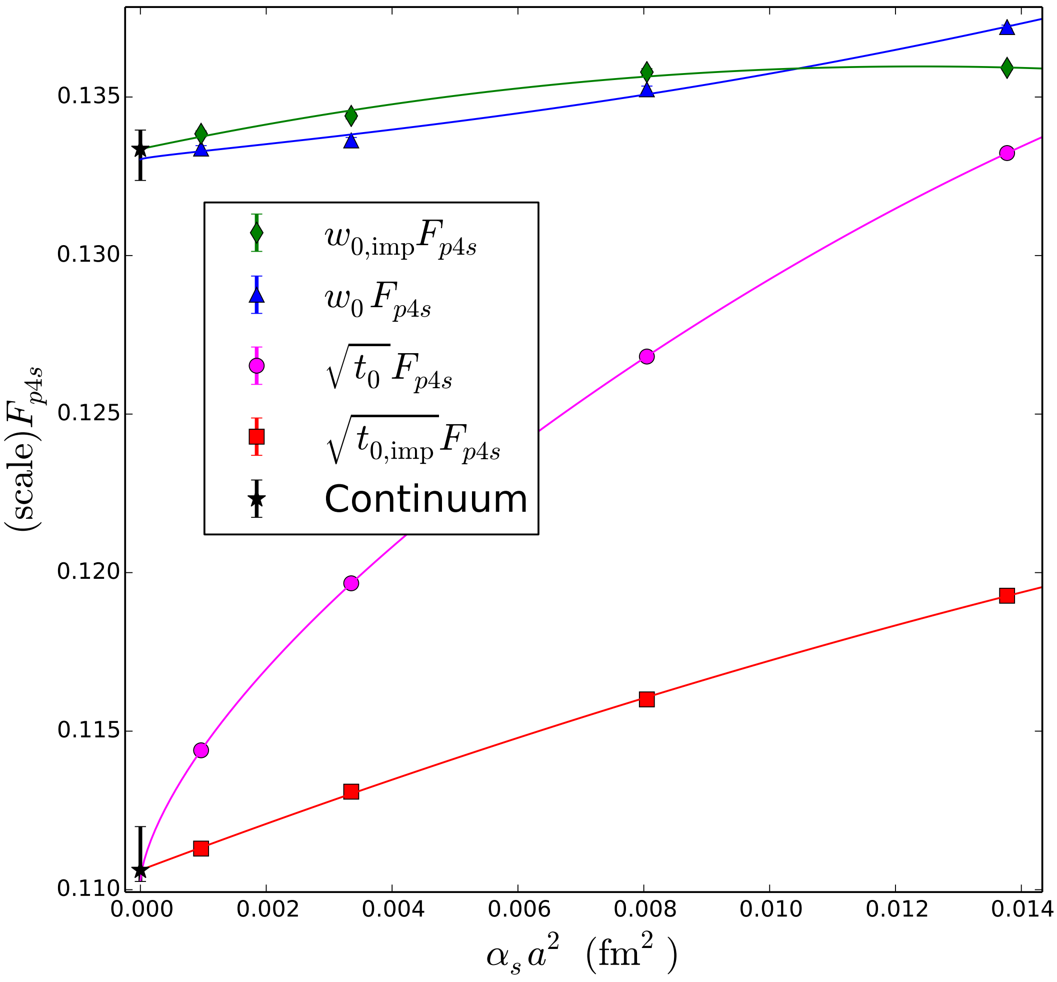

Figure 2.5: Simple continuum extrapolations for the original ($\sqrt{t_0}$ and w_0) and improved ($\sqrt{t_{0,\mathrm{imp}}}$ and $w_{0,\mathrm{imp}}$) gradient-flow scales times F_{p4s} over only physical quark mass ensembles. Quadratic fits in a^2 or $\alpha_s a^2$ over all four physical-mass ensembles are shown for the original and improved scales, respectively. The continuum extrapolation points, calculated from the improved scales, are shown in black with error bars representing the sum of statistical and systematic uncertainties in quadrature.

directly affects aF_{p4s}) for a statistically precise measurement. The importance of higher order terms in a^2 and $\alpha_s a^2$ can be seen directly in the differences between the improved and original w_0, as well as the difference between $\sqrt{t_0}$ and w_0. The situation is further complicated by effects of quark-mass mistunings between ensembles with approximately same ratio m_l/m_s. This is explored in more detail in the full fit analysis in Sec. 2.3.3. For now, we include linear fits in a^2 with or without the coarsest $a \approx 0.15$fm ensemble and quadratic fits in a^2 to all four ensembles.

Figure 2.5 compares the improved scales $\sqrt{t_{0,\text{imp}}}$ and $w_{0,\text{imp}}$ with the unimproved ones $\sqrt{t_0}$ and w_0. As before, we consider linear fits with or without the coarsest ($a \approx 0.15$ fm) ensemble, and quadratic fits with all four ensembles. For the unimproved scales the fit curves are functions of a^2; for the improved scales they are functions of $\alpha_s a^2$, since tree-level discretization errors have been removed. The improvement at tree level is obvious for $\sqrt{t_0}$, where the $\alpha_s a^2$ dependence of $\sqrt{t_{0,\text{imp}}}$ is close to linear, and the slope is considerably less steep than for $\sqrt{t_0}$. The difference between w_0 and $w_{0,\text{imp}}$ is much smaller, and is contaminated here by mistuning effects, so we postpone discussion until after we correct for such mistunings.

The continuum value is extracted from the quadratic fit in $\alpha_s a^2$ to the full dataset on the improved scales. The systematic error from the extrapolation is estimated by the largest differences between this fit and the other fits considered. This yields the simple estimates for the gradient-flow scales $\sqrt{t_0} = 0.1418(1)(^{+18}_{-\ 4})$ fm and $w_0 = 0.1710(4)(^{+\ 7}_{-12})$ fm. Here we do not included any errors (statistical or systematic) from the determination of F_{p4s} so that we can make a cleaner comparison with the extrapolations of the full dataset.

2.3.3 Full Continuum Extrapolation

Using all of the ensembles listed in Tables 2.1 and 2.2, we now perform a combined continuum extrapolation and interpolation to physical quark masses. Compared with the simple continuum extrapolation over the physical quark-mass ensembles only, the full approach has greater statistics, provides a handle for precise tuning of the light-quark and strange-quark masses to their physical values, and allows for better control and analysis of the systematic errors from discretization effects.

We break the analysis into two main sections. First, the functional forms and parameter variations for controlling mass and lattice-spacing dependence are outlined. Second, we present the results from our fits of the lattice data to the models from the first section.

Models of Mass and Lattice-Spacing Dependence

To perform the combined continuum extrapolation/quark-mass interpolation there are three functional forms that must be chosen: quark-mass terms, lattice-spacing terms, and terms that combine both (cross terms).

For the mass dependence we use the chiral expansion outlined in Sec. 2.2.2 with M_π and M_K as independent variables, standing in for the quark-mass dependence. For each fit we include the expansion up to LO (just a constant), NLO (which adds an analytic term linear in the squared meson masses, but no chiral logarithms), or NNLO (chiral logarithms and terms up to quadratic in the squared meson masses). In the fits to Eq. (2.6) the rho meson mass is used for μ and F_{p4s} is used for f. Since F_{p4s} is larger than typical stand-ins for f, we also repeated all the fits while scaling the values of F_{p4s} by the physical ratio of f_π/F_{p4s}. This did not appreciably alter the results.

For the lattice-spacing dependence we use a Taylor-series ansatz in powers of a^2, $\alpha_s a^2$, and $\alpha_s^2 a^2$. For the unimproved scales $\sqrt{t_0}$ and w_0, the first order term in lattice spacing, a^2, is always included. Higher orders are optionally included up to a^6, $\alpha_s a^2$, and $\alpha_s^2 a^2$. For the improved scales $\sqrt{t_{0,\text{imp}}}$ and $w_{0,\text{imp}}$, the first order in either $\alpha_s a^2$ or a^4 is always included. Higher orders are optionally included up to a^8, $(\alpha_s a^2)^2$, and $\alpha_s^2 a^2$. Even though the scales $\sqrt{t_{0,\text{imp}}}$ and $w_{0,\text{imp}}$ are improved to order a^8 at tree level, the a^4 through a^8 terms are included for fits because aF_{p4s} has leading corrections of $\alpha_s a^2$ and a^4. For both scales, the number of lattice-spacing terms in a single fit is not allowed to exceed three. Together with the value of the scale in the continuum limit, this ensures that at most four parameters describe the a-dependence of the data from our four unique lattice spacings.

For cross terms, we include all products of chiral and lattice-spacing terms whose total order is no higher than the largest non-cross term included in the fit function. Also no cross terms are constructed from the highest orders of mass or lattice-spacing terms. For example, a fit including a^6 and the chiral expansion to NNLO would include a term like $a^4(M/(4\pi f))^2$ but not include $a^2(M/(4\pi f))^4$. For the purpose of counting orders, we assume the power counting $\alpha_s \sim (\Lambda_{\text{QCD}}a)^2 \sim (M/(4\pi f))^2$.

Once the functional form is chosen, we also consider various restrictions of the dataset. As already suggested from the naive fit to the physical quark-mass ensembles only, the $a \approx 0.15$ fm ensembles may require higher orders of a^2 to be included. So we consider fits that include or drop these ensembles. Furthermore, when the $a \approx 0.15$ fm ensembles are dropped, we do not include more than two lattice spacing terms to ensure the three unique lattice spacings represented by the dataset are not parameterized by four or more variables. A second restriction on the dataset is determined by the kaon mass. The lighter-than-physical strange-quark ensembles have strange-quark masses all the way down to 1/10 the physical strange-quark mass. Including these ensembles along with the physical-mass ensembles that comprise most of out data requires more complex chiral forms to cover the large range in m'_s. We therefore consider seven different lower bounds for the kaon masses included in the fit, ranging from just below the physical strange-quark mass, to near zero, which includes all the ensembles.

For all scales, there are three chiral expansions and seven choices of lower bound for the kaon mass. For the original scales $\sqrt{t_0}$ and w_0, there are six lattice-spacing expansions with the $a \approx 0.15$ fm ensembles included and three lattice-spacing expansions with the $a \approx 0.15$ fm ensembles *not* included. This produces a total of $3 \times (6 + 3) \times 7 = 189$ different fits. For the improved scales $\sqrt{t_{0,\text{imp}}}$ and $w_{0,\text{imp}}$, there are nine lattice-spacing expansions with the $a \approx 0.15$ fm ensembles included, five lattice-spacing expansions with the $a \approx 0.15$ fm ensembles *not* included. This produces a total of $3 \times (9 + 5) \times 7 = 294$ different fits.

Fits to the Lattice Data

We gauge the acceptability of each of the fits outlined in Sec. 2.3.3 using the p-value and the proximity of the fit curve to the data from our most important ensemble, the one with physical quark-masses and $a \approx 0.06$ fm. Figure 2.6 shows the acceptability for the original and improved scales with p-value as the x-axis, deviation from the physical $a \approx 0.06$ fm ensemble as the y-axis, and the size (radius) of each data point proportional to the number of degrees of freedom. We define 'acceptable' fits as those with $p > 0.01$. Acceptable fits are those inside the black box in Fig. 2.6. Note that, for all the scales considered, fits with acceptable p-values are usually close to the result from the $a \approx 0.06$ fm physical-mass ensemble. For $w_0 F_{p4s}$ and $\sqrt{t_0} F_{p4s}$ no acceptable fit deviates by more than 2 or 2.5 σ_{stat}, respectively.

To determine a central value and systematic error from the choice of fit we construct histograms in Fig. 2.7 of the continuum results from fits with $p > 0.01$. The number of acceptable fits is further refined after careful examination of the continuum extrapolations for outlying members. For $\sqrt{t_0} F_{p4s}$, all fits simultaneously including a^2, $\alpha_s a^2$, and $\alpha_s^2 a^2$ were dropped due to extreme curvature outside the interpolation range (0.06 fm $< a <$ 0.15 fm). For the same reason, all fits to $w_{0,\text{imp}} F_{p4s}$ simultaneously including a^4, $\alpha_s a^2$, and $(\alpha_s a^2)^2$ were dropped.

For each of $\sqrt{t_0}$ and w_0, the central fit is chosen by locating the fits closest to the median with $p > 0.1$ and at least twice as many data points as parameters. If there are several fits that satisfy this criteria, the fit with the larger dataset is chosen. The central fits to $\sqrt{t_0}$ and w_0 are both to the improved scales, include the chiral expansion to NNLO with the full range of kaon masses, and include the lattice spacing terms a^4, a^6, and $\alpha_s a^2$ with the full range of lattice spacings. The fits have 9 free parameters and 20 data points. For $\sqrt{t_0}$ the central fit has $\chi^2/dof = 12.9/11$, $p = 0.30$, and is 0.7σ higher than the physical $a \approx 0.06$ fm ensemble. For w_0 the central fit has $\chi^2/dof = 12.0/11$, $p = 0.37$, and is 0.5σ higher than the physical $a \approx 0.06$ fm ensemble. The central fits are shown in Fig. 2.8. The dashed lines indicate how well the fit describes the data by showing the fit function evaluated at the same masses and lattice spacing as the data points. The three solid bands show the lattice-spacing dependence at fixed quark masses, tuned to a physical value for the strange-quark mass and the indicated ratio of the light-quark to strange-quark mass. One clearly sees the effects of retuning the quark masses from their run values.

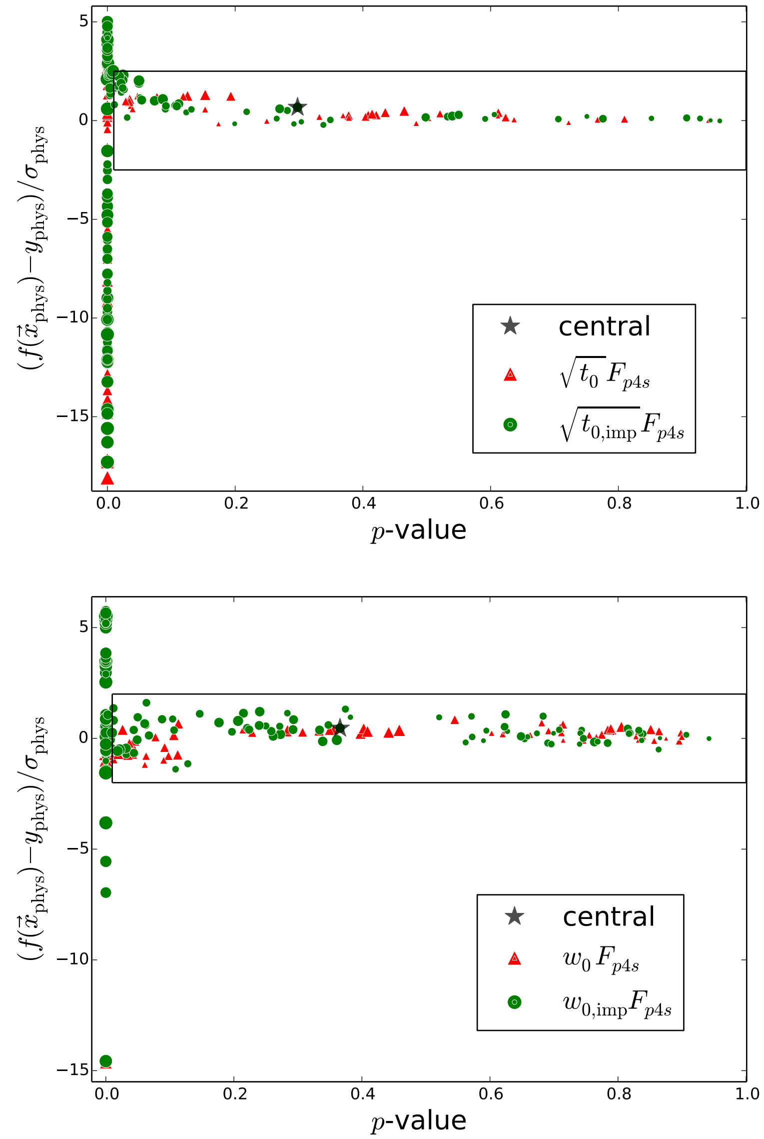

Figure 2.6: The 'acceptability' for the various fits considered for the t_0 scales ($\sqrt{t_0}$ and $\sqrt{t_{0,\rm imp}}$, top) and w_0 scales (w_0 and $w_{0,\rm imp}$, bottom). Fit acceptability is parameterized by the p-value (x-axis) and the proximity to the results from the physical-mass $a \approx 0.06$ fm ensemble in units of σ_{stat} (y-axis). The size of the points is proportional to the number of degrees of freedom. The space within the black box contains fits with $0.01 < p < 1.0$ and a deviation of less than $2.5\sigma_{\rm stat}$ or $2\sigma_{\rm stat}$ for the t_0 and w_0 scales, respectively. This box determines the acceptable subset of fits considered in subsequent analysis. The central fit chosen from this analysis is denoted by the star.

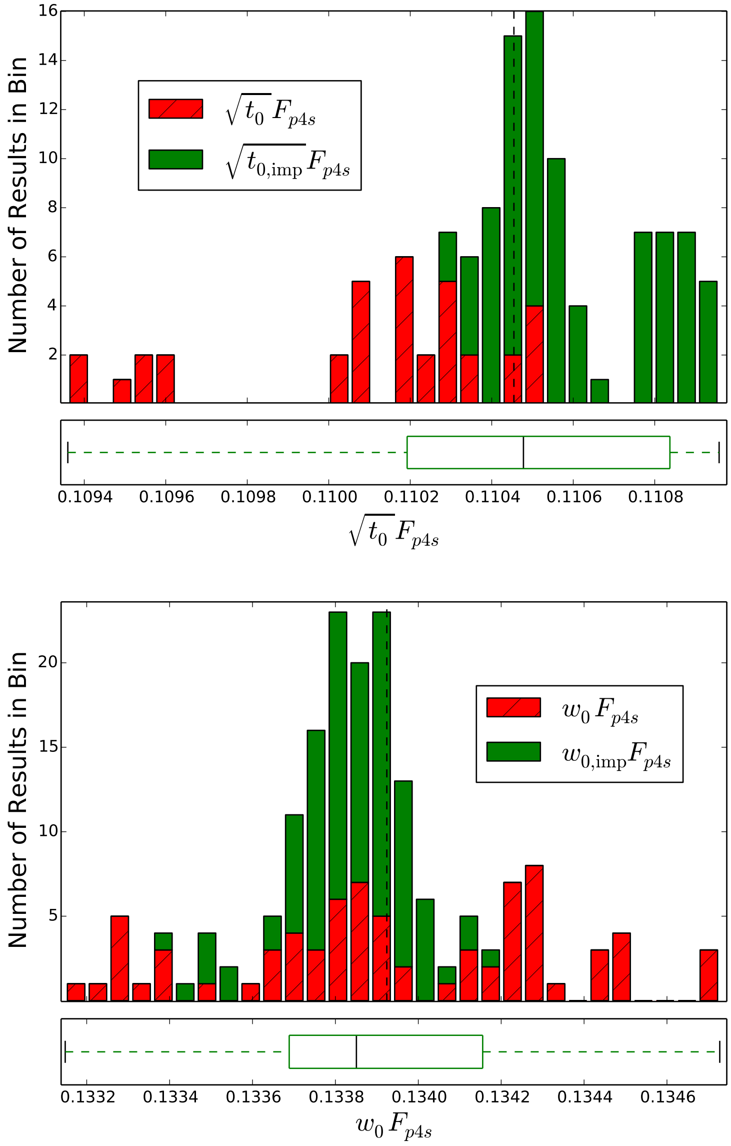

Figure 2.7: Histograms of the continuum extrapolations for $\sqrt{t_0}F_{p4s}$(top) and w_0F_{p4s}(bottom) for all 'acceptable' fits (see text). Each histogram is a stacked combination of continuum extrapolations from the original ($\sqrt{t_0}$ and w_0) and improved scales ($\sqrt{t_{0,\mathrm{imp}}}$ and $w_{0,\mathrm{imp}}$), represented by the red, hashed and green, solid bars, respectively. The box and error bars along the bottom denote the minimum, mean, maximum, and central 68% of the distribution. The vertical dashed line for each distribution marks the continuum result of the associated central fit.

For the fit to $\sqrt{t_{0,\text{imp}}}F_{p4s}$, the lattice-spacing dependence at finer lattice spacings ($a \le 0.09$ fm) is dominated by the $\alpha_s a^2$ contribution. The a^4 contributions start to become comparable to those from $\alpha_s a^2$ for $a \gtrsim 0.12$ and produce the curvature evident in Fig. 2.8 (top). The lattice-spacing dependence of $w_{0,\text{imp}}F_{p4s}$ is milder than for $\sqrt{t_{0,\text{imp}}}F_{p4s}$, but also more complicated. For small $a \lesssim 0.06$ fm, the contributions to the fit from both $\alpha_s a^2$ and a^4 are fairly small and have opposite signs, so that $w_{0,\text{imp}}F_{p4s}$ is nearly constant as $a \to 0$. For intermediate lattice spacings a^4 contributions dominate, while a^6 starts to become important for $a > 0.12$ fm, giving the downward curvature for large a seen in in Fig. 2.8 (bottom).

It is revealing to examine the central extrapolations plotted through only the physical mass ensembles for all four gradient-flow scales, as was done in Fig. 2.5 for the naive extrapolation. This plot is presented in Fig. 2.9. Compared to the simpler fits to just the physical-mass ensembles in Sec. 2.3.2, quark-mass mistunings in the physical quark-mass ensembles are accounted for here. This leads the two coarsest physical-mass ensembles ($a = 0.12$ and $a = 0.15$ fm) to shift down when retuned to the precise ratio $m_l/m_s = 1/27$. For the fits to $\sqrt{t_0}F_{p4s}$ and $\sqrt{t_{0,\text{imp}}}F_{p4s}$ the difference is visible but has only a small effect on the continuum extrapolation. For the fits to $w_0 F_{p4s}$ and $w_{0,\text{imp}}F_{p4s}$ the shift is very important as the fluctuation in the data points across the range of a^2 is comparable to the size of the effect of mass retuning.

