

© Copyright by Damian Paul Menscher, 2005

CHARMONIUM AND CHARMED MESONS WITH IMPROVED LATTICE QCD

BY

DAMIAN PAUL MENSCHER

B.S., Brigham Young University, 1997

M.S., University of Illinois at Urbana-Champaign, 1998

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Physics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2005

Urbana, Illinois

Abstract

I present calculations of the properties of charmonium and charmed mesons with lattice QCD. The calculations use improved actions and 2+1 flavors of dynamical quarks. After a review of the relevant phenomenology, I provide an introduction to lattice techniques including the specific actions used in this analysis. Details of the computational techniques used are given, and the reader is guided through the process of extracting physical results from the data. Results include calculations of the charmonium, D_s , and D mass spectra and the decay constants f_{D_s} and f_D .

To my father, who introduced me to science and computing,
and
to my mother, who wouldn't let me give up.

Acknowledgments

I thank my advisor, Aida El-Khadra, for her help and guidance in this work. I also thank my collaborators in the Fermilab Lattice, MILC, and HPQCD collaborations. This work was supported in part by the U.S. Department of Energy under grant DE FG02 91ER40677. Additional funding came from the following organizations within UIUC: the Department of Physics, CCSO/CITES, and the Beckman Imaging Technology Group.

Table of Contents

Chapter 1	Introduction	1
Chapter 2	Phenomenology	3
2.1	The Charmonium System	6
2.2	The D_s and D Systems	8
2.3	The CKM Matrix	10
2.3.1	Hadronic Matrix Elements	12
2.3.2	Current Constraints on the CKM Matrix	15
Chapter 3	Lattice Theory	17
3.1	QCD in the Continuum	17
3.2	Discretization of Space-Time	18
3.2.1	Lattice Gauge Action	21
3.2.2	Naïve Fermion Action	22
3.3	The Doubling Problem	22
3.3.1	Wilson Fermions	23
3.3.2	Staggered Fermions	28
3.4	Improvement	30
3.4.1	Symanzik Improvement	31
3.4.2	On-Shell Improvement	34
3.4.3	Tadpole Improvement	37
3.4.4	Classical Improvement of the Staggered Action: The Naik Term	39
3.4.5	Improving the Taste Symmetry of the Staggered Action	39
3.4.6	Improving the Gauge Action	42
3.4.7	Putting it All Together	43
3.5	Calculating Observable Quantities	44
3.5.1	Hadron Spectra	46
3.5.2	Decay Constants	48
3.6	Gold-Plated Quantities	49
Chapter 4	Methodology	50
4.1	Computational Resources	50
4.1.1	MILC Resources	50
4.1.2	FNAL Resources	50
4.2	Generating the Gauge Configurations	52
4.3	Obtaining Physical Results	53
4.4	Fitting to Non-Linear Multi-Dimensional Functions	55

4.5	Constrained Curvefitting	56
4.5.1	Testing the Priors	57
4.6	Extracting Physics	58
4.6.1	Splittings	58
4.6.2	Kinetic Masses	58
4.7	Summary of Fitting Procedure	59
4.7.1	Goals	59
4.7.2	Plan of Attack	59
4.8	Systematic Errors: Estimation or Removal	62
4.8.1	Finite Lattice-Spacing Errors	62
4.8.2	Finite Volume Effects	63
4.8.3	Violation of the Chiral Limit	63
Chapter 5 Fitting Procedure in Detail		64
5.1	Preparation: Statistical and Numerical Techniques	64
5.1.1	Jackknife Method	64
5.1.2	Statistical Uncertainties	65
5.1.3	Correlations and Blocking	67
5.1.4	Automatic Differentiation	67
5.2	Problems and Solutions	69
5.3	Charmonium Gives a^{-1}	69
5.3.1	The Model Function	69
5.3.2	Effective Mass	70
5.3.3	Local, $1S$, and $2S$ Smearings	72
5.3.4	Chaining Procedure for Starting Values	72
5.3.5	Setting the Priors	73
5.3.6	Initial Results	75
5.3.7	Cutting Out Anomalous Bootstraps	77
5.3.8	A More Robust Fitting Procedure	79
5.3.9	Testing for Bias	82
5.3.10	Splittings and Determination of a^{-1}	85
5.4	D_s Gives m_c	85
5.4.1	Introducing a Light Quark	87
5.4.2	The Heavy-Light Model Function	87
5.4.3	Kinetic Masses	88
5.4.4	Tuning m_c	90
5.4.5	The D_s Decay Constant	94
5.5	Extrapolation to the D	95
5.5.1	Chiral Extrapolations in PQQCD	95
5.5.2	Mass Splittings	95
5.5.3	Staggered Chiral Perturbation Theory	98
Chapter 6 Results		103
6.1	The Charmonium System	103
6.2	The D_s and D Systems	108
Chapter 7 Conclusions		111

Appendix A	Estimation of Operator Effects	113
Appendix B	Grassmann Variables	117
Appendix C	Numerical Techniques	119
	C.1 Singular Matrices	119
	C.2 Solving a System of Linear Equations	120
	C.3 Multidimensional Curvefitting	121
	C.4 Singular Value Decomposition	124
Appendix D	Model Function Derivatives	126
Appendix E	Partially Quenched Staggered χPT	129
Appendix F	Asymmetric Actions	132
References	136

List of Tables

2.1	A summary of the fermions and bosons in the Standard Model	4
3.1	Transformation of various operators under parity conjugation	33
3.2	Coefficients for the asqtad derivative	42
3.3	Gamma-matrix structure for various quantum states	46
4.1	Parameters of the MILC lattices	51
5.1	The priors used in the charmonium analysis	75
5.2	Two plausible local minima for the h_c $1S$ - $1S$ smearing	80
5.3	Lattice spacings measured on the MILC lattices	86
5.4	The momentum states used in the kinetic mass fits	90
5.5	Rest and kinetic masses for heavy-heavy and heavy-light mesons	92
6.1	Mass splittings in the charmonium spectrum	105
6.2	Calculated splittings in the D_s system	109
6.3	Calculated splittings in the D system	109
6.4	Decay constants f_{D_s} and f_D compared to experiment	110
A.1	Table of potential model values	115
A.2	Potential model predictions for various charmonium states	115
A.3	Potential model predictions for various bottomonium states	115
E.1	Taste splittings	131

List of Figures

2.1	The charmonium spectrum, as seen by experiment	7
2.2	A pant-leg hadronic decay mode	7
2.3	The D_s spectrum as seen by experiment	9
2.4	The D spectrum as seen by experiment	10
2.5	A fit to the apex of the Unitarity Triangle shown in the $(\bar{\rho}, \bar{\eta})$ plane	12
2.6	The quark-level diagram of leptonic D_s^+ decay	13
3.1	A small segment of the hypercubic lattice, showing a site, a link, and a plaquette	21
3.2	Effects of unquenching on several observables	29
3.3	Lattice Feynman diagrams	37
3.4	A graphical representation of terms used in the asqtad action	41
3.5	Improvement operators for the gauge action	42
5.1	Example of the effect of priors and different types of errorbars	67
5.2	Charmonium effective mass plot: δ - δ	71
5.3	Charmonium effective mass plot: $1S$ - $1S$	72
5.4	A test of the effect of the priors	74
5.5	Charmonium fits	75
5.6	Fits to the η_c and the η'_c showing the effect of smearing	76
5.7	Effective mass plots corresponding to our matrix of smearings	78
5.8	Fit results as a function of starting guess	80
5.9	Test of stability with respect to n_{states}	83
5.10	Charmonium splittings	86
5.11	D_s fits	88
5.12	D_s splittings	89
5.13	Fit to the energy-momentum dispersion relation	91
5.14	The mass of the D_s for two different values of κ_{charm}	91
5.15	A m_{val} extrapolation in the D system	96
5.16	The selection of sea- and valence-quark masses we used for our PQS χ PT ex- trapolation	97
5.17	Slices through the PQS χ PT extrapolation. Correlated fits for each am_{sea} are shown	99
5.18	A slice through our PQS χ PT extrapolation corresponding to full QCD	100
5.19	A slice corresponding to full QCD through the PQS χ PT extrapolation for the f_D/f_{D_s} ratio	102
6.1	The inverse lattice spacings for the coarse lattices	104
6.2	The charmonium spectrum shown in terms of mass splittings	104

6.3	The $\overline{2S} - \overline{1S}$ splitting	105
6.4	The $J/\psi - \eta_c$ hyperfine splitting	106
6.5	The $\psi' - \eta'_c$ hyperfine splitting	106
6.6	The ratio of hyperfine splittings: $(\psi' - \eta'_c)/(J/\psi - \eta_c)$	107
6.7	The $h_c - \chi_{c0}$ splitting	107
6.8	The $h_c - \chi_{c1}$ splitting	108
6.9	The kinetic mass of the charmonium $1S$	109
A.1	Plots of the Cornell and Richardson potential wells	116
A.2	Plots of the Schrödinger wavefunctions for charmonium and bottomonium . . .	116
C.1	A pictorial representation of the importance of the condition number of a linear system	120
C.2	Possible steps for the amoeba algorithm	122
D.1	Sample code for derivatives	128

Chapter 1

Introduction

Nature appears to be described by four primary forces: strong, electromagnetic, weak, and gravitational. The Standard Model (SM) is widely regarded as a correct theory of the strong and electroweak interactions. Within the SM, Quantum Chromodynamics (QCD) describes the strong interaction, which binds quarks and gluons together into hadrons. QCD is asymptotically free, which means that the strong coupling constant, α_s , becomes small at the high energies probed by particle accelerators. In this regime, perturbation theory can be used in a small- α_s expansion to calculate observables. QCD is also confining, which is a statement of the fact that we do not observe free quarks, or, more precisely, that all observed particles are colorless. Although the mechanism leading to confinement is not yet well understood, we know that hadronization corresponds to $\alpha_s \sim 1$, so perturbation theory may not be used. In the low-energy regime it is therefore necessary to use a non-perturbative approach. One non-perturbative approach is given by *Lattice QCD*, the application of lattice field theory to QCD.

Lattice QCD was introduced by Kenneth Wilson in 1974 as a tool to investigate confinement [1]. In lattice QCD, the free parameters of QCD are mapped directly onto a discretized theory. Solving the discrete theory (where possible) is therefore equivalent to solving the continuum theory. The technique of discretizing space-time has several advantageous properties. It provides a non-perturbative approach with an implicit ultraviolet cutoff, therefore solving at the outset any questions of how to regularize results. Also, it is systematically improvable, which means all sources of error can be estimated and reduced in a systematic way. We can therefore use this technique to produce definitive results.