For both $\sqrt{t_0}$ and w_0, the tree-level improved version of each scale eliminates a^2 errors and reduces a^4 and a^6 contributions. The improvement in $\sqrt{t_0}$ is obvious in Fig. 2.9. For w_0 the quark-mass retuning allows the reduced a^2 dependence of $w_{0,\text{imp}}F_{p4s}$ relative to $w_0 F_{p4s}$ to become visible at small lattice spacing. Specifically for $a \lesssim 0.06$ fm (the extreme left of the plot), the dominant contribution to $w_0 F_{p4s}$ is the a^2 term, while $w_{0,\text{imp}}F_{p4s}$ is a mixture of the a^4 and $\alpha_s a^2$ contributions of opposite signs, with each contribution approximately one third

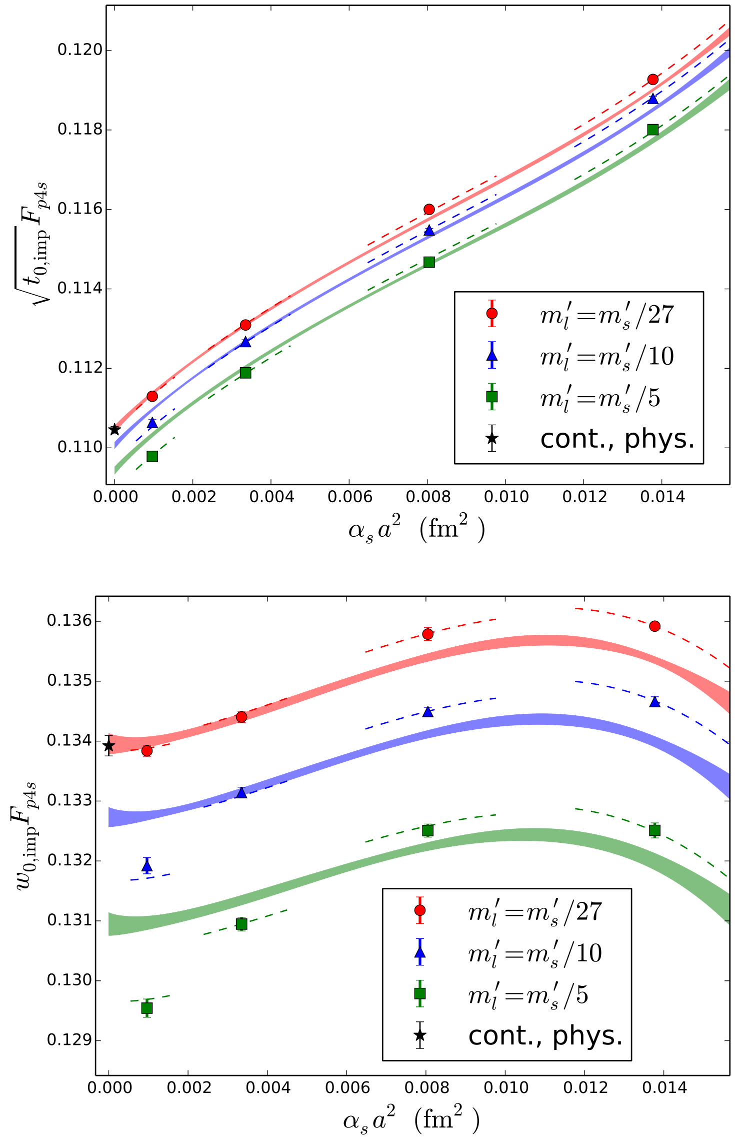

Figure 2.8: The central fits to the gradient-flow scales $\sqrt{t_{0,\mathrm{imp}}}F_{p4s}$ and $w_{0,\mathrm{imp}}F_{p4s}$, plotted as a function of $\alpha_s a^2$. These are used to compute $\sqrt{t_0}$ (top) and w_0 (bottom) at physical quark masses and in the continuum, as indicated by the black stars. Only $m_s' \approx m_s$ ensembles are plotted, but the fits include all $m_s' \leq m_s$ ensembles. Dashed lines represent the fit through each ensembles' actual quark masses and lattice spacing, while the solid bands are for varying lattice spacing at fixed quark masses retuned to the physical strange-quark mass and the ratio of m_l'/m_s' specified in the legend.

the magnitude of the a^2 term for $w_0 F_{p4s}$. Thus the slope of $w_{0,\mathrm{imp}} F_{p4s}$ as $a \to 0$ is significantly smaller than that of $w_0 F_{p4s}$. However, for intermediate and larger lattice spacings, there is no reduction in the lattice-spacing dependence of $w_{0,\mathrm{imp}} F_{p4s}$ relative to $w_0 F_{p4s}$.

2.4 Continuum Results

2.4.1 Scales in Physical Units

We compute our final estimate of the gradient-flow scales in physical units by evaluating the continuum-extrapolated, physical-quark-mass-interpolated value of $\sqrt{t_0} F_{p4s}$ and $w_0 F_{p4s}$ for the best fit in Sec. 2.3.3 and dividing by the physical value of F_{p4s} (see Sec. 2.3).

$$\sqrt{t_0} = 0.1416(1)_{stat}(^{+6}_{-2})_{t_0,extrap}(^{+3}_{-2})_{F_{p4s},extrap}(2)_{\mathrm{FV}}(3)_{f_\pi\ \mathrm{PDG}}\ \mathrm{fm} \tag{2.15}$$

$$w_0 = 0.1717(2)_{stat}(^{+10}_{-10})_{w_0,extrap}(^{+3}_{-2})_{F_{p4s},extrap}(2)_{\mathrm{FV}}(3)_{f_\pi\ \mathrm{PDG}}\ \mathrm{fm} \tag{2.16}$$

The first error is statistical and is from the corresponding central fit discussed in Sec. 2.3.3. The remaining, systematic, errors are from: continuum extrapolation/chiral interpolation (estimated by variations among fits), corresponding continuum and chiral errors on F_{p4s} in physical units, residual finite-volume effects on F_{p4s}, and the error in F_{p4s} from the experimental error in f_π [66], respectively. The error from the choice of fit for the gradient-flow scale is estimated using the histograms in Fig. 2.7. For $\sqrt{t_0}$, where tree-level improvement dramatically reduces our observed discretization errors, we consider only the range of results from the continuum extrapolation of $\sqrt{t_{0,\mathrm{imp}}}$; we take the maximum differences of those fits from the central fit as the error. In other words, we use the full range of the fits shown in green in Fig. 2.7, but do not consider the red outliers at the left of the histogram. For

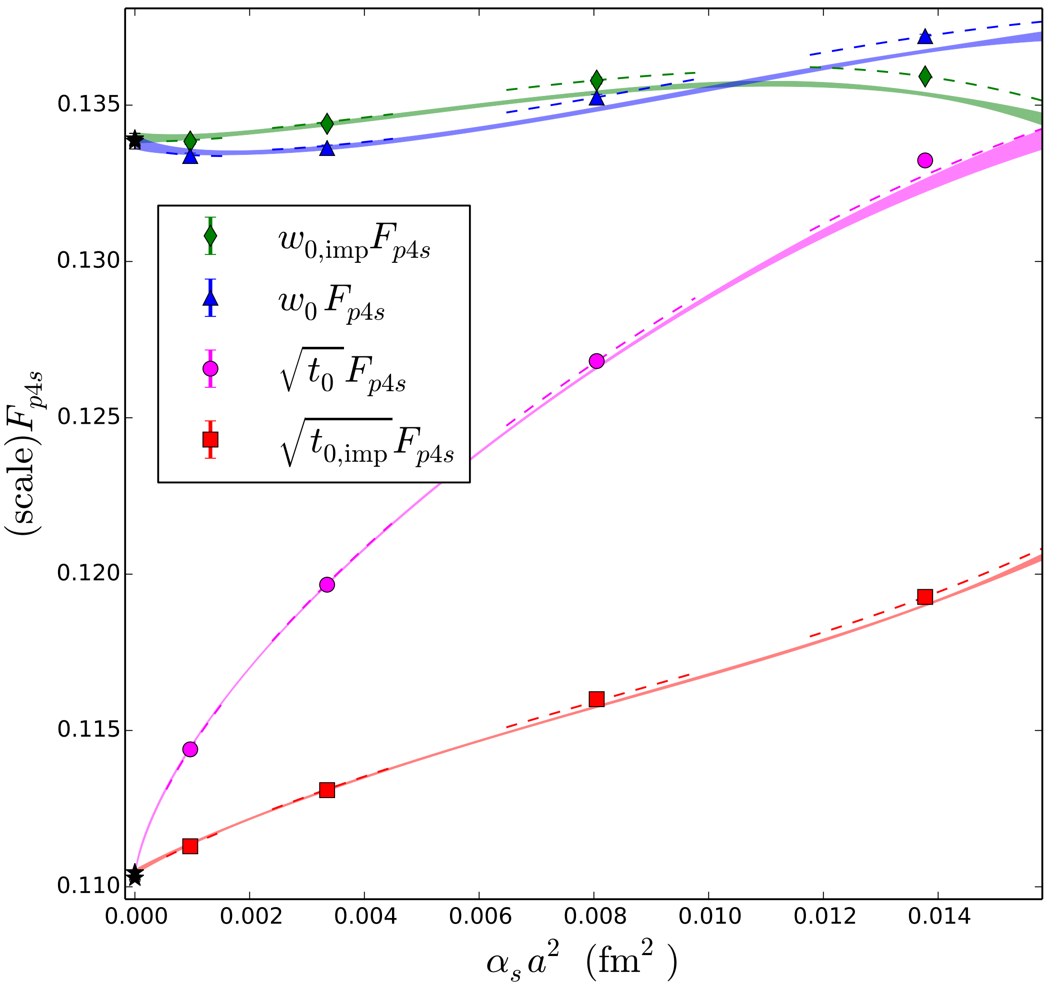

Figure 2.9: Continuum extrapolations for the original ($\sqrt{t_0}$ and w_0) and improved ($\sqrt{t_{0,\text{imp}}}$ and $w_{0,\text{imp}}$) gradient-flow scales times F_{p4s} plotted for physical quark-mass ensembles only. All fits include the chiral expansion to NNLO and are fit to the full range of kaon masses. For $\sqrt{t_0}F_{p4s}$ the fit is quadratic in a^2 and drops the coarsest ensembles at $a \approx 0.15$fm. For $w_0 F_{p4s}$ the fit is cubic in a^2 and covers the full range of a. For the improved scales the plotted lines are from the central fits discussed in this section. The continuum-extrapolation points are shown in black with error bars representing only the statistical error.

w_0, only a mild improvement in the observed a^2 dependence occurs after removing tree-level discretization errors. We consider all fits, both to w_0 and to $w_{0,\text{imp}}$ and take, conservatively, the maximum differences of all those fits (both green and red in Fig. 2.7) from the central fit as the error. The remaining extrapolation errors, residual finite-volume effects, and error from the experimental value of f_π come directly from the analysis of F_{p4s} [13].

The results in Eqs. (2.15) and (2.16) may be compared to the earlier, simple estimates of $\sqrt{t_0} = 0.1418(1)(^{+18}_{-\ 4})$ fm and $w_0 = 0.1710(4)(^{+\ 7}_{-12})$ fm from the physical quark-mass ensembles in Sec. 2.3.2. For both $\sqrt{t_0}$ and w_0, the extrapolated values agree, within the earlier systematic errors. (Note that the earlier result did not include the uncertainties from F_{p4s} and f_π, which give the last three errors in Eqs. (2.15) and (2.16).) For $\sqrt{t_0}$, the agreement is exact within quoted precision and both extrapolations lead to similar statistical uncertainties. The main improvement of extrapolating $\sqrt{t_{0,\text{imp}}}F_{p4s}$ over the full set of ensembles is the narrower systematic uncertainty in the continuum, physical mass extrapolation. For w_0, the central value from the simpler fit is slightly lower. This shift is attributable to the quark-mass re-tuning and higher-order discretization terms only accessible to the full extrapolation. Also, the full extrapolation to $w_{0,\text{imp}}F_{p4s}$ leads to a statistical error two times smaller. Overall, the addition of non-physical quark mass ensembles reduces uncertainties and improves control over the continuum extrapolation without significantly deviating from our initial estimate.

The results presented in this work have evolved from preliminary results presented previously. In chronological order, the estimates from two earlier proceedings are $w_0 = 0.1711(2)(8)$ fm in Ref. [16], and $\sqrt{t_0} = 0.1422(2)(5)$ fm and $w_0 = 0.1732(4)(8)$ fm in Ref. [15]. For comparison to the results in this work, we have altered the original results by keeping only the statistical and systematic error from the choice of fit form to $\sqrt{t_0}F_{p4s}$ or w_0F_{p4s}. We have dropped all other systematic errors, which are shared across all results. For both scales, all results agree within 2σ of the current results. Compared to the result in Ref. [16], those in Ref. [15] account for charm-quark mass mistunings, use aF_{p4s}, instead of af_π, to set the scale, and consider a larger selection of discretization terms. However Ref. [15] uses an incorrect value of am_c for the physical quark-mass, $a \approx 0.06\,\mathrm{fm}$ ensemble when adjusting for charm-quark mass mistunings. The mistake is fixed in the current work and is responsible for most of the downward shift relative to the scales presented in Ref. [15]. Compared to Ref. [15], the current work also incorporates the tree-level improved versions of each scale and refines the selection of discretization terms.

2.4.2 Continuum Meson-Mass Dependence

Using the best fits and datasets chosen in Sec. 2.3.3, we determine the continuum meson-mass dependence of w_0 under a mass-independent scale-setting scheme. The resulting function is useful for prediction of the scales on future ensembles, as well as for explicit comparison of the mass dependence of w_0 to that of other scale-setting quantities. To predict a scale one measures w_0/a (or $w_{0,\mathrm{imp}}/a$), aM_π, and aM_K on a subset of the ensemble to be generated. Then, by evaluating the function at the corresponding dimensionless variables $P = (w_0 M_\pi)^2$ and $K = (w_0 M_K)^2$ one can determine the continuum value of w_0 in physical units at those masses, $w_0(P, K)$, and compute the resulting scale $a = w_0(P, K)/(w_0/a)$. This procedure was originally suggested in Ref. [31].

The functional form of the meson-mass dependence $w_0(P, K)$ is chosen to be the same as the chiral expansion to NNLO, in agreement with the best fit chosen in Sec. 2.3.3. The coefficients are determined by solving the implicit equation

$$w_0 = f(P = (w_0 M_\pi)^2, K = (w_0 M_K)^2) \tag{2.17}$$

numerically for $w_0 = w_0(P, K)$. Using the best fit $h(a, (M_\pi/F_{p4s})^2, (M_K/F_{p4s})^2) = w_0 F_{p4s}$ of Sec. 2.3.3, the implicit function is defined as

$$w_0(P, K) = h(0, P/(w_0 F_{p4s})^2, K/(w_0 F_{p4s})^2)/F_{p4s}\ , \tag{2.18}$$

where F_{p4s} is evaluated at physical quark masses and in the continuum. Note, the first parameter is set to 0, denoting the continuum limit. We find

$$\begin{aligned} w_0(P, K) &= 0.1852 - 0.0474(2K + P) - 0.0448\, P\mu_P - 0.1229\, K\mu_K \qquad (2.19) \\ &- 0.0158\,(P - 4K)\mu_\eta - 0.0307\,\eta\,\mu_\eta + 0.0094\,(2K + P)^2 + 0.0447\,(K - P)^2 \end{aligned}$$

where $\mu_z = z\log(z/\Lambda)$, with $\Lambda = (M_\rho/\sqrt{8}\pi F_{p4s})^2 \approx 0.3170$, and $\eta = (4K - P)/3$. The error in $w_0(P, K)$ is approximately the same as for our continuum determination of w_0 at physical masses. Figure 2.10 plots this function over a large range of values of P and K. Values corresponding to the HISQ ensembles and to the physical-mass point are overlaid to give a sense of the range of meson masses for which this function is valid. The leading $(2K + P)$ dependence can be observed in the roughly linear shape for each line of constant K and the approximately constant vertical gap between lines of fixed K, independent of P. The separation of points within the clusters of physical strange-quark mass ensembles that were simulated close to the physical ratios $m_l/m_s = 1/5, 1/10$, and $1/27$ is due to quark-mass mistunings and discretization errors.

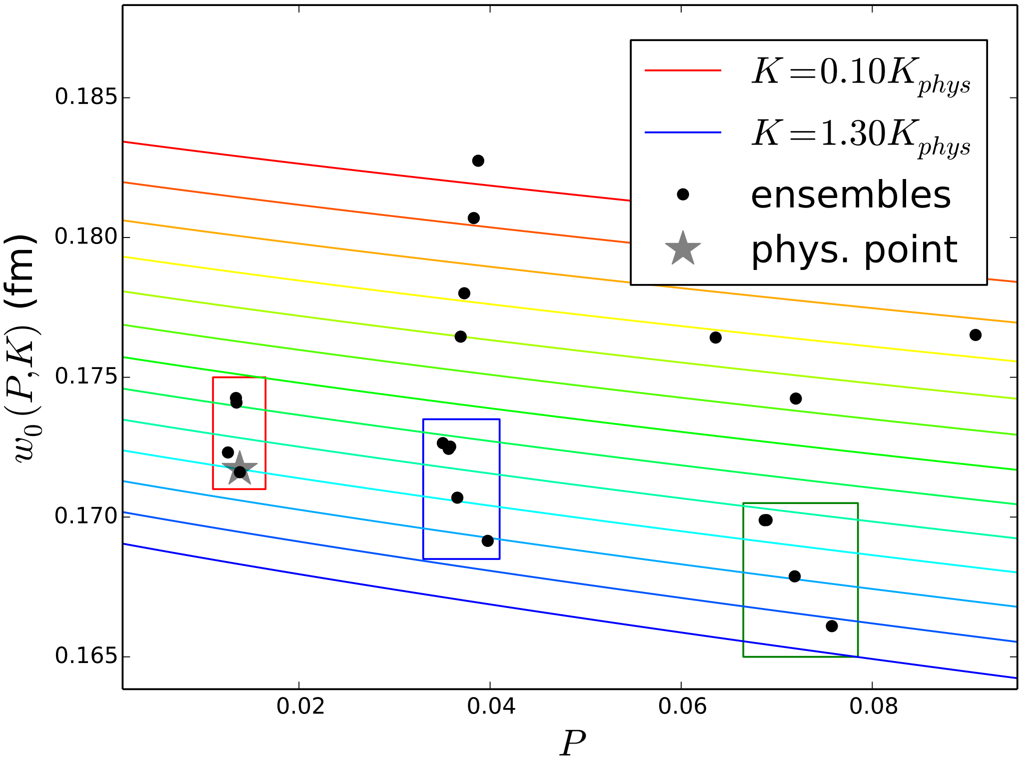

Figure 2.10: The continuum mass dependence of w_0 as a function of $P = (w_0 M_\pi)^2$ for fixed values of $K = (w_0 M_K)^2$. The black points and star illustrate the values of the pion and kaon mass that correspond to various HISQ ensembles and to the physical point, respectively. The three boxes enclose the physical strange mass ensembles with different ratios of m_l'/m_s'. From the left to rightmost box the ratios are $m_l'/m_s' = 1/27$, 1/10, and 1/5.

Using Eq. (2.19) and the results for $w_{0,\mathrm{imp}}/a$ on the HISQ ensembles, we recalculate a(fm) for each ensemble and check that the results are consistent with the original lattice spacings set through F_{p4s}. Table 2.8 lists the lattice spacings determined through F_{p4s} in Ref. [13] and w_0 in this work. The scales determined from w_0 are almost independent of quark masses for fixed β, showing that the procedure is working as designed, and can be used to find consistent scales of new ensembles, even if they do not have physical quark masses. Lattice spacings determined from F_{p4s} and w_0 on the physical quark-mass ensembles agree as the continuum limit is approached, and are close over the whole range of lattice spacings. This fitting procedure may be repeated to find $\sqrt{t_0}$ as a function of $P' = (\sqrt{t_0}M_\pi)^2$ and $K' = (\sqrt{t_0}M_K)^2$. As might be expected from the large slope seen for $\sqrt{t_0}$ in Fig. 2.5, the resulting function shows large discretization effects at the coarser scales. The discretization errors appear as large variations in the scale determinations on coarse ensembles for different quark masses with fixed β. We thus do not include the results in Table 2.8.

2.5 Discussion and Conclusions

With the continuum results complete, we compare with computations of gradient-flow scales performed by other collaborations. Table 2.9 shows a selection of those calculations and their final results in comparison with our own. The same results are also plotted in Fig. 2.11. Differences are shown divided by the joint error, except for HPQCD. Because HPQCD uses a subset of the HISQ ensembles employed here, we do not use the joint sigma, which would double count several sources of error; instead, we use the larger of the two collaborations' total error. Our results for both scales are compatible with those of the three other published continuum-limit calculations by HPQCD, HotQCD, and BMW; the largest difference is 1.9σ.

β	m'_l/m'_s	$(aF_{p4s})/F_{p4s}$(fm)	$w_0/(w_{0,\text{imp}}/a)$(fm)
5.80	1/5	–	$0.1515(^{+11}_{-\ 9})$
5.80	1/10	–	$0.1513(^{+11}_{-10})$
5.80	1/27	$0.15305(^{+57}_{-41})$	$0.1512(^{+11}_{-\ 9})$
6.00	1/5	–	$0.1210(^{+9}_{-8})$
6.00	1/10 ($32^3 \times 64$)	–	$0.1208(^{+9}_{-8})$
6.00	1/10 ($40^3 \times 64$)	–	$0.1209(^{+9}_{-8})$
6.00	1/27	$0.12232(^{+45}_{-33})$	$0.1208(^{+9}_{-7})$
6.30	1/5	–	$0.0876(^{+7}_{-5})$
6.30	1/10	–	$0.0876(^{+6}_{-6})$
6.30	1/27	$0.08791(^{+33}_{-24})$	$0.0876(^{+6}_{-5})$
6.72	1/5	–	$0.0568(^{+4}_{-4})$
6.72	1/10	–	$0.0566(^{+4}_{-4})$
6.72	1/27	$0.05672(^{+21}_{-16})$	$0.0567(^{+4}_{-3})$

Table 2.8: Values of the lattice spacing determined from aF_{p4s}[13] and $w_{0,\text{imp}}/a$ on the physical strange-quark HISQ ensembles listed in Table 2.1. The first two columns list the coupling β and ratio of light to strange sea-quark mass, with the lattice dimensions appended as needed to uniquely identify each ensemble. Since we do not have a function corresponding to Eq. (2.19) for F_{p4s}, mass-independent scale setting with F_{p4s} is performed on the physical quark-mass ensembles only.

Collaboration	N_f	$\sqrt{t_0}$ (fm)	$\Delta\sqrt{t_0}/\sigma$	w_0 (fm)	$\Delta w_0/\sigma$
MILC [This work]	2+1+1	$0.1416(1)(^{+8}_{-5})$	–	$0.1717(2)(^{+12}_{-11})$	–
HPQCD [42]	2+1+1	0.1420(8)	+0.5	0.1715(9)	−0.2
ETMC* [41]	2+1+1	–	–	0.1782	–
HotQCD [14]	2+1	–	–	0.1749(14)	+1.8
BMW [31]	2+1	0.1465(21)(13)	+1.9	0.1755(18)(04)	+1.7
QCDSF-UKQCD* [48]	2+1	0.153(7)	+1.6	0.179(6)	+1.2
ALPHA* [33]	2	0.1535(12)	+8.3	0.1757(13)	+2.3

Table 2.9: Continuum results for the gradient-flow scales $\sqrt{t_0}$ and w_0 from different collaborations [42, 41, 31, 33, 48, 14]. The last two columns tabulate the difference between the results of other collaborations and this work, relative to one joint sigma. For HPQCD, whose errors are not independent of ours, we simply use the larger error for the comparison. Results of the three collaborations marked with an asterisk are preliminary conference results.

Our best agreement is with HPQCD, the latter of which performed an independent analysis on the same HISQ configurations but without the $a = 0.06$ fm ensembles. We also agree with the published, single-lattice-spacing result for $\sqrt{t_0} = 0.1414(7)(5)$ fm from TWQCD [34]. Furthermore, we agree within 2σ with all but one collaboration's preliminary results: $\sqrt{t_0}$ and w_0 calculated by the ALPHA collaboration. This may be due to the difference in the number of flavors; however, it is unclear why the N_f dependence would be so much stronger for $\sqrt{t_0}$ than for w_0.

Finally, we compare the relative lattice scale found from $\sqrt{t_0}$, w_0, and other quantities used for scale setting. Here, we only compare the relative percent statistical error, since the sources of systematic error vary considerably between the quantities considered. In addition, while discretization errors are common to all scale-setting quantities, the errors due to a single quantity cannot be definitively identified, since one always calculates dimensionless ratios of two quantities. Table 2.10 compares the percent error for various scale-setting quantities in lattice units measured on the HISQ physical quark-mass ensembles. Both gradient-flow scales are determined more precisely than r_1/a and af_π. The precision of $\sqrt{t_0}/a$ is on a

$\approx a$(fm)	Statistical Error (%)				
	r_1/a	af_π	aF_{p4s}	$\sqrt{t_0}/a$	w_0/a
0.15	0.36	0.11	0.03	0.02	0.03
0.12	0.25	0.08	0.03	0.04	0.07
0.09	0.33	0.09	0.03	0.03	0.07
0.06	0.12	0.11	0.03	0.03	0.06

Table 2.10: Percent statistical error for several scale-setting quantities including r_1, f_π, F_{p4s}, and the gradient-flow scales $\sqrt{t_0}/a$ and w_0/a on the physical quark mass HISQ ensembles listed in Tables 2.1 and 2.4. The statistical errors in the improved scales $\sqrt{t_{0,\text{imp}}}$ and $w_{0,\text{imp}}$ are comparable to the original gradient-flow scales. The first column is the approximate lattice spacing and can be used to uniquely identify each ensemble.

par with the most precise, aF_{p4s}. This low statistical error was an original motivation for computing the scale from gradient flow. Note further that $\sqrt{t_0}/a$ and w_0/a have only been determined on a small subset of the configurations at finer lattice spacings, while the aF_{p4s} values come from the entire ensembles, so there is considerable room for improvement for the gradient-flow scales. In addition, lower systematic errors — in particular, low dependence on quark masses — may make the gradient-flow scales preferable to aF_{p4s} for relative scale setting. Statistical errors for w_0/a are larger than those of $\sqrt{t_0}/a$. This is one factor that leads to our slightly more precise continuum extrapolated value for $\sqrt{t_0}$ compared to w_0. On the other hand, Fig. 2.9 makes a fairly convincing case that the discretization effects for w_0 are much smaller than those for $\sqrt{t_0}$. It is conceivable that the small slope for w_0 is due to an accidental cancellation between its discretization errors and those of F_{p4s}. However, the argument following Eq. (2.12) above, as well as the empirical evidence given in Ref. [31], indicate that w_0 does in fact have significantly smaller a^2 dependence than $\sqrt{t_0}$. Finally, we remark that the small error of aF_{p4s}, in comparison with that of af_π, is what motivates us to use aF_{p4s} for our continuum extrapolations of the gradient- flow scales, as discussed in Sec. 2.3.2.

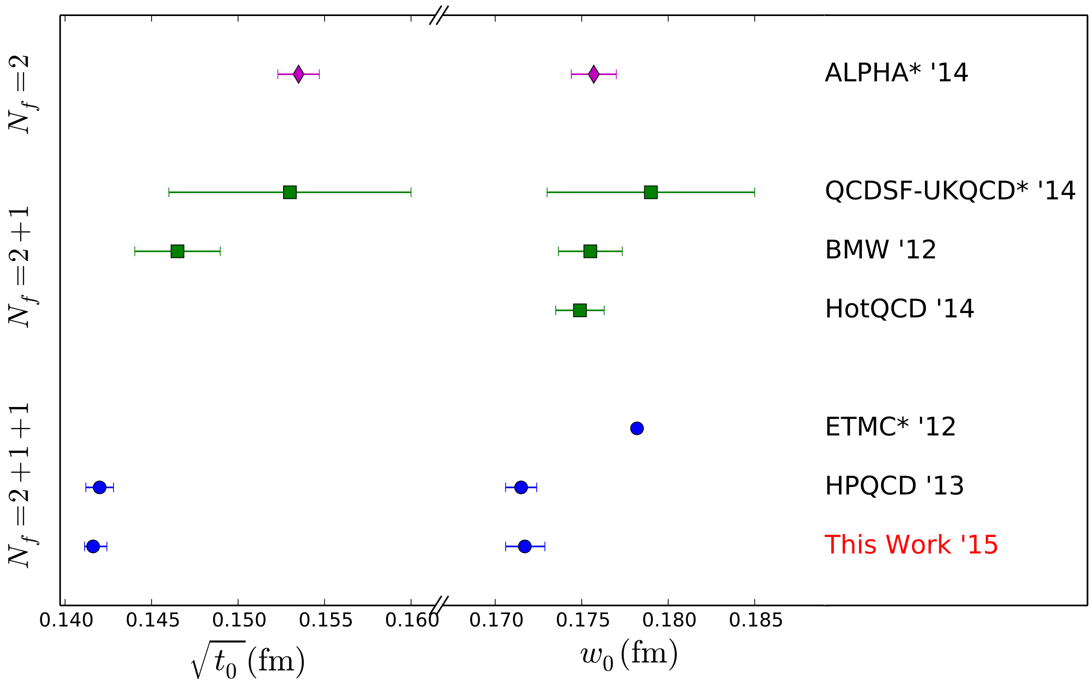

Figure 2.11: The continuum values of $\sqrt{t_0}$ and w_0 separated by collaboration and grouped by the number of flavors. References for each collaborations' work can be found in Table 2.9. Those results for collaborations marked with an asterisk are preliminary. Our results for $\sqrt{t_0}$ and w_0 are consistent within two standard deviations to all other results except the preliminary calculations from the ALPHA collaboration.

In conclusion, we have computed the continuum, physical mass values of $\sqrt{t_0}$ and w_0, and find $\sqrt{t_0} = 0.1416(^{+8}_{-5})$ fm and $w_0 = 0.1717(^{+12}_{-11})$ fm, in reasonable agreement with most independent calculations, and with excellent agreement with the results of HPQCD, who used a subset of the same HISQ ensembles employed here. We have estimated an upper bound on the integrated autocorrelation lengths at different lattice spacings and found no autocorrelation lengths above 65 molecular-dynamics time units. Compared to our preliminary work, the continuum extrapolation is better controlled through the removal of tree-level discretization errors and the use of aF_{p4s} over af_π to set the scale, the quark-mass interpolation is constrained using chiral perturbation theory, and the charm-quark mass has been adjusted to correct for mistunings. Finally, we have calculated the continuum meson-mass dependence for use in future scale-setting applications.

Chapter 3

Chiral Analysis

This chapter develops preliminary results for the quark masses and light pseudoscalar decay constants. The results are the culmination of several improvements to previous MILC results, including the highly improved staggered quark action, inclusion of the charm quark as a fourth dynamical flavor, finer lattices with light quarks near the physical light masses, using gradient flow for improved relative scale setting, and more.

Because quarks and gluons are the elementary particles of QCD, any refinement of quark properties is useful to a wide range of other applications. For example, quark masses are involved in determinations of the Higgs coupling to QCD and flavor structures of beyond Standard Model physics. Light quark masses are especially useful to studies of chiral perturbation theory as they determine the amount of explicit chiral symmetry breaking.

The primary use of the pseudoscalar decay constants and other form factors are their impact on determinations of the CKM matrix elements. For example, in the purely leptonic decay of a kaon into an intermediate W boson, the decay width depends on both the QCD based decay constant f_{K^+} and weak matrix decay element $|V_{us}|$. By determining the decay constants precisely in lattice computations and combining these results with experimental decay widths, the strength of the flavor-changing CKM elements can be determined. Similar procedures can be applied to extract $|V_{us}|$ from the kaon semileptonic form factor $f_+(0)$ and experimental results for K semileptonic decays [9]. In both cases, $|V_{us}|$ can in turn be used to bound potential unitarity violations of the CKM matrix predicted by beyond Standard Model physics. In addition, a clearly significant difference between the values of $|V_{us}|$ extracted from each decay would indicate a failure of the Standard Model, and the existence of new physics.

After a short introduction, this chapter is broken into two primary parts: theory and analysis. The primary goal of the theory section is providing formulas for masses and decay constants developed from partially quenched, rooted staggered chiral perturbation theory. The analysis section then incorporates these formulas to extrapolate MILC HISQ ensemble results to the continuum, physical quark-mass limit.

3.1 Introduction

For lattice observables to be directly compared to experimental results they must be determined at realistic values of the quark masses and in the continuum. However, including simulations at quark masses that are heavier than physical light quark masses still provides several advantages.

1. Gauge configuration generation algorithms slow down as the quark mass becomes lighter. Therefore, more precise lattice results can be generated at heavier quarks for the same computational cost.

2. Quark masses can only be approximated before the simulation is run. Having simulations with additional quark masses allows results to be interpolated to the precise physical quark masses with lower systematic errors

3. Several results of significance to chiral effective theories need to be computed at the chiral limit (vanishing quark mass). From this perspective, the physical light mass is just one of several useful points for controlling extrapolations to the chiral limit.

4. In several lattice effective theories, such as staggered chiral perturbation theory, the continuum extrapolation and quark mass interpolation are related. Therefore, improving control of the quark mass interpolation by including a wider variety of masses also helps reduce systematic errors in the continuum extrapolation.

To achieve high precision lattice results, these simulations at non-physical quark masses are essential.

A related essential tool for achieving high precision lattice results, especially when additional computational resources are not an option, are Symanzik effective theories (SET) [80]. A SET models a given lattice action as a continuum effective field theory plus higher-dimension correction operators multiplied by powers of the lattice spacing, where the functional form is the most general expression up to a certain order in a that maintains all the symmetries of the lattice action. This allows models for continuum extrapolations to be constrained based on the specific action being simulated, reducing the resulting systematic errors on continuum results. Lattice artifacts can also be reduced at finite lattice spacing by using the terms of the effective Lagrangian to create improved lattice actions that vanish up to a given order in the SET.

In this work, SET is used to extend the meson momentum p^2 and quark mass m_q expansion of χPT to include the discretization corrections from staggered quarks in powers of a^2. The combined effective theory, called staggered chiral perturbation theory, then provides a systematic framework for simultaneously extrapolating in the lattice spacing a and quark mass m_q.

3.2 Theoretical Framework

In Secs. 1.1.1 and 1.1.5 the foundations for continuum chiral perturbation theory (χPT) were outlined, based on the spontaneously broken left and right-handed chiral symmetries of the continuum theory. The associated continuum Lagrangian could in principle be used to model the quark mass dependence of lattice observables; however, this is only possible after results are extrapolated close enough to the continuum that remaining discretization artifacts are negligible. In the case of staggered fermions, this approach is not practical due to the significant size of taste artifacts that depend on the lattice spacing. Instead, χPT results are reformulated to explicitly account for staggered quarks and allow quark-mass and continuum extrapolations to be performed simultaneously.

This section focuses on the reformulation of χPT for staggered quarks on the lattice, called staggered chiral perturbation theory (SχPT). With the addition of the fourth root procedure the theory is called rooted staggered chiral perturbation theory (rSχPT). Section 3.2.1 focuses on deriving the rSχPT Lagrangian from the new symmetries of the lattice and replica procedure. Partial quenching is also briefly discussed. This allows for quark masses in internal loops to be different from external quark masses, improving the available range of external quark masses without necessitating completely new sets of gauge configurations.

Finally, the results for light pseudoscalar masses and decay constants in partially quenched rSχPT are outlined in Sec. 3.2.2.

3.2.1 Staggered Chiral Perturbation Theory [4, 5, 53]

Staggered Symanzik Effective Theory

To extend continuum chiral perturbation theory to staggered quarks, a Symanzik effective theory (SET) must be developed [80]. The theory only applies to physical momenta, $p \ll \pi/a$. In other words, the lattice spacing is small enough that the momentum cutoff, π/a, is much larger than any physical momentum of interest. The general form of a SET is therefore an expansion in inverse powers of the cutoff

$$L_{\mathrm{SET}} = \sum_{n=0}^{\infty} a^n L_{(n+4)}, \tag{3.1}$$

where $L_{(n+4)}$ is of mass dimension $n + 4$. Each term $L_{(n+4)}$ must contain the most generic expression of dimension $n + 4$ that still satisfies all the symmetries of the original lattice Lagrangian. Furthermore, symmetries that appear in the different limits of the lattice Lagrangian must be respected by L_{SET}. For example, the staggered fermion action has an exact $SU(4)$ taste symmetry and an approximate $SU_L(4N_f) \times SU_R(4N_f)$ chiral symmetry in the continuum limit. This symmetry will be preserved in $L_{(4)}$, though the higher dimension terms may break either or both.

To finish building the SET, each term in Eq. (3.1) must be identified up to the desired order in a. The leading order term $L_{(4)}$ must be the Lagrangian of continuum QCD, with the corresponding χPT. In the staggered formalism, there are no dimension five operators that satisfy all the lattice symmetries. Therefore, the corrections to the continuum χPT Lagrangian start at $O(a^2)$. These corrections appear as four-quark operators (dimension $3/2 \times 4 = 6$), where one or more gluons with net momentum π/a are exchanged between two quark lines. Because the exchanged gluons could change color, spin, or taste, but not flavor, the flavors of the quark and antiquark fields creating and annihilating each line must be summed together to ensure the correction is a scalar and color singlet. The possible transformations of the remaining degrees of freedom can be captured by inserting one of the associated matrices of $SU(3)$ or $SU(4)$ for color and spin/taste, respectively.

$$O_{ss'tt'cc'} = \bar{q}_i \left(\gamma_s \otimes \xi_t \otimes \lambda_c\right) q_i \bar{q}_j \left(\gamma_{s'} \otimes \xi_{t'} \otimes \lambda_{c'}\right) q_j \tag{3.2}$$

In the preceeding equation i, j denote flavor, s, s' denote spin, t, t' denote taste, and c, c' denote color indices. To establish SχPT the combined chiral and taste indices need to be explicit for each operator so that a spurion analysis can be performed (see Sec. 1.1.5); however, the chiral symmetry is independent of the color space. Therefore, the color indices are not explicitly needed to match terms in SET to those in SχPT. Also, while the coefficients $d_{cc'}$ of each operator are potentially different, when they are matched to the encompassing SχPT operator the coefficient will be arbitrary. Therefore, for this work, all distinct color singlet contractions are implicitly summed over and a simplified set of operators are considered.

$$O_{ss'tt'} = \sum_{cc'} d_{cc'} O_{ss'tt'cc'} = \bar{q}_i \left(\gamma_s \otimes \xi_t\right) q_i \bar{q}_j \left(\gamma_{s'} \otimes \xi_{t'}\right) q_j \tag{3.3}$$

To find the set of allowed $O_{ss'tt'}$ for the SET the staggered symmetries must be enforced. The four symmetries that need to be considered to fully constrain the operators are the $U(1)_\epsilon$ symmetry of Eq. (1.68), shift symmetry, rotations, and parity. Each symmetry leads to a specific new constraint.

The $U(1)_A$ symmetry can be applied to each flavor independently, which requires each of the two quark bilinears in $O_{ss'tt'}$ to be invariant. For the quark bilinear to be invariant, $\gamma_s \otimes \xi_t$ must anticommute with $\gamma_5 \otimes \xi_5$. In other words, one of either the spin or taste generators must anticommute with γ_5 or ξ_5 while the remaining generator commutes. This implies scalar (S), pseudoscalar (P), and tensor (T) generators can only be paired with vector (V), and axial-vector (A) generators.

Under the appropriate redefinition of the staggered fields $q(y)$, the shift symmetry can be defined as a vector taste transformation of the staggered field

$$q(y) \to (I \otimes \xi_\mu) q(y), \qquad \bar{q}(y) \to \bar{q}(y)(I \otimes \xi_\mu). \tag{3.4}$$

Under this transformation, each quark bilinear ismultiplied by a factor of ±1 depending on the value of t and μ. For $O_{ss'tt'}$ to remain invariant, each quark bilinear must be multiplied by the same factor for every value of μ. Since the set of ±1 factors is distinct for each value of ξ_t, this implies $\xi_t = \xi_{t'}$.

Finally, the rotational and parity symmetries are enforced. On the lattice, the rotational symmetry is limited to 90° rotations. This slightly relaxes the usual Lorentz constraint on contracting indexes; Lorentz indices (space-time, spin, and taste) can be repeated any even number of times before contracting, not just twice. Because $\chi_t = \chi_{t'}$, γ_s must be able to contract with $\gamma_{s'}$. Parity then requires that $\gamma_s = \gamma_{s'}$. Finally the possible contractions of spin and taste indices must be determined.

There are two ways the spin and taste indices may be contracted, separately or together. Denote the two types of operators as type A and type B, respectively. If the spin and taste indices are contracted separately, then there are no additional constraints on the allowed values of t, t', s and s'. There are therefore 12 allowed type-A operators corresponding to the 12 allowed quark bilinear taste-spin pairings.

$$\begin{aligned}\{&[S \otimes V],\ [S \otimes A],\ [P \otimes V],\ [P \otimes A],\ [T \otimes V],\ [T \otimes A],\\ &[V \otimes S],\ [V \otimes P],\ [V \otimes T],\ [A \otimes S],\ [A \otimes P],\ [A \otimes T]\}\end{aligned} \tag{3.5}$$

If the spin and taste indices are contracted together, then there must be four, six, or eight spin/taste indices available to contract. However, vector and axial-vector spins cannot be paired with vector or axial-vector tastes in the quark bilinears due to the $U(1)_A$ symmetry, and similarly for tensor taste/spins. Therefore, the only available contraction is with six spin/taste indices. This requires one of the spin or taste matrices to be a tensor, with the other being a vector or axial-vector. There are therefore four type-B operators, again corresponding to the allowed quark bilinear taste-spin pairings.

$$\{[V \otimes T],\ [A \otimes T],\ [T \otimes V],\ [T \otimes A]\} \tag{3.6}$$

Power Counting

The full SχPT Lagrangian is constructed by systematically expanding in powers of the momentum p^2, quark masses m_q, and taste splittings $a^2\delta$. The effects of the first two parameters have been explored in Sec. 1.1.5; the momentum p^2 of the pseudo-Goldstone bosons arises from the chirally-invariant partial derivatives in the continuum Lagrangian, and the quark masses μm_q arise from the explicitly chiral-symmetry-braking quark-mass matrix M. The last parameter, the pseudo-Goldstone taste splittings, is a new parameter arising from the introduction of the discretization errors of the staggered quarks into chiral perturbation theory.

In order to expand simultaneously in powers of all three parameters, a power counting scheme needs to be established. Simulation results show that the typical taste splitting of squared meson masses constructed from two light quarks is about the same order as M_π^2. [12]. Therefore, for light pseudoscalar observables, $p^2 \sim M^2 \sim \mu m_q \sim a^2\delta$. This means the taste breaking artifacts of the staggered formalism are considered leading order and must be included in the leading order SχPT Lagrangian. However, care must be made to revisit this power counting scheme as lattice actions improve and taste-artifacts are reduced. Also, if one were to focus exclusively on ensembles with $a \approx 0.09$fm or finer, it might be possible to consider taste splitting artifacts as next-to-leading order.

Staggered Chiral Perturbation Theory

Just as in the derivation of continuum χPT, SχPT is built on the assumption that the full chiral symmetry in the continuum, chiral limit is spontaneously broken to the vector symmetry. One difference introduced by the staggered fermions is the additional continuum

$SU(4)$ taste symmetry. This implies the full chiral symmetry in the continuum limit is a larger $SU_L(4N_f) \times SU_R(4N_f) \times U(1)_V$ symmetry. When this symmetry is spontaneously broken to the vector subgroup, $SU(4N_f) \times U(1)_V$, the taste symmetry remains and $(4N_f)^2 - 1$ Goldstone bosons are created. In a similar fashion to what was done before, the Goldstone-boson fields are packaged into a single field

$$\Sigma(x) = \exp i\Phi f \tag{3.7}$$

where Φ is a $4N_f \times 4N_f$ matrix and the factor of two is dropped so that the additional factors of 4 from traces over taste in the continuum Lagrangian are cancelled. The elements of Φ are have distinct flavors and taste. When expressed in a flavor-block form, Φ is

$$\Phi = \begin{pmatrix} U & \pi^+ & K^+ & \cdots \\ \pi^- & D & K^0 & \cdots \\ K^- & \bar{K}^0 & S & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{3.8}$$

where each 4×4 block Q is a matrix in taste space expanded in the 16 taste generators T_a.

$$Q = \sum_{a=1}^{16} Q_a T_a, \qquad T_a = \{\xi_5, i\xi_{\mu 5}, i\xi_{\mu\nu}(\mu < \nu), \xi_\mu, I\} \tag{3.9}$$

With these definitions, the leading order continuum chiral Lagrangian is

$$L_{\mathrm{LO}}^{\mathrm{cont}} = \frac{f^2}{8}\mathrm{Tr}\left(\partial_\mu \Sigma \partial_\mu \Sigma^\dagger\right) - \frac{1}{4}\mu f^2 \mathrm{Tr}\left(M\Sigma^\dagger + M\Sigma\right) + \frac{m_0^2}{24}[\mathrm{Tr}(\Phi)]^2 \tag{3.10}$$

where the last term is the only one not already derived in the Lagrangian in Eq. (1.45). This term accounts for the anomalous contribution to the mass of η', which is a flavor and taste singlet. The η' will decouple as it gets heavier [75]. Normally, Φ must be traceless in order to ensure Σ has determinant 1; however, this complicates the quark content of the diagonal elements in Φ. Including m_0 explicitly in the Lagrangian and delaying taking the limit $m_0 \to \infty$ allows conditions on the trace of Φ to be avoided, which simplifies U, D, and the other diagonal elements to $u\bar{u}$, $d\bar{d}$, etc.

To extend the continuum-limit staggered chiral Lagrangian to finite lattice spacing, $O(a^2)$ chiral operators need to be added that reflect the same sets of symmetries as the four quark operators of the SET. To start, the original quark fields are expanded into left and right handed components $q_i = q_i^L + q_i^R$ using the usual projections. In each quark bilinear, the substitution results in four possible terms

$$\bar{q}A_{s,t}q = \bar{q}^L A_{s,t} q^L + \bar{q}^R A_{s,t} q^R + \bar{q}^L A_{s,t} q^R + \bar{q}^R A_{s,t} q^L \tag{3.11}$$

where $A_{s,t} = \gamma_s \otimes \xi_t$. If γ_s is S, P, or T then $[\gamma_5, \gamma_s] = 0$ and the terms with uniform chirality vanish. For all other cases γ_s is either V or A and the terms with mixed chirality vanish.

$$\bar{q}A_{s,t}q = \begin{cases} \bar{q}^L A_{s,t} q^L + \bar{q}^R A_{s,t} q^R & \gamma_s = \mathrm{V}, \mathrm{A} \\ \bar{q}^L A_{s,t} q^R + \bar{q}^R A_{s,t} q^L & \gamma_s = \mathrm{S}, \mathrm{P}, \mathrm{T} \end{cases} \tag{3.12}$$

Because these operators explicitly break the combined chiral-taste symmetry (except when ξ_t = S), the taste matrices ξ_t are temporarily elevated to the spurions F_{t1} and F_{t2} with transformation properties designed to keep the symmetry in tact. The corresponding transformation properties in $SU_L(4N_f) \times SU_R(4N_f)$ depend on the chirality of the quark fields surrounding it.

$$\bar{q}\xi_t q \to \begin{cases} \left(\bar{q}^L L^\dagger\right)\left(L F_{t1} L^\dagger\right)\left(L q^L\right) + \left(\bar{q}^R R^\dagger\right)\left(R F_{t2} R^\dagger\right)\left(R q^R\right) & \xi_t = \mathrm{P}, \mathrm{T} \\ \left(\bar{q}^L L^\dagger\right)\left(L F_{t1} R^\dagger\right)\left(R q^R\right) + \left(\bar{q}^R R^\dagger\right)\left(R F_{t2} L^\dagger\right)\left(L q^L\right) & \xi_t = \mathrm{V}, \mathrm{A} \end{cases} \tag{3.13}$$

After the Lagrangian is built, both spurions are set equal to the values they replaced

$$F_{t1,t2} \to a\xi_t \otimes I_{\text{flavor}} \tag{3.14}$$

The complete set of operators now transforming under $SU_L(4N_f) \times SU_R(4N_f)$ rotations includes Σ, M, their hermitian conjugates, and the set of 10 spurions F_{t1}, F_{t2} for $\xi_t = [S, P, V, A, T]$. In order to maintain shift symmetry, F_{t1} and F_{t2} must contract among themselves after the spurions are set to their original values. This guarantees that any term with a spurion has another spurion and is at least $O(a^2)$. The mass matrices M and $M^\dagger$ and derivatives $\partial_\mu \Sigma$ and $\partial_\mu \Sigma^\dagger$ can not be included with the spurion fields at $O(a^2)$. This is because the mass matrices are $O(m_q) \sim O(a^2)$[5] and the derivatives must always come in pairs with $O(p^2) \sim O(a^2)$. Therefore, the only fields that can combine with spurions at leading order are Σ or $\Sigma^\dagger$.

---

[5] In such expressions, factors of the QCD scale Λ_{QCD}, needed to make the dimensions match, are omitted for simplicity.

For the case of $\xi_t = P$ or T there is only one possible term

$$\mathrm{Tr}(F_{t1}\Sigma F_{t2}\Sigma^\dagger) \to a^2\mathrm{Tr}(\xi_t\Sigma\xi_t^\dagger\Sigma^\dagger) \tag{3.15}$$

For the case of $\xi_t = V$ or A there are three possible terms

$$\mathrm{Tr}(F_{t1}\Sigma^\dagger)\mathrm{Tr}(F_{t2}\Sigma) \quad\to\quad a^2\mathrm{Tr}(\xi_t\Sigma^\dagger)\mathrm{Tr}(\xi_t^\dagger\Sigma) \tag{3.16}$$

$$\mathrm{Tr}(F_{t1}\Sigma^\dagger)\mathrm{Tr}(F_{t1}\Sigma^\dagger)+\mathrm{h.c.} \quad\to\quad a^2\mathrm{Tr}(\xi_t\Sigma^\dagger)\mathrm{Tr}(\xi_t\Sigma^\dagger)+\mathrm{h.c.} \tag{3.17}$$

$$\mathrm{Tr}(F_{t1}\Sigma^\dagger F_{t1}\Sigma^\dagger)+\mathrm{h.c.} \quad\to\quad a^2\mathrm{Tr}(\xi_t\Sigma^\dagger\xi_t\Sigma^\dagger)+\mathrm{h.c.} \tag{3.18}$$

where h.c. stands for hermitian conjugate and $F_{t1}^\dagger = F_{t2}$. Adding everything up leads to eight linearly independent chiral operators of type-A at leading order.