Previous results from lattice QCD used the quenched approximation, in which the ability of gluons to split into a pair of sea quarks is suppressed. This was an uncontrolled approximation which led to $\sim 20\%$ errors. We now have better algorithms which allow us to include realistic sea quark effects. In addition, improved actions provide reduced discretization errors. For some quantities, the remaining errors are particularly small. These “gold-plated” quantities allow us to test lattice QCD.

This work explores QCD on a lattice, and gives results for a selection of observables. Our results are the first to study mesons involving a charm quark which account for all errors. In particular, we use highly-improved actions, simulate 2+1 dynamical quarks, and provide chiral extrapolations to the physical light quark mass. I begin with a brief review of QCD in Chapter 2 and list the phenomenology of charmonium and charmed mesons which will be calculated in this work. Chapter 3 describes how one can discretize QCD on a lattice, and introduces how the naïve actions are improved to reduce discretization errors. This chapter also provides the specific quark and gluon actions we used in this work. The focus of this work was on extracting results from the lattice calculation, and the sophisticated data-analysis techniques which are required. An overview of our approach and the required techniques are described in Chapter 4. A complete walk-through of our fitting procedure, including details of problems encountered and their solutions, is presented in Chapter 5. Finally, a summary of our results for the charmonium, D_s , and D systems is given in Chapter 6. I conclude with some remarks on our findings and comment on the outlook for future research in this field.

Chapter 2

Phenomenology

The Standard Model (SM) describes particle physics interactions via the group $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$. The $SU(3)_c$ group describes the color charge of the strong force, which gives rise to quantum chromodynamics (QCD). The $SU(2)_L \otimes U(1)_Y$ electroweak symmetry breaks, giving the three weak bosons mass and leaving an unbroken $U(1)_{EM}$ theory which describes the massless photon of quantum electrodynamics (QED). We observe excellent agreement between SM predictions and observed phenomena (neglecting gravity), and therefore have high confidence that the theory correctly describes physics at the energy scales we can probe. Before we can look for physics beyond the SM, we must first improve our understanding of the parameters of this theory.

Within the SM, there are three generations of fermions and four types of gauge bosons. Some fundamental properties of each particle¹ are shown in Table 2.1. Each of the three generations of fermions consists of an up-type quark, and down-type quark, a massive lepton, and a massless² neutrino. The weak force, carried by the two W s and the Z , affects all of the fermions. The electromagnetic force, carried by the photon, only interacts with electrically-charged particles. The strong force, carried by the eight gluons, only interacts with particles which have color-charge — the quarks and other gluons.

The SM requires, as inputs, the 9 charged fermion masses (for the 6 quarks and 3 massive leptons). The strong, electromagnetic, and weak forces provide 3 couplings: α_s , α_{EM} , and

¹We use the traditional units of high energy physics, in which $\hbar = 1 = c$. In these units, the mass of a particle is commonly reported in units of energy rather than the more precise energy/ c^2 .

²The Standard Model defines the neutrinos to be massless. We have since discovered neutrino mixing, which implies the neutrinos do have mass.

Table 2.1: A summary of the fermions and bosons in the Standard Model. For each particle, several properties are shown: whether it carries color-charge (C), its electric charge in units of the positron charge (Q), and its mass in high-energy units (m). The neutrinos have recently been shown to have non-zero mass, but the SM defines them to be massless.

		FERMIONS			BOSONS
QUARKS		<u>u</u>	<u>c</u>	<u>t</u>	<u>g</u>
		$C=\text{yes}$	$C=\text{yes}$	$C=\text{yes}$	$C=\text{yes (2)}$
		$Q=+2/3$	$Q=+2/3$	$Q=+2/3$	$Q=0$
		$m=1.5\text{--}4 \text{ MeV}$	$m=1.15\text{--}1.35 \text{ GeV}$	$m=175\text{--}178 \text{ GeV}$	$m=0$
		<u>d</u>	<u>s</u>	<u>b</u>	<u>γ</u>
		$C=\text{yes}$	$C=\text{yes}$	$C=\text{yes}$	$C=\text{no}$
	$Q=-1/3$	$Q=-1/3$	$Q=-1/3$	$Q=0$	
	$m=4\text{--}8 \text{ MeV}$	$m=80\text{--}130 \text{ MeV}$	$m=4.1\text{--}4.4 \text{ GeV}$	$m=0$	
LEPTONS		<u>e^-</u>	<u>μ^-</u>	<u>τ^-</u>	<u>W^\pm</u>
		$C=\text{no}$	$C=\text{no}$	$C=\text{no}$	$C=\text{no}$
		$Q=-1$	$Q=-1$	$Q=-1$	$Q=\pm 1$
		$m=0.511 \text{ MeV}$	$m=106 \text{ MeV}$	$m=1777 \text{ MeV}$	$m=80.4 \text{ GeV}$
		<u>ν_e</u>	<u>ν_μ</u>	<u>ν_τ</u>	<u>Z</u>
		$C=\text{no}$	$C=\text{no}$	$C=\text{no}$	$C=\text{no}$
	$Q=0$	$Q=0$	$Q=0$	$Q=0$	
	$m=0$	$m=0$	$m=0$	$m=91.2 \text{ GeV}$	