Due to the contraction of space-time and taste indices, type-B operators must contain derivatives $\partial_\mu\Sigma$ or $\partial_\mu\Sigma^\dagger$ to provide the space-time indices that contract with taste indices on ξ_t. Because the derivative and taste matrix pairs each provide factors of $O(a^2)$, the overall order of any type-B operator is at least $O(a^4)$. Therefore, they can be neglected at leading order. This leaves the eight type-A operators as the only terms. The complete, leading-order SχPT Lagrangian is

$$L_{\mathrm{LO}} = \frac{f^2}{8}\mathrm{Tr}(\partial_\mu\Sigma\partial_\mu\Sigma^\dagger) - \frac{1}{4}\mu f^2\mathrm{Tr}(M\Sigma + M\Sigma^\dagger) + \frac{m_0^2}{24}(\mathrm{Tr}(\Phi))^2 + a^2\mathcal{V}\ , \tag{3.19}$$

with the taste-violating potential

$$\begin{aligned}
-\mathcal{V} \;=\;& C_1\mathrm{Tr}(\xi_5\Sigma\xi_5\Sigma^\dagger) + \frac{C_3}{2}[\mathrm{Tr}(\xi_\nu\Sigma\xi_\nu\Sigma) + \mathrm{h.c.}] \\
&+\frac{C_4}{2}[\mathrm{Tr}(\xi_{\nu5}\Sigma\xi_{5\nu}\Sigma) + \mathrm{h.c.}] + \frac{C_6}{2}\ \mathrm{Tr}(\xi_{\mu\nu}\Sigma\xi_{\nu\mu}\Sigma^\dagger) \\
&+\frac{C_{2V}}{4}[\mathrm{Tr}(\xi_\nu\Sigma)\mathrm{Tr}(\xi_\nu\Sigma) + \mathrm{h.c.}] + \frac{C_{2A}}{4}[\mathrm{Tr}(\xi_{\nu5}\Sigma)\mathrm{Tr}(\xi_{5\nu}\Sigma) + \mathrm{h.c.}] \\
&+\frac{C_{5V}}{2}[\mathrm{Tr}(\xi_\nu\Sigma)\mathrm{Tr}(\xi_\nu\Sigma^\dagger)] + \frac{C_{5A}}{2}[\mathrm{Tr}(\xi_{\nu5}\Sigma)\mathrm{Tr}(\xi_{5\nu}\Sigma^\dagger)],
\end{aligned} \tag{3.20}$$

Replicas, Rooting, and Partial Quenching SχPT

With the leading order Lagrangian constructed for SχPT, the fourth root procedure needs to be accounted for in order to remove the unwanted tastes from the chiral effective theory. These extra tastes enter expectation values through the fermion operator determinant, and a determinant factor appears for each flavor and taste of quark appearing in the action. In Feynman diagrams, contributions from the determinant correspond to closed internal quark loops, and taking the determinant to the 1/4 power is equivalent to multiplying each such loop by factor of 1/4. Therefore, one way to address the effect of extra tastes would be to parameterize the number of quarks in the action, evaluate observables, and then take the limit as the parameter approaches 1/4. This is known as the replica procedure [20, 38, 22].

To perform the replica procedure, each quark (flavor and taste) is replicated in the action n_r times. This replication factor could be distinct for each flavor and taste; however, it only *needs* to be distinct if the final value is going to be different. In this case, every unaltered staggered quark is left with four degenerate tastes in the continuum limit. So, instead of setting $n_r = 1$ after computing expansions, n_r is set to 1/4.

An alternative way to to interpret the replica procedure that produces equivalent results, is through quark flow [72, 4]. In quark flow, each term and corresponding meson diagrams are re-expressed in terms of the underlying quark fields. For some diagrams, this results in internal quark lines that form completely closed loops, such as in Fig. 3.3(a). For each such loop, a factor of 1/4 can be manually inserted to remove the unwanted tastes. While the quark flow procedure is not obviously systematic, like the replica procedure, it does provide a clearer illustration of how the quarks are acting within χPT diagrams in QCD.

Replicas can be applied in more cases than the fourth-root procedure. Any scenario where the number of quarks participating in an internal loop needs to be controlled is amenable to the technique. One such scenario is partial quenching [24]. In partial quenching, the quark masses appearing in expectation values are split into two sets: valence $\{m_{\text{val}}\}$ and sea $\{m_{\text{sea}}\}$.[6] The valence quark masses $\{m_{\text{val}}\}$ are present in operator expectation values, while the sea quark masses $\{m_{\text{sea}}\}$ contribute through the determinant and therefore only appear in internal loops.

$$\langle O \rangle = \int [dU] \det[M(\{m_{\text{sea}}\})] e^{-S[\{m_{\text{sea}}\},\{m_{\text{val}}\}]} \hat{O}(\{m_{\text{val}}\}) \tag{3.21}$$

To apply the replica procedure to reproduce the effects of partial quenching in χPT, new flavors of quarks are added to the action that represent the valence quarks. Then, the replica procedure is used separately for the sea and valence quark sets. Valence quarks need to be removed from internal quark loops, so their replication factor is set to zero. Sea quarks still act as normal ($n_r = 1$ in SχPT, or $n_r = 1/4$ in rSχPT).

---

[6] The separation of quarks used by partial quenching can be generalized to mixed actions where different quark actions are used for the valence and sea; this work only focuses on the use of different masses.

The reason why partial quenching is useful simplifies down to reducing systematic errors with minimal cost. The low energy constants (LECs) of rSχPT scale different linear combinations of the sea and valence quark masses. By manipulating the valence quark masses separately from sea quark masses, the dependence of observables on each of the LECs can be extracted. Additionally, the computational cost associated with generating different gauge ensembles is usually much higher than the cost of evaluating correlation functions on the ensembles. It is therefore cheaper to get more statistics through varying the valence masses than the sea masses. Modern simulations usually choose many different values of valence quark masses for each particular choice of sea quark masses.

3.2.2 Masses and Decay Constants

Meson Masses at Leading Order

To find the leading order corrections to the Goldstone masses, Eq. (3.19) is expanded to quadratic order in the field Φ. Both the chiral-violating quark mass term and taste-violating potential may in principle alter the masses of the Goldstone bosons. The effects from the mass term were previously worked out in Eq. (1.46). Now the contributions from $\mathcal{V}$ may be added.

Upon a quick inspection of the potential $\mathcal{V}$, there are two types of terms distinguished by the number of traces multiplied together. When expanded to quadratic order in Φ, the terms with two traces multiplied together have only one field in each trace and therefore do not mix flavors. Therefore, terms with the product of two traces only result in mass terms for the flavor-neutral mesons U, D, S, etc. at leading order (LO). On the other hand, terms with a single trace will mix flavors even at LO and are therefore capable of altering any meson mass. The full potential $\mathcal{V}$ is often split into two potentials U, U' based on this property. The potential U collects all of the terms with a single trace, and U' collects all the terms with a product of two traces.

Focusing on the flavor mixing potential U, the corrections to the meson masses at tree level are all of the form

$$M^2_{xy,t} = \mu(m_x + m_y) + a^2\Delta(\xi_t) \tag{3.22}$$

where m_x and m_y are the masses of the valence masses making up the meson and $a^2\Delta(\xi_t)$ is the taste splitting arising from U. The taste-splitting factor $\Delta(\xi_t)$ does not depend on the lattice spacing (except weakly through powers of α_s) or quark masses (sea or valence). The factor does depend on the taste-structure of the meson, denoted by ξ_t. This is a result of the 'accidental' $SO(4)$ rotational symmetry preserved by type A four quark operators, but not by type B operators. Since type B operators do not enter at this order of the Lagrangian, the taste splittings are degenerate within any of the groups $\{\mathrm{P}, \mathrm{V}, \mathrm{T}, \mathrm{A}, \mathrm{S}\}$. The five different taste splittings are

$$\begin{aligned}
\Delta(\xi_5) &= 0 && (3.23)\\
\Delta(\xi_\mu) &= \frac{16}{f^2}\left(C_1 + 3C_3 + C_4 + 3C_6\right) && (3.24)\\
\Delta(\xi_{\mu\nu}) &= \frac{16}{f^2}\left(2C_3 + 2C_4 + 4C_6\right) && (3.25)\\
\Delta(\xi_{\mu 5}) &= \frac{16}{f^2}\left(C_1 + C_3 + 3C_4 + 3C_6\right) && (3.26)\\
\Delta(I) &= \frac{16}{f^2}\left(4C_3 + 4C_4\right) && (3.27)
\end{aligned}$$

The taste-pseudoscalar meson is the only one with no taste splitting, due to the $U_\epsilon(1)$ axial symmetry for each flavor.

For flavor-mixing mesons, Eq. (3.22) is sufficient because the remaining contributions from U' can only effect flavor-neutral mesons. For the flavor-neutral mesons, expanding U' to quadratic order reveals two new vertices that mix the flavor-neutral mesons with each other. The effect of these vertices on the flavor-neutral propagators needs to be combined with the explicit mass terms of U. Once the new propagators are derived, the masses can be inferred from the poles of the full propagator.

Due to the $SO(4)$ symmetry at this order, the different vertices can be grouped by the taste structure of the mesons, similar to the taste splittings from U. No new vertices appear for the scalar, pseudoscalar, and tensor tastes because every term in U' only contains traces of Σ with the vector ξ_μ or axial-vector $\xi_{\mu 5}$ generators. The remaining taste structures, V and A, appear symmetrically in the potential U'. This implies that the $O(\Phi^2)$ contributions to the Lagrangian and associated vertices for each taste structure will follow the same form (just substitute V or A for t).

$$\begin{aligned} L_{\mathrm{LO}} &= \ldots + \frac{1}{2}a^2\delta'_t\left(U_t + D_t + S_t + \ldots\right)^2 + \ldots && (3.28) \\ -a^2\delta'_t &\equiv -a^2\frac{16}{f^2}\left(C_{2t} - C_{5t}\right) && (3.29) \end{aligned}$$

In the expansion of $(U + D + S + \ldots)^2$ it is clear that each vertex allows one flavor neutral meson to connect to another, with potentially different flavor. While the chiral vertices are clearly connected, the underlying quark flow diagram is not connected. Because of the

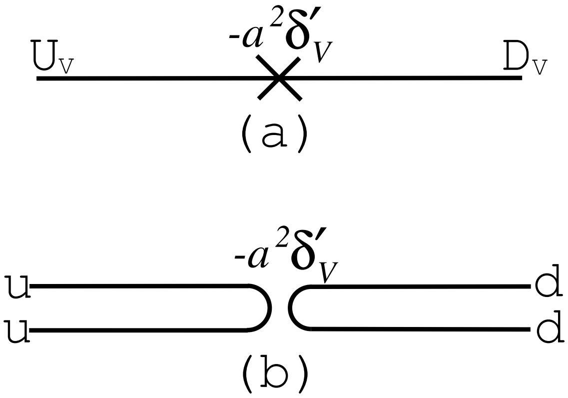

Figure 3.1: An illustration of the hairpin diagram for mixing a vector U with a vector D meson [4]. The sub figure (a) shows the vertex in the chiral theory while (b) shows the disconnect quark flow diagram. The taste and flavor of the mesons may be changed to any exchange of flavor neutral mesons U, D, and S and any one of the tastes A, V, I. Both diagrams only vary by the vertex factor and exterior labels when U and D under such an exchange.

underlying disconnected quark flow, the diagram referring to these first order corrections to the flavor-neutral propagators in chiral effective theory is often referred to as a hairpin diagram (see Fig. 3.1).

To reformulate the neutral meson propagators and extract their masses, the hairpin diagram needs to be summed to all orders with all possible internal meson lines. A convenient procedure for carrying out the re-summation was first introduced in [74]. Some of the important results are summarized here.

Denote the free, meson propagator with neutral flavor M (e.g. U, D, S, ...) and taste structure t by $G_{0,t,M}$. The inverse free propagator is read off from the Lagrangian as

$$G^{-1}_{0,t,M} = q^2 + m^2_{M_t}, \tag{3.30}$$

where $m^2_{M_t}$ are the masses with corrections from U in Eq. (3.22). In the interacting theory, there will be propagators connecting one flavor to another. Is is therefore useful to promote the free propagator into flavor space.

$$\left(G^{-1}_{0,t}\right)_{MN} = G^{-1}_{0,t,M}\delta_{MN} \tag{3.31}$$

In flavor space, the interaction vertices for a given taste H_t can then be combined with the inverse free propagator to form the inverse propagator G^{-1}_t in the interacting theory.

$$G^{-1}_t = G^{-1}_{0,t} + H_t, \qquad (H_t)_{MN} = a^2\delta'_t \tag{3.32}$$

Note, the elements of H_t are constant for all flavors M and N.

Performing the re-summation of the hairpin diagrams is equivalent to inverting Eq. (3.32). After performing several manipulations, the fully interacting propagator is

$$G_t = G_{0,t} + \mathcal{D}_t, \qquad \mathcal{D}_t = -G_{0,t}H_tG_{0,t}\frac{\det G^{-1}_{0,t}}{\det G^{-1}_t} \tag{3.33}$$

The elements of the disconnected contribution $\mathcal{D}_t$ can be expressed explicitly in terms of the quark masses by expanding each determinant in terms of the operator's eigenvalues.

$$(\mathcal{D}_t)_{MN} = -a^2\delta'_t\frac{1}{(q^2+m^2_{M_t})(q^2+m^2_{N_t})}\frac{\prod_O(q^2+m^2_{O_t})}{\prod_F(q^2+m^2_{F_t})} \tag{3.34}$$

The first two factors in the denominator come from $(G_0 H G_0)_{MN}$ and are due to the external mesons of the propagator. If directly evaluating the two-point correlation function, the masses m_{M_t} and m_{Nt} will be the valence quark masses of the external lines. If the propagator appears in an internal loop, then these masses will correspond to each of the possible sea and valence quark masses allowed in the loop.

The second fraction contains the two eigenvalue expansions of the determinants. The numerator is a product over the eigenvalues of the inverse, free propagator, with masses for each of the flavor-neutral mesons in the original basis: U_t, D_t, S_t, etc. The denominator is a product over the eigenvalues of the inverse, full propagator. As can be inferred from Eq. (3.33), the poles of the full propagator occur at the eigenvalues of the full, neutral-meson, mass matrix. For both of these eigenvalue products, the masses come from the iteration over internal quark loops formed between the vertices of two hairpin diagrams. As such, only sea quark masses will appear.

With the full propagator derived in terms of the quark masses, the poles can now be extracted. The two poles corresponding to the external quark masses (valence or sea) are easy to infer. Since they come from the inverse, free propagator $G_{0,t}^{-1}$ they will correspond to the masses already derived from U in Eq. (3.22). The only remaining poles are each of the $m_{F_t}^2$ resulting from the eigenvalues of the inverse, full propagator G_t^{-1}. The mass matrix that needs to be diagonalized is X, where

$$G_t^{-1} = q^2 I + X_t, \qquad X_t = \begin{pmatrix} m_{U_t}^2 + a^2\delta_t & a^2\delta_t & a^2\delta_t & \dots \\ a^2\delta_t & m_{D_t}^2 + a^2\delta_t & a^2\delta_t & \dots \\ a^2\delta_t & a^2\delta_t & m_{S_t}^2 + a^2\delta_t & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{3.35}$$

In the case of $N_f = 2+1$ flavors without replication ($n_r = 1$), the eigenvalues are,

$$\begin{aligned} m^2_{\pi^0_t} &= m^2_{U_t} = m^2_{D_t} && (3.36)\\ m^2_{\eta_t} &= \frac{1}{2}\left(m^2_{U_t} + m^2_{S_t} + 3a^2\delta'_t - Z\right) && (3.37)\\ m^2_{\eta'_t} &= \frac{1}{2}\left(m^2_{U_t} + m^2_{S_t} + 3a^2\delta'_t + Z\right) && (3.38)\\ Z &= \sqrt{\left(m^2_{S_t} - m^2_{U_t}\right)^2 - 2a^2\delta'_t\left(m^2_{S_t} - m^2_{U_t}\right) + 9(a^2\delta'_t)^2} \end{aligned}$$

When replicas are taken into account two important modifications occur. With n_r replicas of N_f flavors, there will be a total of $n_r \times N_f$ eigenvalues. Out of all the eigenvalues, most will be one of the N_f distinct values corresponding to the original mass basis $m^2_{U_t}$, $m^2_{D_t}$, m^{S_t} etc, each of which has a multiplicity of $n_r - 1$. These will cancel against the additional $n_r - 1$ factor for each original mass in the numerator of Eq. (3.34). The remaining N_f eigenvalues correspond to more complicated combinations of the original mass basis and are almost the same as the results in Eqs. (3.36) through (3.38). The only difference introduced into these equations by the replica procedure is the replacement of δ'_t by $n_r\delta'_t$.

Before moving on to the one loop calculations, there is one more interaction in the leading order Lagrangian whose effects need to be determined: the singlet mass term

$$\frac{1}{24}m_0^2\left(\mathrm{Tr}\Phi\right)^2 = \frac{2}{3}m_0^2\left(U_I + D_I + S_I + \ldots\right)^2. \qquad (3.39)$$

The singlet interaction is almost identical to the previous vector and axial taste interactions for flavor-neutral mesons. The only differences are the vertex factor of $-\frac{4}{3}m_0^2$ instead of $-a^2\delta'_t$, and the taste index referring to the singlet instead of vector or axial taste. Both of these differences can be accounted for through a simple substitution into the previous results, with one additional step: taking the limit $m_0 \to \infty$ for observable quantities. Evaluating Eqs. (3.36) through (3.38) with the above substitutions and with the replica factor of $n_r = 1/4$ taken into account yields

$$
\begin{aligned}
m^2_{\pi^0_I} &= m^2_{U_I} = m^2_{D_I} && (3.40)\\
m^2_{\eta_I} &= \frac{1}{3}\left(m^2_{U_I} + 2m^2_{S_I}\right) && (3.41)\\
m^2_{\eta'_I} &\sim m^2_0. && (3.42)
\end{aligned}
$$

Taking the limit $m_0 \to \infty$ in Eq. (3.34) leads to a cancellation between the explicit vertex factor δ'_I and the denominator factor of $(q^2 + m^2_0)$. Thus, the taste singlet η' decouples from the theory.

Next-to-Leading Order rSχPT Results

With the leading order (LO) rSχPT Lagrangian in hand, one loop results can now be calculated. While results could be calculated for any of the non-singlet taste structures, of particular interest in this work are the true Goldstone (pseudoscalar-taste) pseudoscalar meson masses and decay constants. Furthermore, all results are evaluated in the partially quenched theory with three dynamical quark flavors (u, d, and s) and two valence flavors (x and y). The pseudoscalar mesons of interest are therefore composed of two valence quarks x and $\bar{y}$. Let these mesons be denoted by P^+_5, where the taste structure is explicit but the quark content is left implicit.

To find next-to-leading order (NLO) results the one loop results need to be combined with new analytic terms at the same order (i.e. second order in the small quantities $p^2 \sim m \sim a^2$). This section focuses on the derivation of the chiral log contributions at one loop and combines the results with analytic terms defined in [45].

To calculate the mass and decay constant for the true Goldstone meson, it turns out it is sufficient to find the self energy $\Sigma(p^2)$ [4, 5]. This is not very surprising for the mass computation; the mass of P_5^+ is provided by the self-energy shift in the pole of the propagator.

$$\left(m_{P_5^+}^{1-\text{loop}}\right)^2 = m_{P_5^+}^2 + \Sigma(-m_{P_5^+}^2) \tag{3.43}$$

For the decay constant, it is not immediately obvious that only the self energy needs to be calculated. The pseudoscalar decay constant $f_{P_5^+}$ is defined in terms of the expectation value of the axial current $j_{\mu 5}^{P_5^+}$.

$$\langle 0 | j_{\mu 5}^{P_5^+} | P_5^+(p) \rangle = -i f_{P_5^+} p_\mu, \qquad j_{\mu 5}^{P_5^+} = \frac{-if^2}{8} \mathrm{Tr}\left[\xi_5 \mathcal{P}^{P+} \left(\partial_\mu \Sigma \Sigma^\dagger + \Sigma^\dagger \partial_\mu \Sigma \right) \right] \tag{3.44}$$

Here, $\mathcal{P}^{P+}$ is a projection operator that selects the flavor content x and $\bar{y}$. When the expectation value is evaluated at one loop, corrections enter from both wavefunction renormalization and the expansion the axial current to $O(\Phi^3)$. It turns out that the axial current expansion term is proportional to the wavefunction renormalization term, with a proportionality constant of -4 [5]. Overall, this results in a correction of

$$f_{P_5^+} = f \left[1 - \frac{3}{2} \frac{d\Sigma(p^2)}{dp^2} \right]_{p^2 = -m_{P_5^+}^2} \tag{3.45}$$

To understand the way valence and sea quarks appear in the contributions to the self energy, it is helpful to characterize the possible vertices and tadpole diagrams by their quark flow content. This is done for diagrams with two external P^+ mesons in Figures 3.2 and 3.3, originally published by C. Aubin and C. Bernard [4].

The vertices that need to be considered come from the LO Lagrangian expanded to $O(\Phi^4)$ and include the kinetic energy term containing two derivatives, the mass term with one mass matrix M, and the two forms of the staggered potential U and U' with single or double trace terms, respectively. The single trace terms (kinetic energy, mass, and U) must share flavor indices with adjacent fields Φ in the trace and this pattern must cycle around the entire product in the trace. Therefore, only connected vertices as in Fig. 3.2(a) and (b) may contribute. Furthermore, the valence quarks entering the loop from vertex (b) must be different and therefore require a disconnected propagator. The only tadpole diagrams satisfying these criterion are Fig. 3.3(a), (g), and (h) for vertex (a) and Fig. 3.3(i) and (j) for vertex (b).

After addressing the connected vertices, the only remaining Lagrangian terms are in U'. Each term in U' is composed of two traces and must therefore produce disconnected vertices like Fig. 3.2(c) and (d). Luckily, the explicit taste structure of any term in U' is either vector or axial, which ensures the number of meson fields in each trace is odd and only vertex (d) is possible. The only compatible tadpole diagrams with vertex (d) are Fig. 3.3(e) and (f). This implies tadpole diagrams Fig. 3.3 (b), (c) and (d) never appear in this calculation.