$\sin^2 \theta_W$. In addition, the SM leaves 4 free angles in the weak mixing matrix (discussed in Section 2.3) which must be determined. The remaining required parameters are the masses of the Z (M_Z) and Higgs (m_h^0), and the QCD θ parameter, which relates to the topology of the theory [2]. Our understanding of the SM therefore depends on our knowledge of these 19 parameters, which must be measured experimentally. Because the SM is a theory of quarks and leptons, it requires measurements of quark properties and interactions to determine its inputs. However, QCD is confining so we only observe hadrons (color-singlet combinations of quarks) in experiments. It is therefore necessary to combine experimental input with theoretical techniques, including the computational method of Lattice QCD, to determine the fundamental parameters of the theory. In particular, Lattice QCD can be used to determine bare quark masses, CKM matrix elements, and the strong coupling α_s .

Heavy quark physics is of current interest because it can provide information on some elements of the quark mixing matrix. The expected unitarity of this matrix, which provides the couplings of the quarks to the W boson, remains to be tested. The matrix is particularly interesting because it contains a phase which leads to an asymmetry between matter and antimatter. The asymmetry is one source of baryogenesis, and baryogenesis is a prerequisite for our existence. On the lattice, there are two techniques for simulating heavy quarks: nonrelativistically (NRQCD [3, 4]) or by using a relativistic method (called the Fermilab approach [5]). The two methods are expected to agree quite well for systems involving a valence bottom quark, though the Fermilab approach is expected to work slightly better for systems with a valence charm quark, as the charm quark is slightly relativistic. This work focuses on mesons which contain a charm quark, and therefore uses the Fermilab approach. We first look at charmonium, as the spectrum provides an easy test of LQCD and allows us to calibrate our results. We then study the D_s and D systems and examine the spectra and leptonic decay constants. One could combine our predictions of the decay constants with measured decay rates to obtain information about $|V_{cs}|$ and $|V_{cd}|$, two elements of the quark mixing matrix.

2.1 The Charmonium System

Charmonium, the bound state of a charm quark and its antiquark, is the strong interaction equivalent of positronium. Like the hydrogen atom, quarkonium systems provide a rich, yet simple, structure for us to study. Just as the hydrogen atom has a ground state and various excited states, the quarkonia have ground and excited states. For the case of the hydrogen atom, the characteristic scale of these splittings is the Rydberg energy, $m_e c^2 \alpha^2 / 2 \approx 13.6$ eV, which is determined by the electromagnetic force that binds the electrons to the nucleus. In contrast to the hydrogen atom, quarkonium states are separated not by a few eV, but rather by the characteristic scale of the strong force: $\Lambda_{\text{QCD}} \approx 300$ MeV. Because of the large energy difference between states, they are typically thought of as different particles, and named as in Figure 2.1. However, it is important to remember that they all are $c\bar{c}$ states, just with different quantum numbers corresponding to radial or orbital excitations.

The fastest decay mode for a hadron proceeds via a pant-leg diagram. For example, the decay $\rho \rightarrow \pi\pi$, shown in Figure 2.2, takes place in a very short time: $\tau \sim 4.4 \times 10^{-24}$ s. For the lower-lying charmonium states such as the S -wave η_c and J/ψ mesons, decays to a pair of D mesons is kinematically forbidden. Instead, the η_c (J/ψ) must decay via an OZI-suppressed channel³ to 2 (3) pions.⁴ This suppression makes them somewhat “stable” — the lifetime of the J/ψ is $\tau \sim 7.2 \times 10^{-21}$ s, which is about 1600 times the lifetime of the ρ meson. The η_c has a somewhat shorter lifetime ($\tau \sim 4 \times 10^{-23}$ s, or about 10 times the lifetime of the ρ meson) because it decays to two gluons, which removes a factor of α_s from the decay. Even the excited states (ψ' and η'_c) lie somewhat below the $D\bar{D}$ threshold (the energy at which decays to a pair of D mesons are no longer kinematically prohibited) and can therefore be simulated with reasonable systematics.⁵ We also expect to be able to obtain reasonable results for the P -wave states: the h_c and the χ_c triplet. The h_c deserves a special mention: it has only been seen by one experiment [6] (with some evidence for it observed by another [7]), and therefore

³Okubo, Zweig, and Iizuka noted that decays that include hard gluons are suppressed as a result of asymptotic freedom.

⁴Decays to a single pion are not allowed because they would require a singlet gluon. Whether the decays include an even or odd number of gluons depends on the spin of the decaying meson — a spin-0 meson decays to an even number of gluons, while a spin-1 meson decays to an odd number of gluons.

⁵Threshold effects cause the masses to be pushed down (away from the threshold). Since our simulations do not take this into account, we might expect our results for these masses to be slightly high.

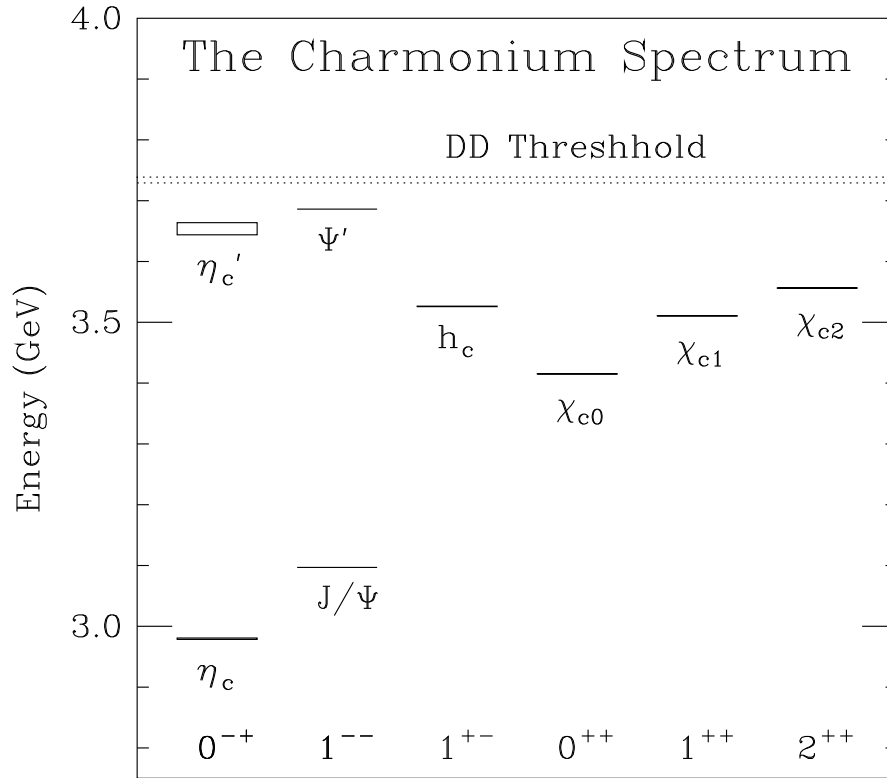


Figure 2.1: The charmonium spectrum, as seen by experiment. The J^{PC} for each particle is shown, where J is the total angular momentum (in units of \hbar), P indicates the particle's behavior under a parity inversion, and C indicates its behavior under a charge inversion. The h_c has only been observed by one experiment [6] (Ref. [7] also found evidence for its existence), and awaits confirmation. Also, its J^{PC} has not been confirmed (though it is believed to be 1^{+-}).

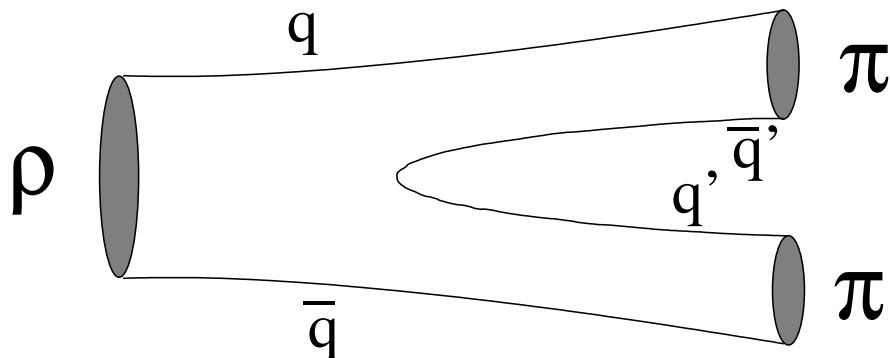


Figure 2.2: The parton-level hadronic decay mode of the ρ . The quarks q and q' in the figure represent the light quarks.

requires confirmation. In addition, its J^P has not been experimentally measured [8].

Note from Figure 2.1 that the charmonium spectrum entails multiple energy scales. The characteristic scale of the strong interaction, Λ_{QCD} , determines the $\overline{2S} - \overline{1S}$ splitting, the splitting between the spin-averaged ψ' and η'_c states and the spin-averaged J/ψ and η_c states. In quarkonium, as in the hydrogen atom, there is a spin-orbit coupling leading to a fine structure, and a spin-spin coupling leading to a hyperfine structure. In contrast to the hydrogen atom, these fields come from the QCD gauge interaction $\sigma_{\mu\nu}F_{\mu\nu}^a$ rather than from the QED equivalent $\sigma_{\mu\nu}F_{\mu\nu}$. As a result, the chromoelectric interaction $\boldsymbol{\sigma} \cdot (\mathbf{D} \times \mathbf{E})$ leads to splittings in the χ_c states of ~ 100 MeV, with dependence on the momentum of the quarks. Meanwhile, the chromomagnetic interaction $\boldsymbol{\sigma} \cdot \mathbf{B}$ leads to $J/\psi - \eta_c$ and $\psi' - \eta'_c$ splittings of ~ 100 MeV, with dependence inversely proportional to the squared mass. The combination of energy scales makes the charmonium spectrum a useful testing-ground of our methods. In addition, there has already been extensive work in understanding quarkonium spectrums through the use of potential models. This work will serve as a useful guide to our current and future studies of these particles.

2.2 The D_s and D Systems

After proving our methods on the charmonium system, we will investigate the D_s and D systems, shown in Figures 2.3 and 2.4, respectively. The D_s and D_s^* states lie below the DK threshold at 2362 MeV, kinematically forbidding the hadronic pant-leg decay. The D_s and D therefore decay weakly, giving them long lifetimes $\tau \sim 10^{-12}$ s. The hadronic decay $D_s^* \rightarrow D_s \pi^0$ is suppressed due to isospin symmetry⁶, so the dominant decay (94%) is $D_s^* \rightarrow D_s \gamma$. The D^* has a similar suppression. They therefore have lifetimes $\tau \sim 3 \times 10^{-22}$ s.