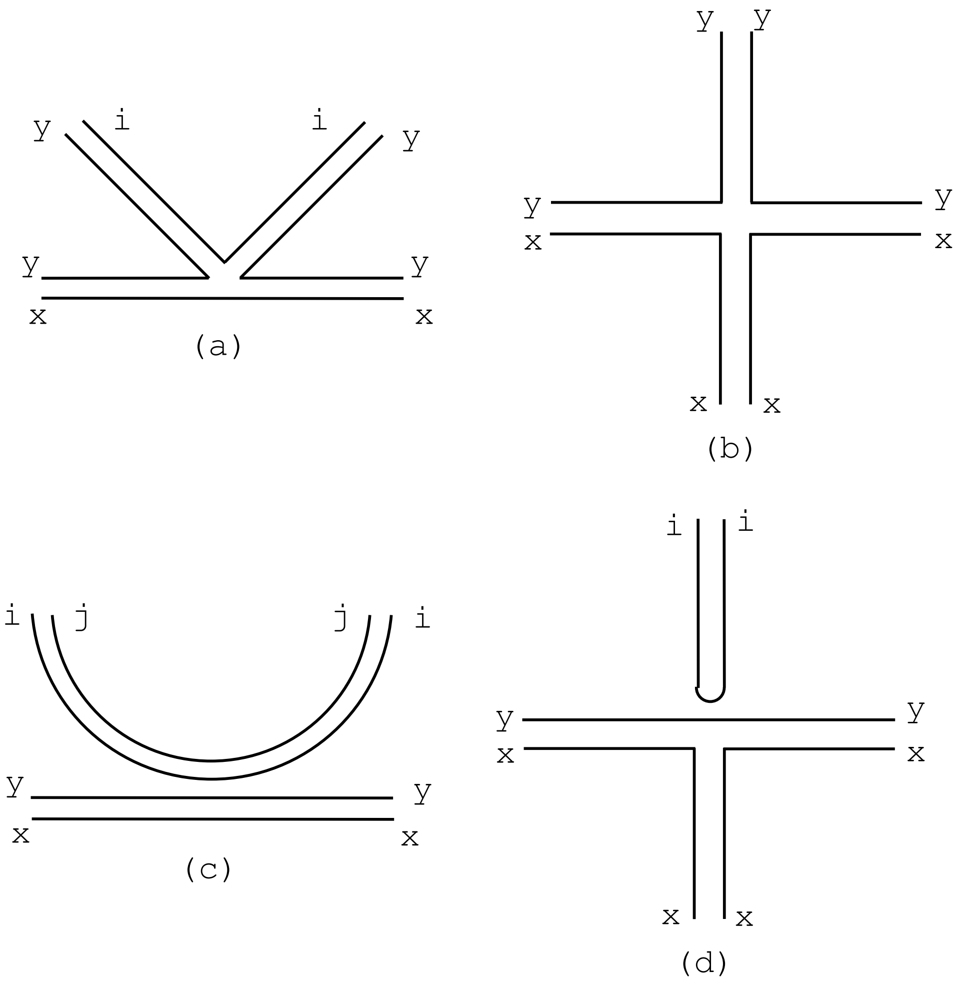

Figure 3.2: Quark flow diagrams for each of the possible flavor contractions of four mesons, where two mesons are the pseudoscalar boson composed of x and $\bar{y}$ quarks. The top two (a) and (b) represent vertices with connected quark flow, while the bottom two (c) and (d) represent disconnected quark flow diagrams. In the chiral and full QCD theory, these are all connected vertices. Note: there is another form of vertex (a) and (d) not shown here, where the two valence quarks x and $\bar{y}$ swap roles.

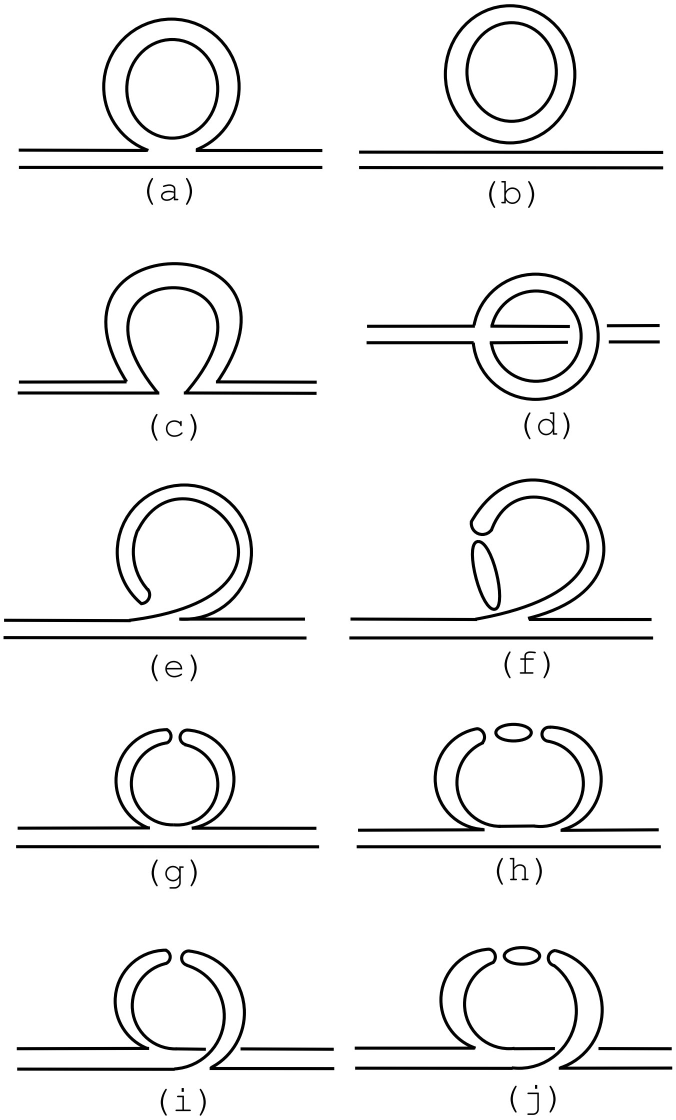

Figure 3.3: Potential quark flow contributions to the self-energy tadpole diagram. The corresponding vertices for each diagram are shown in Figure 3.2. Fig. 3.2(a) corresponds to tadpoles (a), (g), and (h); Fig. 3.2(b) corresponds to tadpoles (i), and (j); Fig. 3.2(c) corresponds to only (b) and (c); and Fig. 3.2(d) corresponds to (e) and (f). Diagrams (f), (h), and (j) represent *full* disconnected propagators in the loops. Diagrams (g) and (i) use a single iteration of the disconnected two-point vertex and have no internal quark loops.

Based on the list of tadpole diagrams, the self energy can be expressed as a simple sum of the contributions for each allowed case listed above.

$$\Sigma(p^2) = \frac{1}{16\pi^2 f^2}\left[(\sigma_a + \sigma_g + \sigma_h)_a + (\sigma_i + \sigma_j)_b + (\sigma_e + \sigma_f)_d\right] \tag{3.46}$$

The subscripts on $\sigma(p^2)$ in Eq. (3.46) denote the contributing tadpole diagram from Fig. 3.3 and the subscript on the groups denotes the vertex in Fig. 3.2.

In each contribution, the vertices tend to contribute in similar ways. For connected vertices, the kinetic energy term contributes $p^2 + q^2$, where the derivative operators either act on both the external (momentum p) or internal (momentum q) meson lines. Similarly, the mass matrix contributes only for connected vertices and yields one continuum, tree-level expression mass term for the external lines, $m^2_{P_5^+} = \mu(m_x + m_y)$, and another for the internal lines. The last of the connected vertex contributions, the single trace potential, contributes the taste splittings $a^2\Delta(\xi_t)$ based on the taste of the internal meson ($\Delta(\xi_5) = 0$ for the P_5^+). The last possible vertices are the disconnected terms of U', each of which yields a factor of $a^2\delta'_t$. [7]

The value of each diagram is then determined by the symmetry factors and possible propagators in each loop. For tadpole (a), each of the flavors Q and tastes B of mixed sea-valence mesons (e.g. $x\bar{u}$) contributes to the final inner loop.

$$\sigma_a(p^2) = -\frac{1}{12}\sum_{Q,B}\int \frac{d^4q}{\pi^2}\left[p^2 + q^2 + \left(m^2_{P_5^+} + m^2_{Q_5}\right) + a^2\Delta(\xi_B)\right]\frac{1}{q^2 + m^2_{Q_B}}. \tag{3.47}$$

---

[7] The disconnected vertex associated with m_0^2 for taste ξ_I does not contribute as it only connects two fields together.

In comparison, not every taste contributes non-zero results to the tadpole diagrams for (g) and (h). Combining the two diagrams together, the results for vector-tastes $V = B$ and axial-tastes $A = B$ are non-zero and identical up to the taste index.

$$\begin{aligned}\left(\sigma_g(p^2) + \sigma_h(p^2)\right)_B &= -\frac{1}{3}\int \frac{d^4q}{\pi^2}\left[\left(p^2 + q^2 + m^2_{P_5^+} + m^2_{X_5} + a^2\Delta(\xi_B)\right)(\mathcal{D}_B)_{XX}\right. \\ &\qquad \left. + \left(p^2 + q^2 + m^2_{P_5^+} + m^2_{Y_5} + a^2\Delta(\xi_B)\right)(\mathcal{D}_B)_{YY}\right] . \qquad (3.48)\end{aligned}$$

Here, $X = \bar{x}x$ and $Y = \bar{y}y$ refer to the flavor-neutral, valence mesons and $\mathcal{D}$ is the potentially mixing flavor-neutral propagator defined in Eq. (3.33) The only other taste structure with a non-vanishing contribution is the singlet. The only difference from the vector/axial case is the overall factor.

$$\begin{aligned}\left(\sigma_g(p^2) + \sigma_h(p^2)\right)_I &= -\frac{1}{12}\int \frac{d^4q}{\pi^2}\left[\left(p^2 + q^2 + m^2_{P_5^+} + m^2_{X_5} + a^2\Delta(\xi_B)\right)(\mathcal{D}_I)_{XX}\right. \\ &\qquad \left. + \left(p^2 + q^2 + m^2_{P_5^+} + m^2_{Y_5} + a^2\Delta(\xi_B)\right)(\mathcal{D}_I)_{YY}\right] \qquad (3.49)\end{aligned}$$

The total contribution from diagrams (g) and (h) is the total over the three non-vanishing tastes: $\sigma_g + \sigma_h = \sum_{B=\mu,\mu5,I}(\sigma_g + \sigma_h)_B$.

Just like diagrams (g) and (h), (i) and (j) only contribute through the vector, axial, and singlet tastes and use a disconnected propagator in the loop. The primary difference in this case is the change in flavor (X to Y or vice versa) over the disconnected propagator. Note, there is no mass contribution m_X^2, m_Y^2 or a related term in the numerator because the mass matrix that previously provided such terms does not have off-diagonal entries containing quark masses (see Eq. (3.35)).

$$\begin{aligned}
\left(\sigma_i(p^2) + \sigma_j(p^2)\right)_B &= -\frac{2}{3}\int \frac{d^4q}{\pi^2}\left[\left(p^2 + q^2 - m_{P_5^+}^2 + a^2\Delta(\xi_B)\right)(\mathcal{D}_B)_{XY}\right] && (3.50)\\
\left(\sigma_i(p^2) + \sigma_j(p^2)\right)_I &= \frac{1}{6}\int \frac{d^4q}{\pi^2}\left[\left(p^2 + q^2 - m_{P_5^+}^2 + a^2\Delta(\xi_I)\right)(\mathcal{D}_I)_{XY}\right] && (3.51)
\end{aligned}$$

Also, in the case of (i) and (j), the singlet taste contribution has the opposite sign of the vector and axial taste contributions. Basically, the sign changes for diagrams with vertex (b) but does not change for diagrams with vertex (a) because the order of the taste matrices arising from external and internal fields in the kinetic energy and mass traces depends on how the flavor indices are contracted at the vertex. For vertex (b), each internal field must appear between the external fields at the vertex. Therefore, to cancel the taste matrices the internal taste matrix (ξ_I, ξ_ν, or $\xi_{\nu 5}$) has to be commuted with the external taste matrix (ξ_5). This results in an overall sign difference between the singlet and vector/axial cases. For vertex (a), the internal fields start out adjacent to each other in the vertex. Therefore, no commutation is required and the overall sign is independent of the internal field's taste.

The final contributions are from the disconnected vertices in U'. The vertex gives an explicit factor of $a^2\delta'_B$ for each of the contributing tastes (vector and axial). Furthermore, because tadpole diagram (e) does not contain any vertices in the internal propagator, it only contributes the simple propagator, in contrast to the full disconnected propagators of diagram (f).

$$\sigma_e = \sum_{B=\mu,\mu 5} -\frac{2}{3}a^2\delta'_B \int \frac{d^4q}{\pi^2}\left[\frac{1}{q^2+m^2_{X_B}} + \frac{1}{q^2+m^2_{Y_B}}\right] \tag{3.52}$$

$$\sigma_f = \sum_{B=\mu,\mu 5} -\frac{2}{3}a^2\delta'_B \int \frac{d^4q}{\pi^2} \sum_{M=U,D,S} ((\mathcal{D}_B)_{XM} + (\mathcal{D}_B)_{YM}) \tag{3.53}$$

Upon review of Eqs. (3.43) through (3.53), we see that the integrals that yield non-vanishing, chiral logs in the self energy and its derivative (evaluated at $p^2 = -m^2_{P_5^+}$) are all of the same basic rational forms. In the self energy, all of the integrands are of the form

$$\mathcal{I}^{[n,k]}\left(\{m\};\{\mu\}\right) \equiv \frac{\prod_{a=1}^{k}(q^2+\mu_a^2)}{\prod_{j=1}^{n}(q^2+m_j^2)} , \tag{3.54}$$

where $\{m\}$ and $\{\mu\}$ denote the set of masses iterated over in the denominator and numerator, respectively, and n and k are the sets' cardinalities. Note, the masses appearing in $\{m\}$ are not necessarily unique; however, for the cases considered here every integrand contains at most one double pole, and one or more single poles. In the case where all poles are single, the integrand can be re-expressed in terms of the residues at each pole

$$\mathcal{I}^{[n,k]}\left(\{m\};\{\mu\}\right) = \sum_{j=1}^{n} \frac{R_j^{[n,k]}\left(\{m\};\{\mu\}\right)}{q^2+m_j^2} , \tag{3.55}$$

where

$$R_j^{[n,k]}\left(\{m\};\{\mu\}\right) \equiv \frac{\prod_{a=1}^{k}(\mu_a^2 - m_j^2)}{\prod_{i\neq j}(m_i^2 - m_j^2)} \ . \tag{3.56}$$

If one double pole at $q^2 = -m_l^2$ is present, a similar procedure may be followed. To apply the residue expansion, the integrand is re-expressed as the derivative of an exclusively simple-pole integrand

$$\begin{aligned} \mathcal{I}_{\mathrm{dp}}^{[n,k]}\left(m_\ell;\{m\};\{\mu\}\right) &= -\frac{d}{dm_\ell^2}\left(\frac{\prod_{a=1}^{k}(q^2+\mu_a^2)}{\prod_{j=1}^{n}(q^2+m_j^2)}\right) && (3.57)\\ &= \frac{R_\ell^{[n,k]}\left(\{m\};\{\mu\}\right)}{(q^2+m_\ell^2)^2} + \sum_{j=1}^{n}\frac{D_{j,\ell}^{[n,k]}\left(\{m\};\{\mu\}\right)}{(q^2+m_j^2)} \ , && (3.58) \end{aligned}$$

with

$$D_{j,\ell}^{[n,k]}\left(\{m\};\{\mu\}\right) \equiv -\frac{d}{dm_\ell^2}R_j^{[n,k]}\left(\{m\};\{\mu\}\right) \ . \tag{3.59}$$

With this substitution, the logarithmic portion of the result for each loop integral is a coefficient in terms of $R_j^{[n,k]}$ and $D_{j,\ell}^{[n,k]}$ times one of two chiral logs.

$$\begin{aligned} \mathcal{I}_1 &\equiv \int\frac{d^4q}{(2\pi)^4}\frac{1}{q^2+m^2} \to \frac{1}{16\pi^2}\ell(m^2) \equiv m^2\ln\frac{m^2}{\Lambda^2} && (3.60)\\ \mathcal{I}_2 &\equiv \int\frac{d^4q}{(2\pi)^4}\frac{1}{(q^2+m^2)^2} = -\frac{\partial}{\partial m^2}\mathcal{I}_1 \to \frac{1}{16\pi^2}\tilde{\ell}(m^2) \equiv -\left(\ln\frac{m^2}{\Lambda^2}+1\right) && (3.61) \end{aligned}$$

where Λ is the chiral scale. Note that both the divergent and analytic terms are dropped from these results. The divergent terms will be absorbed into the renormalized LECs and the analytic terms will be part of the more general analytic expression appended to the final NLO result.

The above forms of the chiral logs $\ell(m^2)$ and $\tilde{\ell}(m^2)$ apply in infinite volume. For the large but finite volumes typical in computer simulations, modifications are added. The modifications arise from replacing the integral over the Fourier momenta in the three spatial dimensions with corresponding discrete Fourier sums, which result in the Bessel functions K_0 and K_1 of imaginary argument. For more details [4] provides a brief overview of the derivation, and further details can be found in [19]. The final results are simply quoted here for completeness.

$$\ell(m^2) \equiv m^2\left(\ln\frac{m^2}{\Lambda^2}+\delta_1(mL)\right), \tag{3.62}$$

$$\tilde{\ell}(m^2) \equiv -\left(\ln\frac{m^2}{\Lambda^2}+1\right)+\delta_3(mL), \tag{3.63}$$

where

$$\delta_1(mL) = \frac{4}{mL}\sum_{\vec{r}\neq 0}\frac{K_1(|\vec{r}|mL)}{|\vec{r}|}, \tag{3.64}$$

$$\delta_3(mL) = 2\sum_{\vec{r}\neq 0}K_0(|\vec{r}|mL), \tag{3.65}$$

and $\vec{r}$ is a three dimensional vector with integer components. After substituting the above results into Eqs. (3.43) and (3.45), the NLO expressions need to be completed by including the analytic terms at order $O(x^2)$ where $x \sim a^2 \sim m_q$. This involves a combination of continuum χPT LECs at NLO (the various L_i's defined in [71]) and new terms from rSχPT corresponding to the next order of the SET. These new operators will all yield terms that are $O(m_q a^2)$, because a^2 has to be included to form operators distinct from the continuum, and $O(a^4)$ operators cannot contribute to these results: the Goldstone mass must vanish in the chiral limit, and any operator contributing to the decay constant must contain two derivatives. Because the leading order meson mass and decay constant come from terms of $O(m_q)$ in the Lagrangian, this allows the effects of the new operators to be combined into single terms of the form a^2C for the mass and a^2F for the decay constant.

Since the analysis of the next section is based on $N_f = 2+1+1$ HISQ ensembles, the mass and decay constant formulas at NLO in the $N_f = 2+1$ rSχPT are quoted here. For the full $N_f = 1+1+1$ rSχPT expressions and various other useful limits, see [4, 5].

The Goldstone mass is

$$\begin{aligned}
\frac{(m_{P_5^+}^{1-\text{loop}})^2}{(m_x+m_y)} \quad = \quad & \mu\Bigg\{1 + \frac{1}{16\pi^2 f^2}\Big(-2a^2\delta'_V \sum_{j_V} R^{[4,2]}_{j_V}\ \ell(m^2_{j_V}) - \Big(V \to A\Big) \\
& +\frac{2}{3}\sum_{j_I} R^{[3,2]}_{j_I}\ \ell(m^2_{j_I})\Big) + \frac{16\mu}{f^2}\left(2L_8 - L_5\right)(m_x+m_y) \\
& +\frac{32\mu}{f^2}\left(2L_6 - L_4\right)(2m_l + m_s) + a^2C\Bigg\},
\end{aligned} \tag{3.66}$$

where the the residue terms for tastes V (or A) and I are shorthand for

$$R^{[4,2]}_{j_V} = R^{[4,2]}_{j_V}\left(\{m_{X_V}, m_{Y_V}, m_{\eta_V}, m_{\eta'_V}\}; \{m_{U_V}, m_{S_V}\}\right), \tag{3.67}$$

$$R^{[3,2]}_{j_I} = R^{[3,2]}_{j_I}\left(\{m_{X_I}, m_{Y_I}, m_{\eta_I}\}; \{m_{U_I}, m_{S_I}\}\right) \ . \tag{3.68}$$

The result for the decay constant is significantly more complicated due to the presence of double poles.

$$\begin{aligned}
f_{P_5^+}^{1-\text{loop}} \quad &= f\Bigg\{1 + \frac{1}{16\pi^2 f^2}\Bigg[-\frac{1}{32}\sum_{Q,B}\ell\left(m_{Q_B}^2\right) + \frac{1}{6}\Bigg(R^{[2,2]}_{X_I}(\{\mathcal{M}^{(5)}_{X_I}\})\tilde{\ell}(m^2_{X_I}) \\
&+R^{[2,2]}_{Y_I}(\{\mathcal{M}^{(5)}_{Y_I}\})\tilde{\ell}(m^2_{Y_I}) + \sum_{j_I} D^{[2,2]}_{j_I,X_I}(\{\mathcal{M}^{(5)}_{X_I}\})\ell(m^2_{j_I}) + \sum_{j_I} D^{[2,2]}_{j_I,Y_I}(\{\mathcal{M}^{(5)}_{Y_I}\})\ell(m^2_{j_I}) \\
&-2\sum_{j_I} R^{[3,2]}_{j_I}(\{\mathcal{M}^{(6)}_{I}\})\ell(m^2_{j_I})\Bigg) + \frac{1}{2}a^2\delta'_V\Bigg(R^{[3,2]}_{X_V}(\{\mathcal{M}^{(7)}_{X_V}\})\tilde{\ell}(m^2_{X_V}) \\
&+R^{[3,2]}_{Y_V}(\{\mathcal{M}^{(7)}_{Y_V}\})\tilde{\ell}(m^2_{Y_V}) + \sum_{j_V} D^{[3,2]}_{j_V,X_V}(\{\mathcal{M}^{(7)}_{X_V}\})\ell(m^2_{j_V}) \\
&+\sum_{j_V} D^{[3,2]}_{j_V,Y_V}(\{\mathcal{M}^{(7)}_{Y_V}\})\ell(m^2_{j_V}) + 2\sum_{j_V} R^{[4,2]}_{j_V}(\{\mathcal{M}^{(8)}_{V}\})\ell(m^2_{j_V})\Bigg) + \Big(V \to A\Big)\Bigg] \\
&+\frac{16\mu}{f^2}\left(2m_\ell + m_s\right)L_4 + \frac{8\mu}{f^2}\left(m_x + m_y\right)L_5 + a^2 F\Bigg\}\ , \qquad (3.69)
\end{aligned}$$

The numerator masses for all of the residue terms are implicitly set to

$$\{\mu\} = \{m_U, m_S\}\ , \qquad (3.70)$$

and the explicit mass sets in the denominators are

$$\begin{aligned}
\{\mathcal{M}_Z^{(5)}\} &\equiv \{m_\eta, m_Z\}\ , \\
\{\mathcal{M}^{(6)}\} &\equiv \{m_\eta, m_X, m_Y\}\ , \\
\{\mathcal{M}_Z^{(7)}\} &\equiv \{m_\eta, m_{\eta'}, m_Z\}\ , \\
\{\mathcal{M}^{(8)}\} &\equiv \{m_\eta, m_{\eta'}, m_X, m_Y\}\ ,
\end{aligned} \tag{3.71}$$

3.3 Computational Analysis

The previous section covered the foundations of staggered chiral perturbation theory (SχPT) and how it can provide a model of the combined lattice spacing and quark mass dependence of the pseudoscalar meson masses and decay constants. This section summarizes the procedure of extracting physical observables from lattice computations using the model from SχPT as a base. There are three parts to the presentation. First, the complete model and inputs to the chiral analysis are summarized. Then, the procedures for optimizing fits of the model to lattice data are outlined. Finally, the optimized fit function is compared to the ensemble results in order to provide a deeper context for the previous two parts, as well as to set the stage for the continuum, physical quark-mass results in the next section.

3.3.1 Fit Structure

For sufficiently light masses and small lattice spacings, a next to leading order expansion in a^2 and m_q may be sufficient; however, that is not the case in these simulations. To study the kaon, χPT must be extended up to the mass of the strange quark. While the strange quark is still significantly lighter than the charm quark, it is not light enough to neglect higher order corrections in m_q. Also, the largest lattice spacing included in the MILC HISQ ensembles is approximately $a \approx 0.15$fm. In order to include the results from the coarsest lattice, higher order effects in a^2 or $\alpha_s a^2$ should be considered. This is particularly relevant when fitting results based on improved actions/observables, such as the HISQ action, which diminish lower order discretization effects.

In this work, the next to leading order results of rSχPT are extended up to fourth order corrections. Specifically, at a fixed lattice spacing a and finite spatial volume $V = L^3$, the full model g can be summarized as

$$g = g_0 \left[\left(1 + g_0^{(1)} + g_1^{(1)} + g_0^{(2)} + g_1^{(2)} + g_2^{(2)} \right)_{\text{sys}} + \left(g_0^{(3)} + g_0^{(4)} \right)_{\text{tree}} \right]. \tag{3.72}$$

The superscript n on $g^{(n)}$ denotes the overall order of the correction in combined powers of m_q and a^2, where the usual power counting $a^2 \sim m_q$ is employed. The subscript denotes the number of internal loops included in contributing diagrams The various orders are grouped into two sets 'sys' and 'tree' depending on whether loops are included. In other words, the first set is a fully systematic χPT expansion while the latter set is only comprised of analytic terms based on the chiral symmetries. For each order and loop content, each term is further elaborated below.