While most particles in these spectra are well understood, a notable exception is the $D_{sJ}^*(2317)$. The recently-observed $D_{sJ}^*(2317)$ is a candidate for the $J^P = 0^+$ state [9, 10]. Like the D_s and D_s^* , the $D_{sJ}^*(2317)$ lies below the DK threshold. It presents a particular

⁶Isospin is the approximate symmetry resulting from the nearly-identical masses of the u and d quarks. This symmetry, which is respected by the strong force, requires that the isospin does not change. Because it is only an approximate symmetry, it only leads to a suppression of isospin-violating processes, not a complete cancellation of them.

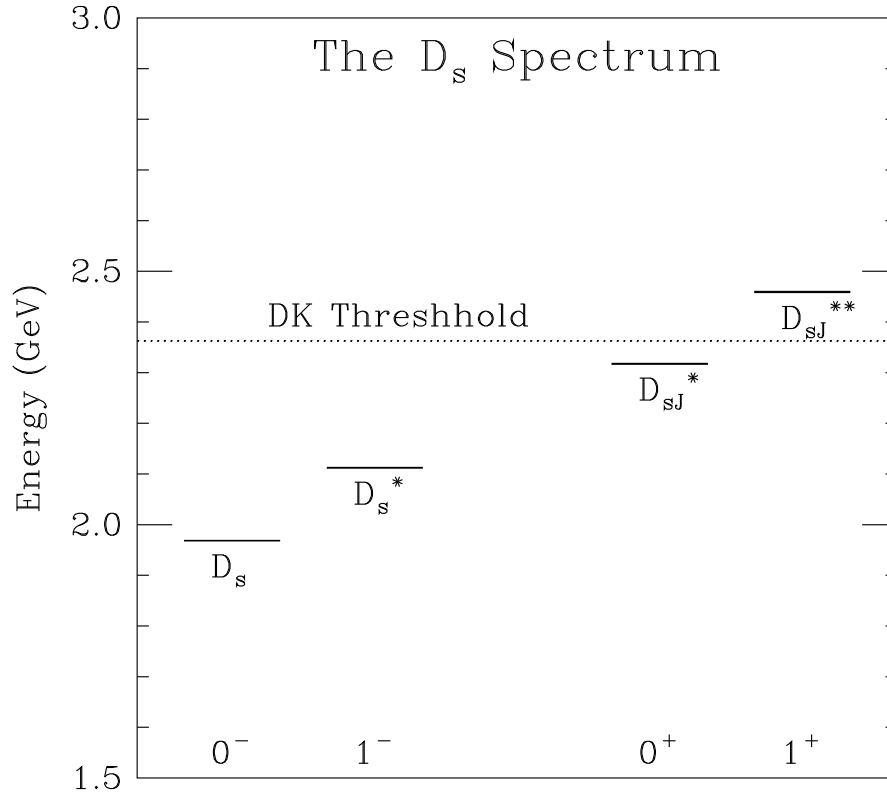


Figure 2.3: The D_s spectrum as seen by experiment. The J^P of each state is shown.

challenge to theory predictions because it is so close to the threshold, and has a narrow width. It has been observed to decay via $D_{sJ}^*(2317) \rightarrow D_s \pi^0$, with a lifetime $\tau > 1.4 \times 10^{-22}$ s. By using non-perturbative methods, we can compute the expected mass of the $D_s(0^+)$ state, and then see whether our results are consistent with the experimentally-observed mass of 2317 MeV. If the measurements are in agreement, then it would provide further evidence that the $D_{sJ}^*(2317)$ is indeed the expected 0^+ state, and not some exotic system. An obvious challenge arises due to the fact that the $D_{sJ}^*(2317)$ lies just below the DK threshold, which may cause our prediction of its mass to be slightly high.⁷

⁷See Footnote 5.

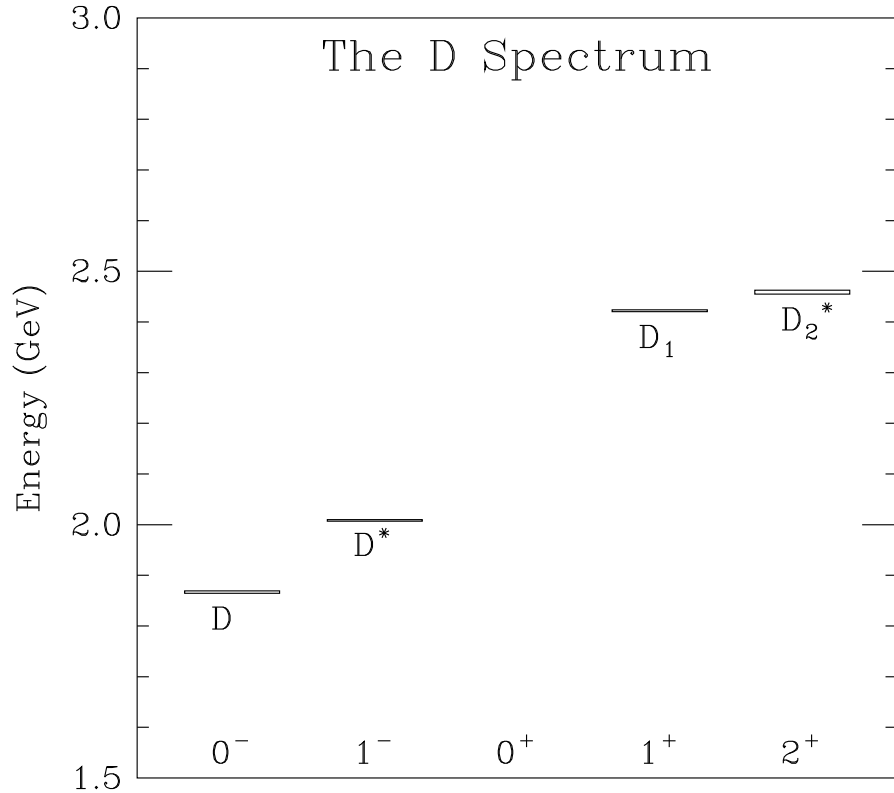


Figure 2.4: The D spectrum as seen by experiment. The J^P of each state is shown.

2.3 The CKM Matrix

The weak interaction allows quarks to change flavor. The weak force doesn't act on the quark mass eigenstates, but rather on “rotated” quark states [11]:

$$\begin{pmatrix} u \\ d' \end{pmatrix}, \quad \begin{pmatrix} c \\ s' \end{pmatrix}, \quad \begin{pmatrix} t \\ b' \end{pmatrix}. \quad (2.1)$$

For the case of two generations, we define the rotation using a unitary matrix as

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \mathbf{U} \begin{pmatrix} d \\ s \end{pmatrix}, \quad (2.2)$$

where

$$\mathbf{U} \equiv \begin{pmatrix} \cos \theta_c & \sin \theta_c \\ -\sin \theta_c & \cos \theta_c \end{pmatrix}. \quad (2.3)$$

We see that the rotation can be represented by a single parameter, θ_c , known as the Cabibbo angle. Experimentally, we find $\sin \theta_c \approx 0.22$.

With the discovery of the third generation of quarks, the unitarity matrix was expanded by Kobayashi and Maskawa [12]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (2.4)$$

With n generations, unitarity allows for $n(n-1)/2$ angles and $(n-1)(n-2)/2$ phases.⁸ For 3 generations, the 3 angles and 1 phase can be chosen as in the Wolfenstein parameterization [14]:

$$\mathbf{V} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \quad (2.5)$$

This is convenient because it organizes the matrix elements in powers of a small parameter, $\lambda \equiv \sin \theta_c$. The remaining parameters, A , ρ , and η , are all real numbers of order unity.⁹

Any two rows or columns of the CKM matrix leads to a unitarity condition.¹⁰ For example, the first and third columns give the condition

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \quad (2.6)$$

The unitarity triangle is simply a geometrical representation of this equation in the complex plane. If it is scaled such that the base is of unit length (forming the line segment from $(0, 0)$

⁸ Unitarity requires that the matrix satisfy $\mathbf{V}^\dagger = \mathbf{V}^{-1}$. An $n \times n$ unitary matrix has n^2 real parameters. Of those, $2n$ change the phase of independent quark states, and do not affect the physics. However, 1 is an overall phase and leaves the matrix invariant. So the physics requires $n^2 - (2n - 1)$ real parameters. Not all can be angles, though, because an orthogonal $n \times n$ matrix has only $\frac{1}{2}n(n-1)$ real parameters. Therefore, $n^2 - (2n - 1) - \frac{1}{2}n(n-1) = \frac{1}{2}(n-1)(n-2)$ of the parameters must be phases [13, § 12.12].

⁹The phase η is responsible for CP violation within the Standard Model. Charge conjugation (C) and parity (P) are symmetries under the strong, electromagnetic, and gravitational interactions, but are violated by the weak interaction. Except in rare cases, however, CP is conserved by the weak interaction. Understanding these cases is a current focus in particle physics, because the CP violation allowed by the SM is insufficient to explain the matter-antimatter asymmetry in the observed universe.

¹⁰Unitarity requires $V^\dagger = V^{-1}$.