The functional form g is used to fit the two observables: $m^2_{P_5^+}/(m_x + m_y)$ and $f_{P_5^+}$. Both quantities must be fit simultaneously because they share some of the LECs of χPT as parameters. Depending on which observable is being predicted, g_0 is either f or μ, the associated LO LEC for χPT.

The next term $g_0^{(1)}$ is the collection of all analytic terms at NLO in rSχPT, summarized at the ends of Eqs. (3.66) and (3.69). The other NLO term $g_1^{(1)}$ collects all the 1-loop chiral logs contained in the rest of those equations. Note, the finite volume dependence of this model only enters through the chiral logs at NLO; higher order corrections for finite volume are not needed here. The remaining three systematic terms at NNLO are based on χPT with continuum quarks (i.e. not SχPT). The first term $g_0^{(2)}$ is a combination of new analytic terms parameterized by two sets of five new LECs, Q_1 through Q_5 for $m^2_{P_5^+}$ and Q_6 through Q_{10} for $f_{P_5^+}$. The second term $g_1^{(2)}$ is a collection of new one loop diagrams, partially based on the analytic NLO terms in $g_0^{(1)}$. The final systematic term $g_2^{(2)}$ is a collection of partially quenched two-loop diagrams first calculated by [27, 29, 28]. The two loop diagrams are complicated expressions and computationally expensive to compute compared to the other terms in g. We wish to thank J. Bijnens for providing the Fortran library used to compute all of the NNLO loop contributions in this work.

The final two terms are not based on a fully systematic χPT expansion; instead they are only comprised of analytic terms in the valence and sea quark masses. While these analytic terms are not explicitly derived from NNNLO χPT , they do obey all the chiral symmetries. This ensures that $g_0^{(3)}$ and $g_0^{(4)}$ are equivalent to the collection of χPT analytic terms that would be derived at each order.

The functional form described in Eq. (3.72) is sufficient for studying any one lattice spacing. Given a value of a, the leading order corrections for each taste of meson are added to the masses in loops; the disconnected vertices $a^2\delta'_t$ are defined; the disconnected propagators are re-expressed in terms of the full mass matrix basis π_0, η, and η'; and the a^2 analytic terms are added to the full NLO expansion. For the HISQ action, these taste artifacts are expected to be order $\alpha_s^2 a^2$, which implies that the disconnected vertices become $\alpha_s^2 a^2 \delta_t'$ for $t = V, A$ and similarly for the analytic terms.

However, one subtle point is that, to this point, we have only included discretization effects that break the continuum symmetries of the theory. This means that any continuum LECs determined by a simulation at finite lattice spacing will possess discretization corrections to their associated continuum values. These discretization corrections are not taste-artifacts, so they will follow the leading *generic* discretization corrections. For this HISQ action this is expected to be $O(\alpha_s a^2)$. Next order corrections could be either a^4, $\alpha_s a^4$, or $(\alpha_s a^2)^2$. For the leading order term g_0, corrections are added up to second order. For all other terms in $g^{(1)}$ through $g^{(3)}$ only first order discretization corrections are added. Based on this, g_0 is replaced by

$$g_0 \to g_0 \left[1 + dg_1 \alpha_s a^2 + dg_2 a^4 + dg_3 \alpha_s a^4 + dg_4 (\alpha_s a^2)^2\right] \tag{3.73}$$

where dg_i are new parameter sets determined independently for each LEC (e.g. $g_0 = \mu$ and $g_0 = f$). Any combination of second order corrections may be considered in individual fits; in this preliminary work we only consider fits with $dg_3 = 0$. Including the the additional form $\alpha_s a^4$ in the fits is not expected to cause any significant changes; this will be tried in the future.

Fit Inputs

The functional form in Eq. (3.72) is dependent on several variables. The complete set of dependencies includes the lattice spacing a, the strong coupling constant α_s, the physical

β	$\approx a$(fm)	$a/w_{0,\mathrm{imp}}$	α_s	am_{p4s}
5.80	0.15	0.8823(3)	0.58801	0.02746(6)
6.00	0.12	0.7040(5)	0.53727	0.02122(5)
6.30	0.09	0.5107(3)	0.42387	0.01452(4)
6.72	0.06	0.3306(2)	0.29134	0.00873(2)

Table 3.1: The values of inputs to the χPT functional form in Eq. (3.72) that are constant for fixed coupling β, excluding taste splittings. From left to right the columns list the gauge coupling constant β, the approximate lattice spacing a, the relative lattice spacing $a/w_{0,\mathrm{imp}}$ set through gradient flow (see chapter 2), the strong coupling constants α_s, and $am_{\mathrm{p4s}} = 0.4am_s$ where m_s is the strange quark mass in lattice units extrapolated to physical values of the dynamical quark masses [13]. For both $a/w_{0,\mathrm{imp}}$ and am_{p4s} the quoted errors are a combination of statistical and systematic errors from retuning to physical quark masses.

β	$\approx a$(fm)	$\mu 5$	$\mu\nu$	μ	I
5.80	0.15	3.119934e-02	6.005524e-02	8.861347e-02	1.176413e-01
6.00	0.12	1.601204e-02	3.064779e-02	4.970139e-02	6.207202e-02
6.30	0.09	5.417633e-03	1.039458e-02	1.570553e-02	1.977797e-02
6.72	0.06	1.108087e-03	2.153787e-03	2.946609e-03	3.858016e-03

Table 3.2: Taste splittings for each of the non-Goldstone tastes and values of the gauge coupling β [12]. Taste splittings are calculated on ensembles with $m_l' = m_s'/5$ at each β, with the two valence quark masses set equal to the light sea quark mass am_l'. After the gauge coupling β, listed in the first column, the remaining columns all show $w_{0,\mathrm{imp}}^2 \Delta(\xi_t) = w_{0,\mathrm{imp}}^2 (m_{P_t^+}^2 - m_{P_5^+}^2)$ where t is indicated by the column header.

value of the strange quark mass in lattice units am_s, the four non-zero taste splittings $a^2\Delta(\xi_t)$ for $t \in \{\mu, \mu\nu, \mu 5, I\}$, the spatial volume $V = L^3$, the collection of $N_f = 2+1+1$ sea quark masses, and the two valence quark masses m_x, m_y. For completeness, each variable is briefly summarized here with tabulated values wherever it is feasible. The values of the first seven variables (a, α_s, am_s, and all $a^2\Delta_t$) are held fixed across ensembles with the same coupling β. Their values are listed in Tables 3.1 and 3.2. The remaining variables are independently set on each gauge ensemble or, in the case of valence masses, each operator measured on each gauge ensemble.

In this chiral analysis, the relative lattice spacing is determined using the gradient flow analysis presented in Ch.2. Specifically, the tree level improved scale $w_{0,\text{imp}}/a$ is chosen for its reduced discretization errors. An additional benefit of using any of the gradient flow scales in this analysis is their minimal mass dependence; this means any unintended remaining mass dependence in the lattice spacings will be minimal compared to the expected mass dependence of $m^2_{P_5^+}$ and $f_{P_5^+}$.

Using the expected $\alpha_s^2 a^2$ dependence of the taste splittings, we may define the strong coupling constant up to a proportionality constant by measuring the average taste splitting

$$\alpha_s \propto \left(\frac{w_{0,\text{imp}}}{a}\right)\sqrt{\frac{1}{16}\left[a^2\Delta(\xi_5) + 4a^2\Delta(\xi_{\mu 5}) + 6a^2\Delta(\xi_{\mu\nu}) + 4a^2\Delta(\xi_\mu) + a^2\Delta(\xi_I)\right]} \qquad (3.74)$$

The overall constant is then set so that $\alpha_s = \alpha_V(q^* = 1.5/a)$ where α_V is the coupling in the asqtad heavy quark potential determined through the average plaquette in [40, 12].

The taste splittings themselves are calculated as the average difference of squares of the meson masses with the specified taste and the true Goldstone $m_{P_5^+}$

$$a^2\Delta(\xi_t) = \langle (am_{P_t^+})^2 - (am_{P_5^+})^2 \rangle. \qquad (3.75)$$

In these simulations, the taste splittings are computed for mesons with equal valence quarks on the $m_l' = m_s'/5$ approximately physical strange quark mass ensembles at each β.

The final variable held fixed for fixed β is the physical value of the strange quark mass in lattice units am_s. This variable is needed to scale the bare quark masses in each simulation into common units. In this analysis, a related quantity called $am_{\mathrm{p4s}} = 0.4am_s$ is used instead. In either case, this variable is given an initial value based on previous simulations [13] and then redetermined as an output of the chiral analysis. The new value of am_{p4s} can then be fed back into the fit to be determined again. This procedure is repeated until there is no statistical difference between what is put into the model and what is predicted by the model.

Out of the remaining three variables input into the functional form in Eq. (3.72), the spatial volume $V/a^3 = N_s^3$ in lattice units is already listed in Tables 1.1 and 1.2. The only quantities that need to be specified are the values of sea and valence quark masses for each ensemble and operator. These values are tabulated for the nearly-physical strange quark mass ensembles in Table 3.3. For all lighter-than-physical strange quark mass ensembles (tabulated in Table 1.2) the valence quark masses are the same as the physical strange quark mass ensemble with $a \approx 0.12$ fm and $m_l' = m_s'/10$. For all of the ensembles, both nearly-physical and lighter than physical strange quark mass, the valence mass generally runs from approximately the physical strange quark mass down either to a tenth of the physical strange quark mass or to the physical light quark mass (for ensembles with nearly-physical light sea quark masses). The values of $m^2_{P_5^+}$ and $f_{P_5^+}$ are computed for all distinct pairs of masses drawn from the available collection on each ensemble. This ensures both the pion and kaon values are accessible to interpolations (i.e. $m_x = m_y = m_l'$, and $m_x = m_l'$ with $m_y = m_s'$).

Fit Outputs

With the variable dependencies of the functional form in Eq. (3.72) summarized, the only remaining input to the fitting procedure is the lattice simulation data to be fit: the Goldstone

β	m'_l/m'_s	am'_l	am'_s	am'_c	m_v/m'_s
5.80	1/5	0.01300	0.0650	0.838	0.1, 0.15, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0
5.80	1/10	0.00640	0.0640	0.828	0.1, 0.15, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0
5.80	1/27	0.00235	0.0647	0.831	0.036, 0.07, 0.1, 0.15, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0
6.00	1/5	0.01020	0.0509	0.635	0.1, 0.15, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0
6.00	1/10	0.00507	0.0507	0.628	0.1, 0.15, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0
6.00	1/27	0.00184	0.0507	0.628	0.036, 0.073, 0.1, 0.15, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0
6.30	1/5	0.00740	0.0370	0.440	0.1, 0.15, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0
6.30	1/10	0.00363	0.0363	0.430	0.1, 0.15, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0
6.30	1/27	0.00120	0.0363	0.432	0.033, 0.066, 0.1, 0.15, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0
6.72	1/5	0.00480	0.0240	0.286	0.05, 0.1, 0.15, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0
6.72	1/10	0.00240	0.0240	0.286	0.05, 0.1, 0.15, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0
6.72	1/27	0.00080	0.0220	0.260	0.036, 0.068, 0.1, 0.15, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0

Table 3.3: The sea and valence quark masses used for computing light pseudoscalar masses and decay constants for ensembles with nearly-physical am'_s. All primed masses denote simulation values of the masses, as opposed to physical values. The ensembles are ordered by the gauge coupling β in the first column, followed by the value of the light sea quark mass relative to the strange quark mass am'_l/am'_s in the second column. For completeness, the next three columns list the three distinct flavors of sea quark masses in lattice units. The final column lists all available valence quark masses for each ensemble. The light pseudoscalar mass and decay constant are computed for every distinct pair of valence quarks m_x,m_y chosen from the values for m_v (with replacement).

masses $m^2_{P_5^+}$ and decay constants $f_{P_5^+}$. For each ensemble and available pair of valence masses, the meson mass is determined by extracting the lowest energy contribution to the zero momentum two-point correlator for P_5^+. Starting from the vacuum state $|0\rangle$, the quark bilinear $O^\dagger_{P_5^+}$ can be used to create a taste- and spin-pseudoscalar meson (the Goldstone meson), where

$$O_{P_5^+}(\vec{x},t) = \bar{q}(\gamma_5 \otimes \xi_5)q. \tag{3.76}$$

The two-point correlator is then defined as the average over space of the vacuum expectation value of creating and later annihilating a Goldstone meson. With the meson created at point x and annihilated a time t later, the zero momentum two -oint correlator is defined as

$$C_{P_5^+,P_5^+}(t) = \frac{1}{L^3}\sum_{\vec{y}} \langle O_{P_5^+}(\vec{y},t) O^\dagger_{P_5^+}(\vec{x},0)\rangle \tag{3.77}$$

The mass can be related to the correlator by inserting a complete set of energy states and using the usual time dependence of the bilinear operator in the Heisenberg picture,

$$C_{P_5^+,P_5^+}(t) = \sum_n c_n(\vec{x}) e^{-(E_n-E_0)t} \sim c_1(\vec{x}) e^{-m_{P_5^+}t} \tag{3.78}$$

where the coefficient $c_1(\vec{x})$ is

$$c_1(\vec{x}) = \frac{1}{L^3}\sum_{\vec{y}} \langle 0|O_{P_5^+}(\vec{y},0)|1\rangle\langle 1|O^\dagger_{P_5^+}(\vec{x},0)|0\rangle \tag{3.79}$$

, where $|1\rangle$ is the first excited state (here, a state of a single Goldstone meson at rest). The mass is finally extracted by fitting to the asymptotic behavior of $C_{P_5^+,P_5^+}(t)$ for $t \to \infty$. The decay constant can also be related to the two-point correlator through the 1 particle creation/annihilation amplitudes contained in the coefficient $c_1(\vec{x})$. [6, 1, 50].

$$f_{P_5^+} = (m_x + m_y)\sqrt{\frac{L^3 c_0(\vec{x})}{4m_{P_5^+}}} \tag{3.80}$$

In practice the point source at x does not provide a good signal for the ground state in Eq. (3.78). In this work the point source is replaced with both random-wall and Coulomb-wall sources to help alleviate the low signal to noise ratio. For more details see [13].

3.3.2 Fitting Procedures

Given all the required inputs and outputs for evaluating Eq. (3.72) on the HISQ gauge ensembles generated by MILC, the next step is to find the collection of parameters at each order that minimizes χ^2 between the fit's predictions and the collection of simulated mass and decay constant data. This fitting procedure is non trivial due to the large number of data points and a parameter space with high dimensionality. Specifically, the full data set includes 1792 measurements of masses/decay constants and the full N^{4}LO fit function has 132 parameters. Reaching an optimal parameter set within a reasonable amount of time is further complicated by the computational complexity of evaluating the two-loop part of the NNLO fit function ($g_2^{(2)}$ in Eq. (3.72)).

To help reach the optimal parameter set, Gaussian priors are utilized to direct the search through parameter space to physically reasonable solutions. For example, all discretization effects are given priors corresponding to $1 - \sigma$ deviations from the continuum values. To help constrain the LECs of SχPT at NLO and χPT at NNLO, the fit is also done in two stages.

As a first stage, only the lighter than physical strange quark mass ensembles at $a \approx 0.12$fm are considered. Furthermore, an upper bound is placed on the quark masses such that all sea quark masses and the sum of the valence quark masses is not greater than 2.0 in m_{p4s} units. Since the resulting dataset is only composed of a single lattice spacing and sufficiently small quark masses, it allows all of the discretization corrections to LECs to be set to zero and for the higher N^{3}LO/N^{4}LO analytic terms to be dropped. This first stage is the only fit that is fully systematic in the sense that all χPT effects (chiral logs from loops as well as analytic terms) are included up through the given order (NNLO).

The second stage is to constrain the LEC's determined from the systematic fit, add the results at larger quark masses and other lattice spacings back into the dataset, and re-perform the fit with discretization corrections and NNNLO/NNNNLO analytic contributions included. In practice, the full dataset is still not utilized because the sum of valence masses can be too large for the analytic terms at NNNNLO to reasonably account for. So, an upperbound of 2.75 in m_{p4s} is implemented for the sum of valence quark masses. With this upper bound, physics of the kaon can still be explored. In this work, the second stage fit will be referred to as the full or non-systematic fit, depending on the context.

Once the second stage is complete, preliminary results can be calculated from the optimized parameter set and fit function evaluated in the continuum limit and at physical values of the quark masses. However, at this point there are several systematic issues that need to be accounted for to reach final results. Several of these issues require re-optimizing the fit with an altered set of constraints or an altered dataset.

For example, altering the dataset by reducing the *effective* number of datapoints is often necessary to converge on an optimal parameter set. As the number of valence quark masses considered grows, the statistical precision at each ensemble remains constant since the underlying set of configurations does not change. Because additional valence quark masses are generally helpful when performing chiral extrapolations, there is a tendency for the number of datapoints to grow without increasing statistical precision. This limited precision can introduce noise into the fit to χ^2. On a percentage basis, small eigenvalues of the covariance matrix will be affected the most, and may lead to fits with unphysically high or low χ^2 values. The solution to this problem is to effectively limit the number of datapoints fed into the fit routine. In this work two approaches are considered for limiting the amount of data fed to the optimized fit: spectral value decomposition (SVD) and manually removing datapoints (thinning).

At the first stage, spectral value decomposition (SVD) is used to eliminate the smallest eigenvalues from the inverse covariance matrix. Below a certain threshold, the small eigenvalues mostly contribute noise to the determination of χ^2 without contributing physically relevant constraints on the parameter space. By setting these eigenvalues to zero, the total value of χ^2 is reduced significantly more than the number of degrees of freedom, which makes it easier for the fit routine to reach a minimum. However, the precise value of the threshold could introduce systematic biases if set too high. Therefore, after reaching a converged fit, any SVD thresholds should be varied and the fit should be allowed to re-converge. An alternative to SVD is to manually thin the data set before constructing the covariance matrix. One advantage of this approach is the lack of an arbitrary threshold; instead, the thinning routine can be based on physical properties of the ensemble.

At the second stage, the same SVD threshold is kept in place, but several features of the dataset are changed. For instance, the fit is run both with and without the $a \approx 0.15$fm ensembles. From earlier studies with gradient flow it is apparent that highly improved actions can sometimes complicate the problem by revealing previously-hidden higher order contributions at small lattice spacings. Removing the coarsest ensembles helps check the size of these higher order contributions.

After the full fit is optimized, the systematic fit will need to be repeated. This is primarily due to the analytic taste-violating LECs and to the disconnected vertex factors $a^2\delta_t$, all of which depend on the lattice spacing. Since the systematic fit only includes a single lattice spacing, these parameters cannot be fully determined by the systematic fit and are therefore left without a prior in the full fit. However, if the full fit significantly modifies the optimal value of any of these factors, then the systematic fit needs to be rerun with them either fixed or constrained within priors from the full fit.

Lastly, the effects of am_{p4s} need to be included. Since am_{p4s} is used to scale the quark masses in the dataset, but is also an output of the optimized fit, the first versions of the optimized fit are likely to predict slightly different values of am_{p4s} than what is initially input. Therefore, to make the fit self-consistent, each fit that returns an inconsistent value of am_{p4s} needs to be re-optimized with a dataset containing quark masses scaled by the new value of am_{p4s}.

In summary, the full fitting procedure is a looping two-stage process. Each iteration of the fit goes through a systematic stage consisting of one lattice spacing but fully systematic NNLO χPT, then a secondary stage with all or most of the lattice spacings and higher order chiral and discretization effects parameterized in the fit functional form. Between iterations of the fits, systematic effects due to higher order discretization effects and lattice-spacing dependent LECs are explored. Finally, the iterations continue until the value of am_{p4s} is determined self-consistently.

For the preliminary results presented in this work, the first iteration has been carried out. While the various systematic effects cannot be explored, the first iteration does yield continuum, physical-mass results that are likely to change by only small amounts with the remaining iterations.

3.3.3 Central Fits

The systematic fit for the first iteration of am_{p4s} is a fit to the seven ensembles at $a \approx 0.12$ fm with $m_x + m_y, m_s' < 2.0$ in m_{p4s} units. The cutoff on the quark masses is determined by attempting fits with several cutoffs and varying orders of the χPT expansion. Specifically, the cutoff of 2.0 is the largest cutoff that does not require $O(\mathrm{N}^3\mathrm{LO})$ or higher terms and still converged with $p > 0.1$. In addition to the quark mass cutoff, an SVD eigenvalue cutoff of 10^{-3} is used to decrease the noise in determinations of χ^2. Without this noise reduction, the fluctuations in χ^2 near a minimum make it impossible for the gradient descent based optimizer to converge on a minimum.

The final systematic fit that serves as the basis for the LECs of the full fit has a $\chi^2/dof = 89.6/99$ and $p = 0.76$ with priors treated as additional degrees of freedom. If the priors are excluded, the fit has $\chi^2/dof = 85.4/94$ and $p = 0.74$. The only priors constraining the systematic fit are on the five LECs L_0 through L_3 and L_7, with the deviation from the prior less than 2σ in all cases. For L_0 through L_2, the deviations are less than 1σ.

From the systematic fit, several variations of a full fit to additional lattice spacings have been attempted. None of the fits reached a state where the norm of the logarithmic derivative of χ^2 with respect to the parameters is small enough to be considered fully converged. As a result, any predictions made with these fits underestimate statistical errors, due to the relatively steep curvature of χ^2 compared to a true minimum. Since such a fit is unreliable for error determination, only the current best fit is explored here.

The current best full fit, keeps the LO and most of the NLO LECs determined from the systematic fit fixed. Among the NLO LECs, the two neutral mixing vertices $a^2\delta'_\mu$, $a^2\delta'_{\mu 5}$ and the two analytic a^2 coefficient F, C are allowed to vary without priors. The NNLO LECs are also allowed to vary, but given Gaussian priors that match the central value and deviations predicted by the systematic fit. Both N^{3}LO and N^{4}LO analytic terms are then added and allowed to vary with no priors. For every LEC below N^{4}LO, discretization terms are also added as specified in Sec. 3.3.1. For the two LO discretization terms, both variations of higher order discretizations in Eq. (3.73) are included. No higher order charm quark mass corrections are included.

The dataset used to optimize the current best fit uses SVD with a cutoff of 10^{-3} (consistent with the systematic fit), and removes all points where $m_x + m_y < 2.75$ in m_{p4s} units. All three $a \approx$0.15 fm ensembles are removed from the analysis, as well as the $a \approx 0.12$ fm, nearly-physical quark mass ensemble. All four ensembles are removed from consideration because they are not essential for the continuum extrapolation and they contribute an abnormally large percentage of the total χ^2/dof compared to other ensembles. For the coarse $a \approx 0.15$ fm ensemble, the large χ^2 is likely due to neglecting higher order discretization corrections. For the $a \approx 0.12$ fm, nearly-physical quark mass ensemble the complication comes from being left out of the systematic fit at $a \approx 0.12$ fm. All of the physical strange-quark mass ensembles at $a \approx 0.12$ fm can only deviate from the systematic fit based on their mass content. This constraint is hardest to fulfill for the lighter quark masses, where higher order effects have relatively less impact.

The current best fit has a $\chi^2/dof = 570.3/366$ and $p = 3.9 \times 10^{-9}$ with priors treated as additional degrees of freedom. If the priors are excluded, the fit has $\chi^2/dof = 477.8/301$ and $p = 9.9 \times 10^{-9}$.

To gain some deeper insight into how the fit is performing, Figures 3.4, 3.5, 3.6, and 3.7 plot the pion-like and kaon-like Goldstone boson mass and decay constant as a function of the average valence quark mass. For all of the plots, only the nearly-physical strange quark mass ensembles that are included in the fit are shown. All of the lighter-than-physical strange-quark mass ensembles are also included in the fit, but do not appear on these plots.

Results for the pion-like Goldstone boson are shown in Figures 3.4 and 3.5. A Goldstone boson is considered pion-like if both valence quark masses are set equal, and the physical point (black star in each figure) is defined where both valence quark masses equal the light sea-quark mass. For each nearly-physical m_s' ensemble included in the plot, each point corresponds to a unique value of $m_x = m_y$, and the line is the interpolation of the central fit evaluated at sea quark masses, lattice spacing, and lattice volume identical to that ensemble. Finally, the continuum line is evaluated at zero lattice spacing, infinite volume, and the physical sea-quark masses determined in Sec. 3.4.1.

Results for the kaon-like Goldstone boson are shown in Figures 3.6 and 3.7. A Goldstone boson is considered kaon-like if one valence quark mass equals the sea strange-quark mass, and the physical point (black star in each figure) is defined where the free valence quark equals the physical up-quark mass determined in Sec. 3.4.1. Each ensemble's points, best

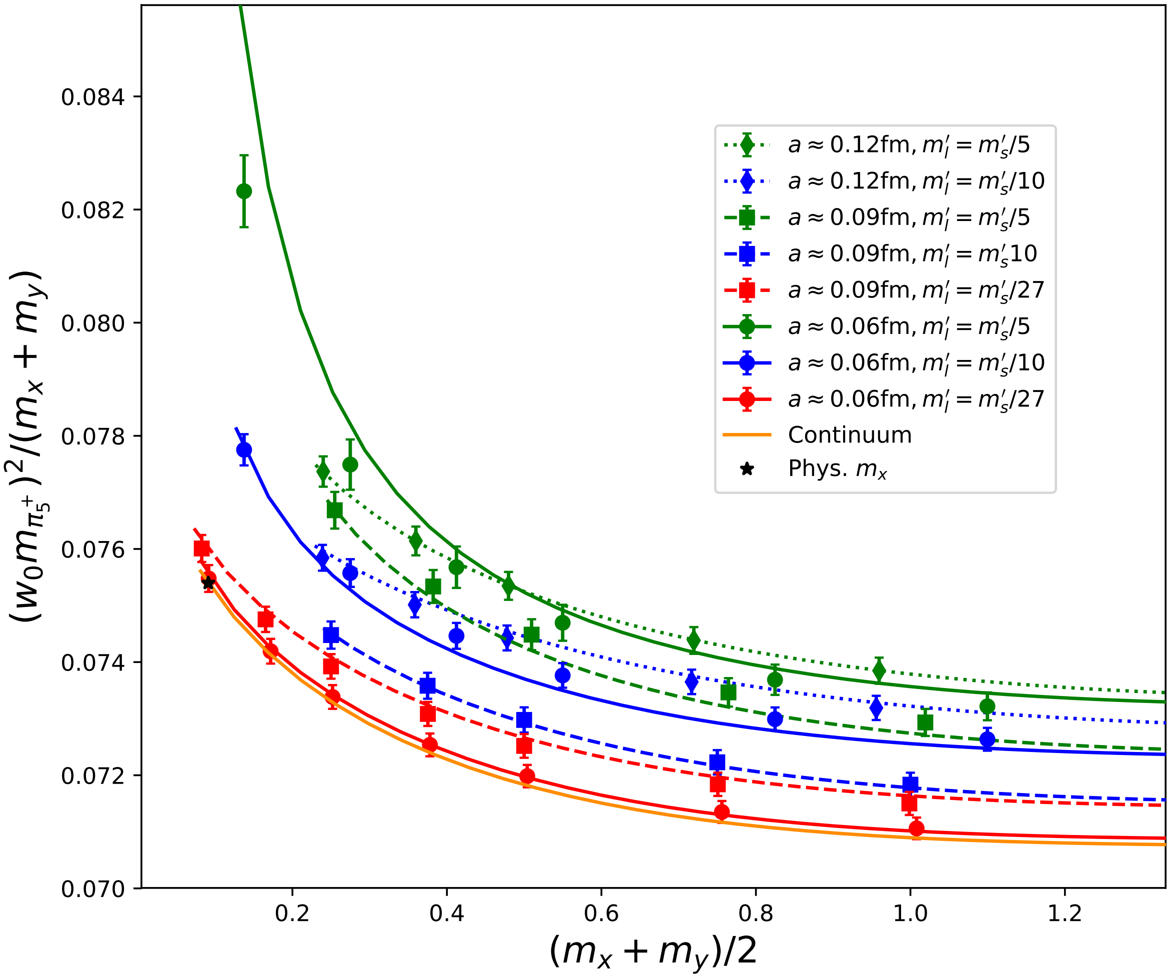

Figure 3.4: Plot of the Goldstone pion-like meson mass $m^2_{\pi_5^+}/(m_x + m_y)$ in $w_{0,\mathrm{imp}}$ units for $m_{\pi_5^+}$ and m_{p4s} units for the quark masses m_x and m_y, versus the average valence quark mass $(m_x + m_y)/2$ for all of the nearly-physical strange quark mass ensembles including in the central fit. All points are pion-like because both valence quarks are set equal to each other. Symbols denote different lattice spacings and the color denotes the ratio of the light and strange sea-quark masses, as per the legend. The fit evaluated at physical sea-quark masses and continuum is denoted by the dark orange "Continuum" line, and the point at physical valence quark masses is indicated by the black star.

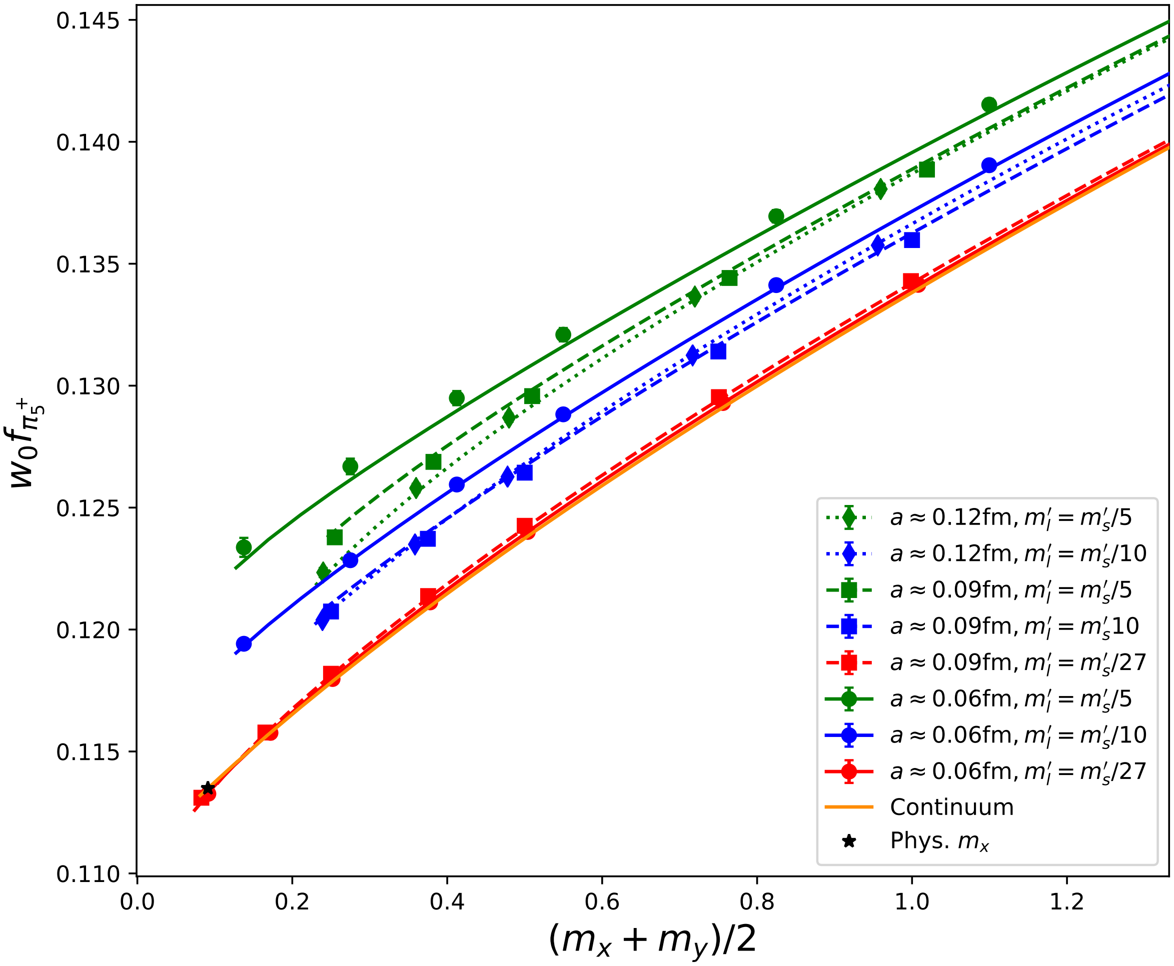

Figure 3.5: Plot of the Goldstone pion-like decay constant $f_{\pi_5^+}$ in $w_{0,\mathrm{imp}}$ units versus the average valence quark mass $(m_x + m_y)/2$ for all of the nearly-physical strange quark mass ensembles including in the central fit. The rest of the figure description is similar to Fig. 3.4.

fit line, and the continuum line are defined similarly to the pion-like figures. One obvious difference for the kaon-like figures is the number of datapoints. Due to the valence quark mass cutoff of $m_x + m_y < 2.75$ in m_{p4s} units, all of the free valence quark masses included in the plot have to be less than 0.45. This is because the lightest, nearly-physical simulation strange quark mass is approximately 2.3 in m_{p4s} units.

Between both the pion-like and kaon-like plots one trend is clear: the continuum line closely follows the two finest nearly-physical quark mass ensembles at $a \approx 0.09$ fm and especially at $a \approx 0.06$ fm. This implies that both the discretization corrections and quark mass mistunings of the $a \approx 0.06$ fm, nearly-physical quark mass ensemble are small, and it suggests that there will be very little variability in the central value of physical predictions. This may explain why the current preliminary results in Sec. 3.4 are in such good agreement with FLAG averages despite the poor confidence of the fit; what matters most is the fit matches data on the $a \approx 0.06$ fm, nearly-physical quark mass ensemble.

For points where the agreement with the fit is not ideal, there are a several features that could be examined. As an example, consider the largest deviation in Fig. 3.4, which occurs on the $a \approx 0.06$ fm, $m_l' = m_s'/5$ ensemble evaluated at the lightest valence quark mass. In this regime, the sea quark masses are at their heaviest while the valence quark masses are considerably lighter than even the light sea-quark mass. As a result, the partially-quenched pion receives significant finite volume corrections; the total finite volume correction is 8% of the total contributions to the pion mass. Considering that the NLO finite volume corrections are on the same order as the entire NLO and NNLO corrections, it's possible that finite volume effects at NNLO are not negligible for this ensemble as well. Finally, the nearby points at slightly heavier valence quark masses also deviate on the same side of the fit

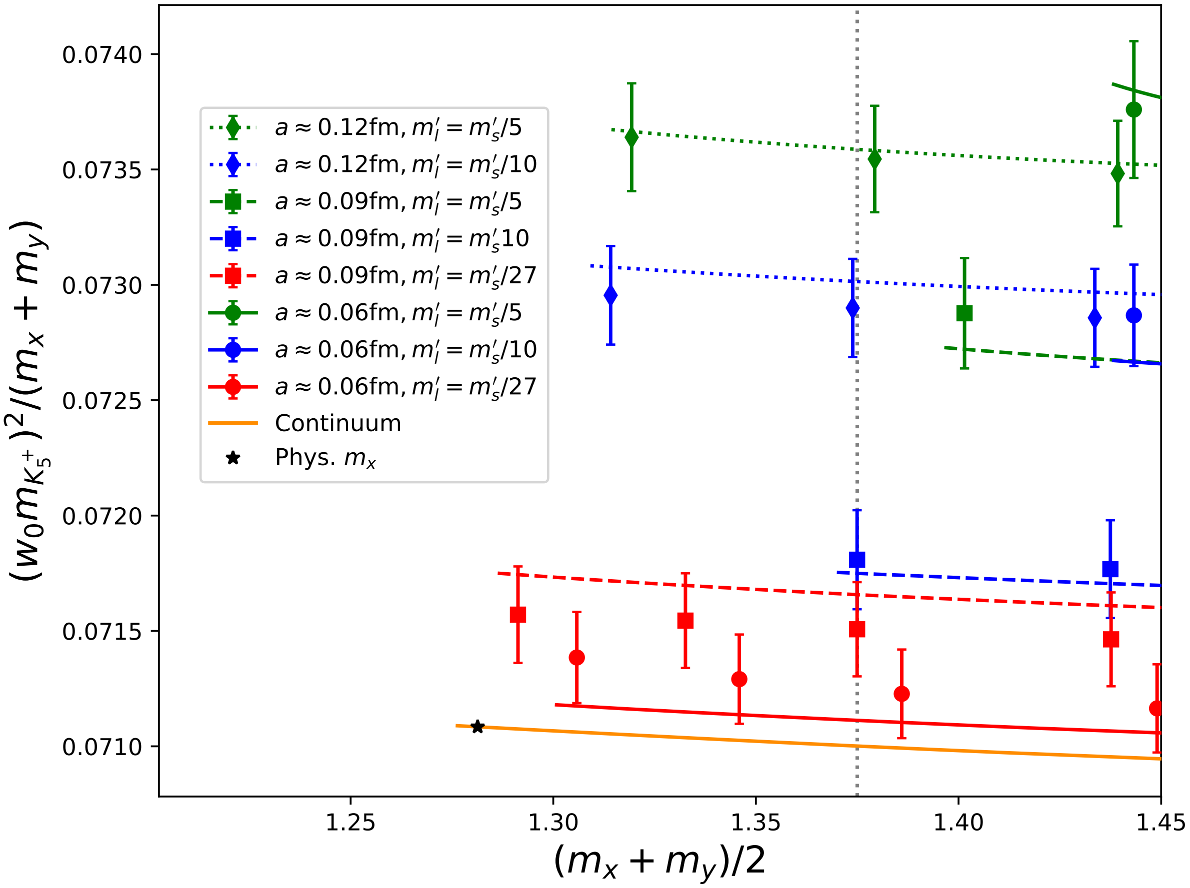

Figure 3.6: Plot of the Goldstone kaon-like meson mass $m^2_{K_5^+}/(m_x + m_y)$ in $w_{0,\mathrm{imp}}$ units for $m_{K_5^+}$ and m_{p4s} units for the quark masses m_x and m_y, versus the average valence quark mass $(m_x + m_y)/2$ for all of the nearly-physical strange quark mass ensembles including in the central fit. All points are kaon-like because one valence mass is set equal to the sea strange-quark mass while the other is free to change. Symbols denote different lattice spacings and the color denotes the ratio of the light and strange sea-quark masses, as per the legend. The fit evaluated at physical sea-quark masses and continuum is denoted by the dark orange "Continuum" line, and the point where the free quark is set equal to the physical m_u is indicated by the black star.

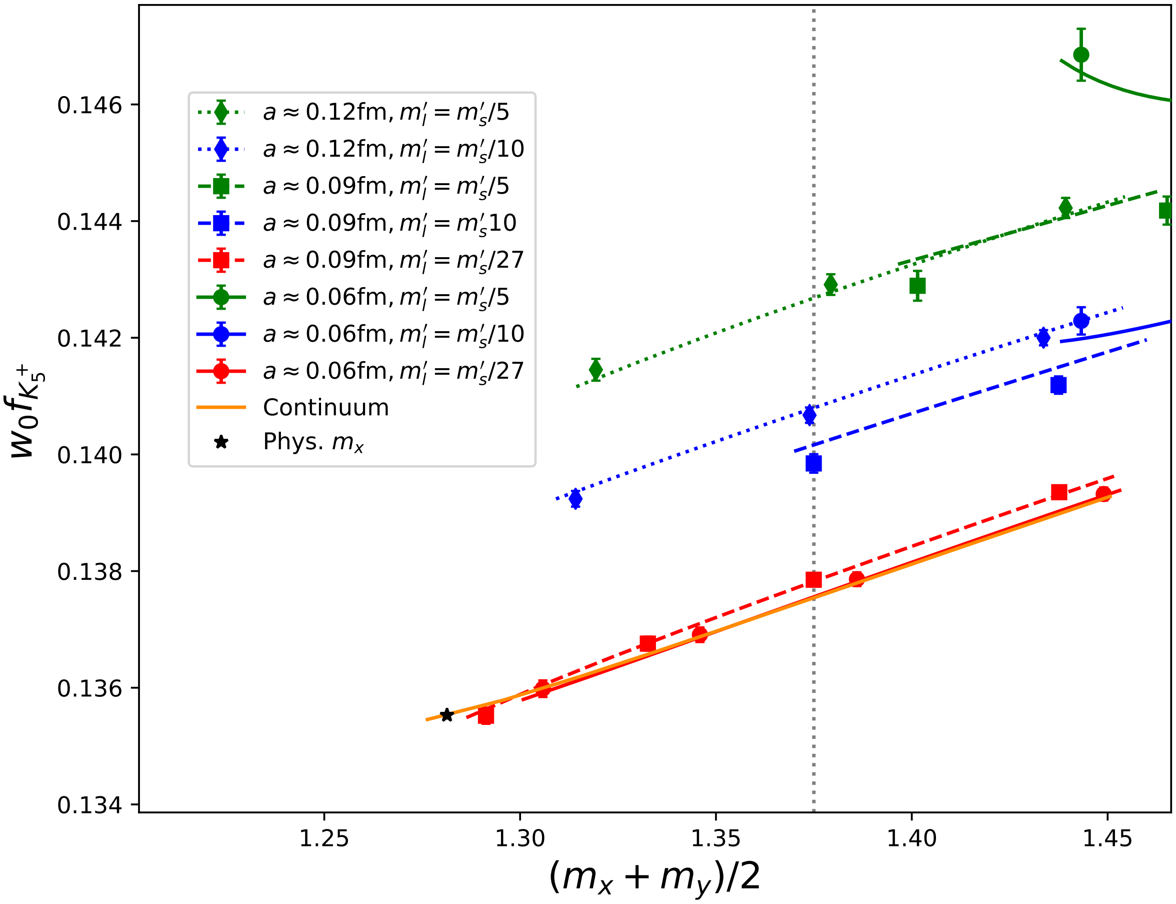

Figure 3.7: Plot of the Goldstone kaon-like decay constant $f_{K_5^+}$ in $w_{0,\text{imp}}$ units versus the average valence quark mass $(m_x + m_y)/2$ for all of the nearly-physical strange quark mass ensembles including in the central fit. The rest of the figure description is similar to Fig. 3.6.

(though within 2σ). While the finite volume effects are likely still large, this may also be explained by the large correlations between measurements on the same underlying ensemble.

3.4 Continuum Results

3.4.1 Extraction Procedure

In order to extract physical results from the central fit, the fit must be evaluated in the continuum and at physical values of each of the light quark masses. The continuum is fairly easy to reach; just set $a = 0$ in the input to the central fit. However, evaluating the fit at physical quark masses is less trivial.

To determine what values of the sea quark masses correspond to the physical values, the predictions of the central fit are tuned to reproduce three experimentally known quantities: m_{π^0}, m_{K^0}, and m_{K^+}. The quark masses that reproduce all three experimental results when corrected for isospin-breaking and electromagnetic effects are then defined as the physical quark masses for m_u, m_d, and m_s.

The value of m_c is not determined from this analysis; however, the dynamical effects from mistunings in the simulation charm quark masses are taken into account. Leading order effects on the meson masses in w_0 units are taken into account using the same procedure described in Sec. 2.3.2. Higher order effects could also be included as corrections to the LO χPT LECs; however, they are not included at this time because their effects are expected to be negligible.

To clarify the procedure for adjusting the u, d, and s quark masses, start with the case of $N_f = 2+1$ flavors and let the central fit function for $m^2_{P_5^+}$ evaluated in the continuum be denoted by $g(m_x, m_y; m_l, m_s)$ where all quark masses are in units of m_{p4s}. Also define the pion-like and kaon-like outputs of the central fit to be $g_\pi(m_l, m_s) \equiv g(m_l, m_l, m_l, m_s)$ and $g_K(m_l, m_s) = g(m_l, m_s, m_l, m_s)$.

Given an initial set of values for the two unique sea masses m_l, m_s, the goal is to determine a new set of values m_l', m_s' such that $g_\pi(m_l', m_s')$ and $g_K(m_l', m_s')$ are closer to the physical values m_π and m_K (in the isospin-symmetric theory with no QED), respectively. This can be accomplished by repeatedly performing linear extrapolations in m_l and then in m_s, updating m_l and m_s after each iteration. Let dm be a very small mass increment. Then the new values of the sea quark masses are given by

$$\begin{aligned} m_l' &= m_l + \frac{2dm}{g_\pi(m_l + dm, m_s) - g_\pi(m_l - dm, m_s)}[m_\pi - g_\pi(m_l, m_s)] && (3.81) \\ m_s' &= m_s + \frac{2dm}{g_K(m_l', m_s + dm) - g_K(m_l', m_s - dm)}[m_K - g_K(m_l', m_s)] && (3.82) \end{aligned}$$

Finally, the update process is repeated until the fit function is within some desired tolerance of both m_π and m_K. The final values of the quark masses are then the physical values of the down and strange quark mass.

To determine the up quark mass, a similar linear extrapolation procedure is carried out. The primary difference is that the sea quarks are now held fixed and one valence quark m_x

is allowed to vary. The update procedure now matches to m_{K^+}[8] and the output of interest is $g_{K^+}(m_x) \equiv g(m_x, m_s, m_l, m_s)$ where m_l and m_s are fixed at the physical values from the previous interpolations. The update step is given by

$$m'_x = m_x + \frac{2dm}{g_{K^+}(m_x + dm) - g_{K^+}(m_x - dm)}[m_{K^+} - g_{K^+}(m_x)] \tag{3.83}$$

After iterating until convergence is reached, the final value of m'_x is defined as the physical value of m_u.

Electromagnetic and Isospin Corrections

As mentioned before, the "physical" meson masses m_π^2, m_K^2, and $m_{K^+}^2$ in this section correspond to pure QCD within a mostly isospin-symmetric theory. To define these three masses in terms of the experimentally measured pions and kaons, both the missing electromagnetic effects and broken isospin need to taken into account.

The experimental neutral pion mass $m_{\pi^0}^2$ is a sufficient substitute for the adjusted pion mass m_π^2. This is because the experimental value of $m_{\pi^0}^2$ has suppressed leading-order isospin corrections of $O[(m_u - m_d)^2]$ and all neutral mesons have vanishing electromagnetic corrections in the chiral limit [2, 39].

---

[8] While the physical isospin-symmetric theory is partially broken by the new light valence mass in m_{K^+}, the sea quark masses are still isospin symmetric and no QED effects are included in the value of m_{K^+} we match to. The excluded effect of isospin violation in the sea is negligible.

Obtaining the value of the completely isospin symmetric kaon mass m_K^2 from experimental masses requires both isospin and electromagnetic corrections. Isospin corrections are handled at leading order by computing the isospin average kaon mass, in the pure QCD theory [2]

$$m_K^2 \equiv \frac{1}{2}\left(m_{K^+}^2 + m_{K^0}^2\right)_{\text{QCD}} \tag{3.84}$$

Next-to-leading order isospin corrections are suppressed, in a similar fashion to those in $m_{\pi^0}^2$.

Electromagnetic contributions to the experimental kaon masses can be determined through the parameter ϵ that quantifies the violations of Dashen's theorem [39], which states the kaon and pion electromagnetic mass splittings are equal in the chiral limit

$$(m_{K^+}^2 - m_{K_0}^2)_{\text{EM}} = (1+\epsilon)(m_{\pi^+}^2 - m_{\pi^0}^2)_{\text{EM}}. \tag{3.85}$$

Since the experimental pion mass difference $\Delta_\pi = m_{\pi^+}^2 - m_{\pi^0}^2$ is almost entirely determined by electromagnetism [2, 45], Eq. (3.85) can be rewritten as

$$(m_{K^+}^2 - m_{K_0}^2)_{\text{EM}} = (1+\epsilon')\Delta_\pi, \tag{3.86}$$

where ϵ' is parameter closely related to ϵ. By separating the electromagnetic kaon mass contributions $\Delta_{K^+}^\gamma$, $\Delta_{K^0}^\gamma$ from the experimental values, solving for the pure QCD masses, and substituting into Eqs. (3.84) and (3.86), the adjusted kaon mass can be determined entirely in terms of experimental masses and the constants ϵ' and $\Delta^{\gamma}_{K^0}$

$$m_K^2 = \frac{1}{2}\left[\left(m_{K^+}^2 + m_{K^0}^2\right)_{\text{expt}} - 2\Delta^{\gamma}_{K^0} - (1+\epsilon')\Delta_\pi\right]. \tag{3.87}$$

The most recent MILC results for ϵ' and $\Delta^{\gamma}_{K^0}$ can be found in Ref. [18].

Finally, the last adjusted mass $m_{K^+}^2$ needs electromagnetic corrections. The different valence mass in $m_{K^+}^2$ accounts for the leading order isospin corrections in χPT, and higher orders can be neglected, as for the isospin average m_K^2. The electromagnetic correction is derived similarly to Eq. (3.87)

$$m_{K^+}^2 = \left(m_{K^+}^2\right)_{\text{expt}} - \Delta^{\gamma}_{K^0} - (1+\epsilon')\Delta_\pi. \tag{3.88}$$

For completeness, the values for all three adjusted masses are

$$m_{\pi^0} = 134.977 \text{ MeV} \tag{3.89}$$

$$m_K = 494.495 \text{ MeV} \tag{3.90}$$

$$m_{K^+} = 491.405 \text{ MeV} \tag{3.91}$$

Experimental values are taken from PDG [69] and all errors are dropped because no systematic errors have been computed on the results in this work, which are preliminary.

3.4.2 Preliminary Results

With the physical values of m_d, m_u, and m_s set, the decay constants f_{K^+} and $f_\pi \approx f_{\pi^+}$ in pure QCD [9] can be calculated by evaluating the central fit in the continuum. The preliminary results for the first iteration's central fit are:

$$\begin{aligned} f_\pi &= 130.7(4)\text{MeV}, && (3.92)\\ f_{K^+} &= 156.0(3)\text{MeV}, && (3.93)\\ f_{K^+}/f_\pi &= 1.194(2). && (3.94)\end{aligned}$$

All errors are only statistical and, because the central fit has not reached a minimum of χ^2, these statistical errors may be underestimated. However, the current best fit is also on a significantly reduced data set, so we ultimately are able to decrease the statistical errors by adding back in more of the data.

The current FLAG averages [2] are used to establish a baseline for comparison within the lattice community[10]. The ratio of the preliminary results with the FLAG averages are

---

[9] f_π is the same as f_{π^+} to leading order in $m_u - m_d$[2].

[10] For f_{π^+}, the $N_f = 2+1$ simulation average is used because there is not a currently listed $N_f = 2+1+1$ result.

$$\begin{aligned}
\frac{f_{\pi^+}}{f_{\pi^+,\mathrm{FLAG}}} &= 1.003(11)[11]\{3\}, && (3.95)\\
\frac{f_{K^+}}{f_{K^+,\mathrm{FLAG}}} &= 1.0008(30)[25]\{17\}, && (3.96)\\
\frac{(f_{K^+}/f_{\pi^+})}{(f_{\pi^+}/f_{K^+})_{\mathrm{FLAG}}} &= 0.9993(45)[42]\{17\}. && (3.97)
\end{aligned}$$

The total propagated error assuming all contributions are independent is listed in parentheses, the total error from the FLAG result is in square brackets, and the statistical error from this work is in braces. All of the results are currently within 1σ for this work's statistical error and within 0.3σ for the total propagated error. This suggests that, even if systematic errors and a potentially larger statistical error are taken into account, the final results will likely not change much. Also, the current FLAG average for f_{K^+} at $N_f = 2+1+1$ is predominantly determined by previous results from the MILC and HPQCD Collaborations on the same HISQ ensembles as this work. Therefore, a conservative comparison for f_{K^+} or the ratio should really neglect the statistical error and just focus on systematic effects. If the most recent MILC/HPQCD results included in the FLAG averages are used and only compared using systematic errors, then all of the results still agree within $1-\sigma$: $f_{K^+} = 155.92(13)(^{+34}_{-23})$ and $f_{K^+}/f_{\pi^+} = 1.1956(10)(^{+26}_{-18})$ [13].

The light quark mass ratios can also be evaluated

$$\begin{aligned}
m_u/m_d &= 0.470(3), && (3.98)\\
m_s/m_l &= 27.48(5) && (3.99)
\end{aligned}$$

and compared to the current FLAG averages

$$
\begin{aligned}
\frac{(m_u/m_d)}{(m_u/m_d)_{\mathrm{FLAG}}} &= 0.999(120)[120]\{6\}, && (3.100)\\
\frac{(m_s/m_l)}{(m_s/m_l)_{\mathrm{FLAG}}} &= 1.0065(127)[125]\{18\}. && (3.101)
\end{aligned}
$$

The quoted error types are the same as the errors quoted for the decay constants. For the quark mass ratios, systematic errors should not be ignored. For instance, systematic errors on the m_u/m_d ratio are expected to be an order of magnitude larger than the current statistical error in this work [2]. Fortunately, the preliminary result for m_u/m_d is in agreement with the FLAG average well within 1σ of this work's statistical error. For the larger m_s/m_l ratio, the neglected systematic errors are required to bring the current result into alignment. When compared with the FLAG average error taken into account ($(m_s/m_l)_{\mathrm{FLAG}} = 27.30(34)$) the difference is only about 0.5σ.

3.5 Conclusions

This chapter presented the first iteration and preliminary results for a high precision chiral analysis of light pseudoscalars on the MILC Collaboration's HISQ ensembles. A systematic chiral expansion at NLO in rSχPT and NNLO in continuum χPT was successfully applied to ensembles for the single lattice spacing $a \approx 0.12$ fm. A combined continuum extrapolation and chiral interpolation based on the LECs from the systematic fit is in progress. The current optimal fit, though not fully converged, is highly constrained by the nearly-physical quark mass ensembles at $a \approx 0.06$ fm and $a \approx 0.09$ fm. As a consequence, the preliminary results for the decay constants f_{π^+} and f_{K^+}, their ratio f_{K^+}/f_{π^+}, and the quark mass ratios m_s/m_l and m_u/m_d are all in close agreement with current lattice community averages.

The quark masses are properties of elementary particles in QCD. As such, many different phenomenological as well as theoretical results rely on them. For example, the light quark masses govern the extent of explicit chiral-symmetry breaking in the continuum. The decay constants are not experimentally measureable, but lattice results can be combined with experimental decay rates to predict the CKM matrix elements $|V_{us}|$ and $|V_{ud}|$. In this preliminary result, the prediction of f_{K^+}, and therefore $|V_{us}|$ from leptonic decays, is in good agreement with FLAG. As a consequence, the existing tension in the values of $|V_{us}|$ determined from leptonic and semileptonic decays will likely be reinforced by this work, or even sharpened if the final precision is higher than previous results.

3.5.1 Future Work

The most pressing next step is reaching a converged fit with a reasonable χ^2 per degree of freedom on a set of ensembles with a range of lattice spacings and quark masses. Although the preliminary results will most likely not change drastically, an error budget cannot be established without first reaching the minimum χ^2 for such a fit.

Once a fully converged fit is reached, the next step is running through the systematic and full fit optimizations again, with the value of am_{p4s} for each lattice spacing replaced by the prediction from the first fully converged fit. Fortunately, the current value of $m_{\mathrm{p4s}}^{(1)} = 0.4m_s$ in the continuum limit and in units of the original $m_{p4s}^{(0)}$ (where the iteration number is denoted by the superscript) is equal to unity within $1\sigma_{\text{stat}}$.

$$\frac{m_{\text{p4s}}^{(1)}}{m_{\text{p4s}}^{(0)}} = 1.002(6) \tag{3.102}$$

Given this agreement, its possible that updating the value of am_{p4s} will not significantly change the optimized fit results. Instead the slight variation in results will go towards a measurement of the systematic error arising from the final imprecision of am_{p4s} determined self-consistently from the fit. It is also possible for the fit to converge on a statistically different value of am_{p4s} on the next iteration due to the removal of the statistical error originally propagated from the external values of am_{p4s}. In this case, several more iterations of am_{p4s} may be required.

Once a fully converged fit is found and the values of am_{p4s} are settled, the final step is calculating a complete error budget. Several variations of the dataset, model, and parameterizations will be run to assess systematic errors from thinning, higher order discretization choices, finite volume effects beyond NLO chiral level that was included in the fit, charm quark mass mistuning, etc. Also, for quantities extrapolated to physical units and/or compared to experimental quantities, there will be external systematic errors coming from the scale setting quantity, experimental inputs, isospin/EM effects, and others. Each of these errors needs to be carefully sourced and tabulated to derive a total error budget. We can also check that the continuum results are independent of the chiral scale Λ chosen in the chiral logarithms, as required by chiral perturbation theory.

Aside from completing the current version of this analysis, there are also several additions that could be made to improve the final analysis. Given the considerable impact of the finer, nearly-physical mass ensembles on final results, it would be very interesting to include the $a \approx 0.042$ fm and $a \approx 0.03$ fm MILC HISQ ensembles in this work. The gradient flow scales for these finer ensembles have been estimated and are included in Appendix A.

Another revision that could prove beneficial is changing how the statistical errors from the relative lattice spacings $w_{0,\mathrm{imp}}/a$ are propagated into the chiral analysis. Currently, the variance of each lattice spacing modifies a corresponding block of the covariance matrix dealing with ensembles at that lattice spacing. This routine has a tendency to inflate the error everywhere and hide smaller correlations. An alternative, recently employed in [18], is to treat the lattice spacings as parameters to the fit, and constrain them using Bayesian priors. This technique avoids modifying the covariance matrix, while still allowing the fit to dynamically account for expected variations in the lattice spacing.

Appendix A

Gradient Flow on New Ensembles

In this appendix we tabulate the results of running gradient flow on three new ensembles. The details of the ensembles can be found in the most recent MILC/Fermilab Collaboration paper [18]. All quoted errors are statistical only.

$\approx a$(fm)	m'_l/m'_s	$\sqrt{t_0}/a$	w_0/a	$\sqrt{t_{0,\text{imp}}}/a$	$w_{0,\text{imp}}/a$	$w_0/(w_{0,\text{imp}}/a)$ fm
0.042	1/5	3.331(7)	3.881(20)	3.277(7)	3.889(20)	0.0440(5)
0.042	1/27	3.3861(15)	4.024(03)	3.3309(15)	4.033(3)	0.0437(4)
0.03	1/5	19.27(4)	5.263(11)	18.97(4)	5.266(13)	0.0316(4)

Table A.1: Values of the gradient-flow scales on the HISQ nearly-physical strange-quark ensembles with $a < 0.06$ fm. described in [18]. The first two columns identify the ensemble by the approximate lattice spacing and sea-quark mass ratio (similar to “Key” in [18]). The next four columns are the results for the four gradient flow scales in lattice units discussed in 2. The last column is the predicted lattice spacing $w_0/(w_{0,\text{imp}}/a)$ in fm, where w_0 is evaluated at the ensemble meson masses by using the continuum mass dependence of w_0 presented in Sec.2.4.2.

References

[1] S. Aoki et al. Pion decay constant for the Kogut-Susskind quark action in quenched lattice QCD. *Phys. Rev.*, D62:094501, 2000.

[2] S. Aoki et al. Review of lattice results concerning low-energy particle physics. *Eur. Phys. J.*, C77(2):112, 2017.

[3] M. Asakawa et al. Thermodynamics of SU(3) Gauge Theory from Gradient Flow. *Phys.Rev.*, D90:011501, 2014.

[4] C. Aubin and C. Bernard. Pion and kaon masses in staggered chiral perturbation theory. *Phys. Rev.*, D68:034014, 2003.

[5] C. Aubin and C. Bernard. Pseudoscalar decay constants in staggered chiral perturbation theory. *Phys. Rev.*, D68:074011, 2003.

[6] C. Aubin, C. Bernard, Carleton E. DeTar, J. Osborn, Steven Gottlieb, E. B. Gregory, D. Toussaint, U. M. Heller, J. E. Hetrick, and R. Sugar. Light pseudoscalar decay constants, quark masses, and low energy constants from three-flavor lattice QCD. *Phys. Rev.*, D70:114501, 2004.

[7] T. Banks, Leonard Susskind, and John Kogut. Strong-coupling calculations of lattice gauge theories: (1 + 1)-dimensional exercises. *Phys. Rev. D*, 13:1043–1053, Feb 1976.

[8] Bär, Oliver and Golterman, Maarten. Chiral perturbation theory for gradient flow observables. *Phys.Rev.*, D89:034505, 2014.

[9] A. Bazavov, C. Bernard, C. M. Bouchard, C. DeTar, Daping Du, A. X. El-Khadra, J. Foley, E. D. Freeland, E. Gámiz, Steven Gottlieb, U. M. Heller, Jongjeong Kim, A. S. Kronfeld, J. Laiho, L. Levkova, P. B. Mackenzie, E. T. Neil, M. B. Oktay, Si-Wei Qiu, J. N. Simone, R. Sugar, D. Toussaint, R. S. Van de Water, and Ran Zhou. Determination of $|V_{us}|$ from a lattice qcd calculation of the $k \to \pi\ell\nu$ semileptonic form factor with physical quark masses. *Phys. Rev. Lett.*, 112:112001, Mar 2014.

[10] A. Bazavov et al. Nonperturbative QCD Simulations with 2+1 Flavors of Improved Staggered Quarks. *Rev. Mod. Phys.*, 82:1349–1417, 2010.

[11] A. Bazavov et al. Scaling studies of QCD with the dynamical HISQ action. *Phys.Rev.*, D82:074501, 2010.

[12] A. Bazavov et al. Lattice QCD ensembles with four flavors of highly improved staggered quarks. *Phys.Rev.*, D87(5):054505, 2013.

[13] A. Bazavov et al. Charmed and light pseudoscalar meson decay constants from four-flavor lattice QCD with physical light quarks. *Phys.Rev.*, D90(7):074509, 2014.

[14] A. Bazavov et al. Equation of state in (2+1)-flavor QCD. *Phys.Rev.*, D90(9):094503, 2014.

[15] A. Bazavov et al. Gradient Flow Analysis on MILC HISQ Ensembles. *PoS*, LATTICE2014:090, 2014.

[16] A. Bazavov et al. Symanzik flow on HISQ ensembles. *PoS*, LATTICE2013:269, 2014.

[17] A. Bazavov et al. Gradient flow and scale setting on MILC HISQ ensembles. *Phys. Rev.*, D93(9):094510, 2016.

[18] A. Bazavov et al. B- and D-meson leptonic decay constants from four-flavor lattice QCD. 2017.

[19] C. Bernard. Chiral logs in the presence of staggered flavor symmetry breaking. *Phys. Rev.*, D65:054031, 2002.

[20] C. Bernard. Staggered chiral perturbation theory and the fourth-root trick. *Phys. Rev.*, D73:114503, 2006.

[21] C. Bernard et al. The Static quark potential in three flavor QCD. *Phys.Rev.*, D62:034503, 2000.

[22] Claude Bernard, Maarten Golterman, and Yigal Shamir. Effective field theories for QCD with rooted staggered fermions. *Phys. Rev.*, D77:074505, 2008.

[23] Claude W. Bernard, Tom Burch, Thomas A. DeGrand, Carleton E. Detar, Steven A. Gottlieb, Urs M. Heller, James E. Hetrick, Kostas Orginos, Bob Sugar, and Doug Toussaint. Scaling tests of the improved Kogut-Susskind quark action. *Phys. Rev.*, D61:111502, 2000.

[24] Claude W. Bernard and Maarten F. L. Golterman. Partially quenched gauge theories and an application to staggered fermions. *Phys. Rev.*, D49:486–494, 1994.

[25] Werner Bernreuther and Werner Wetzel. Decoupling of Heavy Quarks in the Minimal Subtraction Scheme. *Nucl.Phys.*, B197:228, 1982.

[26] Werner Bernreuther and Werner Wetzel. Erratum: Decoupling of Heavy Quarks in the Minimal Subtraction Scheme. *Erratum-ibid*, B513:758, 1998.

[27] Johan Bijnens, Niclas Danielsson, and Timo A. Lahde. The Pseudoscalar meson mass to two loops in three-flavor partially quenched chiPT. *Phys. Rev.*, D70:111503, 2004.

[28] Johan Bijnens, Niclas Danielsson, and Timo A. Lahde. Three-flavor partially quenched chiral perturbation theory at NNLO for meson masses and decay constants. *Phys. Rev.*, D73:074509, 2006.

[29] Johan Bijnens and Timo A. Lahde. Decay constants of pseudoscalar mesons to two loops in three-flavor partially quenched (chi)PT. *Phys. Rev.*, D71:094502, 2005.

[30] Tom Blum, Carleton E. Detar, Steven A. Gottlieb, Kari Rummukainen, Urs M. Heller, James E. Hetrick, Doug Toussaint, R. L. Sugar, and Matthew Wingate. Improving flavor symmetry in the Kogut-Susskind hadron spectrum. *Phys. Rev.*, D55:R1133–R1137, 1997.

[31] Borsanyi, S. and others. High-precision scale setting in lattice QCD. *JHEP*, 1209:010, 2012.

[32] L. S. Brown. *Quantum field theory.* Cambridge University Press, 1994.

[33] Mattia Bruno and Rainer Sommer. On the N_f-dependence of gluonic observables. *PoS*, LATTICE2013:321, 2014.

[34] W. Chen et al. Decay Constants of Pseudoscalar *D*-mesons in Lattice QCD with Domain-Wall Fermion. *Phys.Lett.*, B736:231–236, 2014.

[35] M. A. Clark and A. D. Kennedy. The RHMC algorithm for two flavors of dynamical staggered fermions. *Nucl. Phys. Proc. Suppl.*, 129:850–852, 2004. [,850(2003)].

[36] M. A. Clark and A. D. Kennedy. Accelerating dynamical fermion computations using the rational hybrid Monte Carlo (RHMC) algorithm with multiple pseudofermion fields. *Phys. Rev. Lett.*, 98:051601, 2007.

[37] M. A. Clark and A. D. Kennedy. Accelerating Staggered Fermion Dynamics with the Rational Hybrid Monte Carlo (RHMC) Algorithm. *Phys. Rev.*, D75:011502, 2007.

[38] P. H. Damgaard and K. Splittorff. Partially quenched chiral perturbation theory and the replica method. *Phys. Rev.*, D62:054509, 2000.

[39] Roger F. Dashen. Chiral SU(3) x SU(3) as a symmetry of the strong interactions. *Phys. Rev.*, 183:1245–1260, 1969.

[40] C. Davies et al. The Determination of alpha(s) from lattice QCD with 2+1 flavors of dynamical quarks. *Nucl.Phys.Proc.Suppl.*, 119:595–597, 2003.

[41] Albert Deuzeman and Urs Wenger. Gradient flow and scale setting for twisted mass fermions. *PoS*, LATTICE2012:162, 2012.

[42] R.J. Dowdall et al. Vus from pi and K decay constants in full lattice QCD with physical u, d, s and c quarks. *Phys.Rev.*, D88:074504, 2013.

[43] Z. Fodor et al. The lattice gradient flow at tree-level and its improvement. *JHEP*, 1409:018, 2014.

[44] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. P. Lepage, J. Shigemitsu, H. Trottier, and K. Wong. Highly improved staggered quarks on the lattice with applications to charm physics. *Phys. Rev. D*, 75:054502, Mar 2007.

[45] J. Gasser and H. Leutwyler. Chiral Perturbation Theory: Expansions in the Mass of the Strange Quark. *Nucl. Phys.*, B250:465–516, 1985.

[46] Maarten F. L. Golterman and Jan Smit. Selfenergy and Flavor Interpretation of Staggered Fermions. *Nucl. Phys.*, B245:61–88, 1984.

[47] Anna Hasenfratz. Improved gradient flow for step scaling function and scale setting. *PoS*, LATTICE2014:talk presented at Lattice 2014, to be published in the proceedings., 2015.

[48] R. Horsley et al. SU(3) flavour symmetry breaking and charmed states. *PoS*, LATTICE2013:249, 2014.

[49] A. D. Kennedy, Ivan Horvath, and Stefan Sint. A New exact method for dynamical fermion computations with nonlocal actions. *Nucl. Phys. Proc. Suppl.*, 73:834–836, 1999. [,834(1998)].

[50] G. W. Kilcup and Stephen R. Sharpe. A Tool Kit for Staggered Fermions. *Nucl. Phys.*, B283:493–550, 1987.

[51] John B. Kogut and Leonard Susskind. Hamiltonian Formulation of Wilson's Lattice Gauge Theories. *Phys. Rev.*, D11:395–408, 1975.

[52] J. F. Lagae and D. K. Sinclair. Improved staggered quark actions with reduced flavor symmetry violations for lattice QCD. *Phys. Rev.*, D59:014511, 1999.

[53] Weon-Jong Lee and Stephen R. Sharpe. Partial flavor symmetry restoration for chiral staggered fermions. *Phys. Rev.*, D60:114503, 1999.

[54] G. Peter Lepage. Flavor symmetry restoration and Symanzik improvement for staggered quarks. *Phys. Rev.*, D59:074502, 1999.

[55] Peter Lepage. Perturbative improvement for lattice QCD: An Update. *Nucl. Phys. Proc. Suppl.*, 60A:267–278, 1998. [,267(1997)].

[56] Yu-bing Luo. Improvement of the staggered fermion operators. *Phys. Rev.*, D55:353–361, 1997.

[57] Lüscher, M. Future applications of the Yang-Mills gradient flow in lattice QCD. *PoS*, LATTICE2013:016, 2014.

[58] Lüscher, Martin. Properties and uses of the Wilson flow in lattice QCD. *JHEP*, 1008:071, 2010.

[59] Lüscher, Martin. Trivializing maps, the Wilson flow and the HMC algorithm. *Commun.Math.Phys.*, 293:899–919, 2010.

[60] Lüscher, Martin and Weisz, Peter. Perturbative analysis of the gradient flow in non-abelian gauge theories. *JHEP*, 1102:051, 2011.

[61] Manohar and Wise. *Heavy Quark Physics.* Cambridge University Press, 2000.

[62] MILC Code. `http://www.physics.utah.edu/~detar/milc/milc_qcd.html`. Version 7.7.11: 12-29-2013.

[63] I. Montvay and G. Munster. *Quantum fields on a lattice.* Cambridge University Press, 1997.

[64] A. Morel. Chiral Logarithms in Quenched QCD. *J. Phys.(France)*, 48:1111–1119, 1987.

[65] R. Narayanan and H. Neuberger. Infinite N phase transitions in continuum Wilson loop operators. *JHEP*, 0603:064, 2006.

[66] K.A. Olive et al. Review of Particle Physics. *Chin.Phys.*, C38:090001, 2014.

[67] Kostas Orginos and Doug Toussaint. Testing improved actions for dynamical Kogut-Susskind quarks. *Phys. Rev.*, D59:014501, 1999.

[68] Kostas Orginos, Doug Toussaint, and R. L. Sugar. Variants of fattening and flavor symmetry restoration. *Phys. Rev.*, D60:054503, 1999.

[69] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016.

[70] A Ramos. Wilson flow and renormalization. *PoS*, LATTICE2014:talk presented at Lattice 2014, to be published in the proceedings, 2015.

[71] Stefan Scherer. Introduction to chiral perturbation theory. *Adv. Nucl. Phys.*, 27:277, 2003. [,277(2002)].

[72] Stephen R. Sharpe. Quenched chiral logarithms. *Phys. Rev.*, D46:3146–3168, 1992.

[73] Stephen R. Sharpe. B(K) Using staggered fermions: An update. *Nucl. Phys. Proc. Suppl.*, 34:403–406, 1994.

[74] Stephen R. Sharpe and Noam Shoresh. Physical results from unphysical simulations. *Phys. Rev.*, D62:094503, 2000.

[75] Stephen R. Sharpe and Noam Shoresh. Partially quenched chiral perturbation theory without Phi0. *Phys. Rev.*, D64:114510, 2001.

[76] R. Sommer. A New way to set the energy scale in lattice gauge theories and its applications to the static force and alpha-s in SU(2) Yang-Mills theory. *Nucl.Phys.*, B411:839–854, 1994.

[77] Rainer Sommer. Scale setting in lattice QCD. *PoS*, LATTICE2013:015, 2014.

[78] M. Srednicki. *Quantum field theory.* Cambridge University Press, 2007.

[79] Leonard Susskind. Lattice Fermions. *Phys. Rev.*, D16:3031–3039, 1977.

[80] K. Symanzik. Continuum Limit and Improved Action in Lattice Theories. 1. Principles and phi**4 Theory. *Nucl. Phys.*, B226:187–204, 1983.