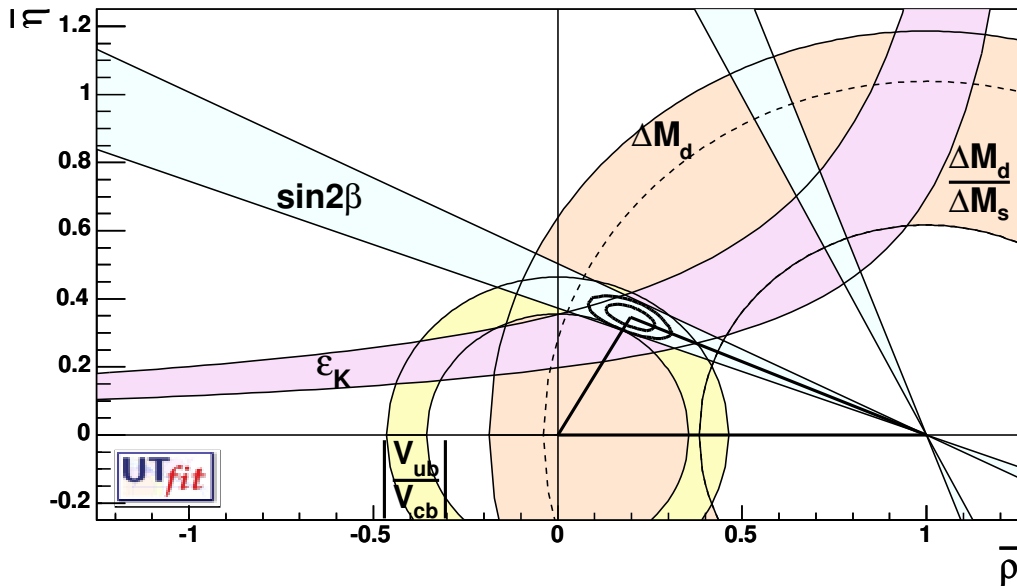


Figure 2.5: A fit to the apex of the Unitarity Triangle shown in the $(\bar{\rho}, \bar{\eta})$ plane. Figure from http://utfit.roma1.infn.it/ckm-results/CKM_fit_today.eps.

to $(1, 0)$) then the apex of the triangle is at $(\bar{\rho}, \bar{\eta})$ where $\bar{\rho} = \rho(1 - \lambda^2/2)$ and $\bar{\eta} = \eta(1 - \lambda^2/2)$ as seen in Figure 2.5. Different experiments can provide different constraints on the apex of the unitarity triangle (examples will be given in Section 2.3.2). Assuming the Standard Model is correct, all experiments should be consistent with a single apex. Therefore, one test of the Standard Model is the careful analysis of the unitarity triangle [15].

2.3.1 Hadronic Matrix Elements

Unfortunately, experimental information from weak decays can not provide enough information to determine the CKM matrix elements. Weak decays depend on two parameters: the CKM matrix element which embodies the short-distance coupling of quarks to the W boson, and a hadronic matrix element that parameterizes our ignorance about the long-distance nonperturbative effects of the hadronic process (see Figure 2.6). Experiments can measure the combined effects of these two parameters, but not the individual parameters. This is because we only measure hadronic properties, and have no direct way to directly compare the hadronic and quark processes. An accurate determination of the CKM matrix elements therefore requires a lattice measurement of the meson decay constants.

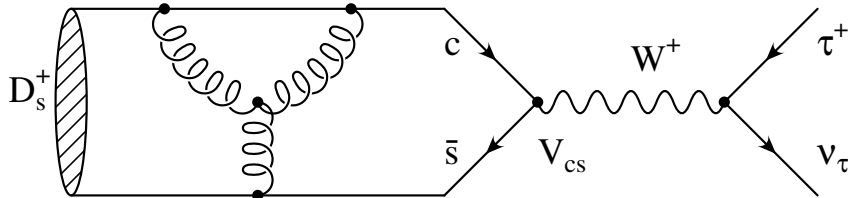


Figure 2.6: The quark-level diagram of leptonic D_s^+ decay. The decay constant f_{D_s} parameterizes our ignorance about the hadronic input to the decay (represented by the virtual gluons in the figure), while the vertex depends only on V_{cs} .

The D_s and D decay weakly to leptons, as shown in Figure 2.6. For a charged pseudoscalar meson P , the decay constant f_P is defined through the matrix element:

$$\langle 0 | A_\mu(0) | P(\mathbf{q}) \rangle \equiv i f_P q_\mu \quad (2.7)$$

where q is the momentum of the meson. Here A_μ is the axial-vector current. For the case of the D_s^+ , one can measure the leptonic decay rate $\Gamma(D_s^+ \rightarrow \ell^+ \nu_\ell)$, which is related to the D_s decay constant through

$$\Gamma(D_s^+ \rightarrow \ell^+ \nu_\ell) = \frac{G_F^2}{8\pi} |V_{cs}|^2 f_{D_s}^2 m_{D_s} m_\ell^2 \left(1 - \frac{m_\ell^2}{m_{D_s}^2}\right)^2 [1 + \mathcal{O}(\alpha)]. \quad (2.8)$$

The leptonic decay to a τ has the largest branching fraction due to the helicity suppression¹¹ of decays to lighter leptons. For the case of the D_s^+ , the helicity suppression causes decays to the τ to be about 10 times more likely than decays to a μ , with decays to the e strongly suppressed. For the lighter case of the D , the helicity suppression is somewhat less severe: the relative widths are in the ratio $2.65 : 1 : 2.3 \times 10^{-5}$ for the $\tau^+ \nu_\tau$, $\mu^+ \nu_\mu$, and $e^+ \nu_e$ final states. Clearly the branching fraction to the electron is too small for precision measurements.

¹¹The D_s has total angular momentum $J = 0$, and therefore the ℓ^+ and ν_ℓ must have opposite spin. Because all neutrinos are left-handed, the ℓ^+ must also be left-handed. However, the W only couples to left-handed particles (right-handed antiparticles), which means it cannot couple to the necessary left-handed ℓ^+ except in a boosted frame. As a result, the probabilities of producing the requisite particles are:

$$P(\text{R-handed}) = \frac{1-v}{2}, \quad P(\text{L-handed}) = \frac{1+v}{2} \quad \text{for a fermion} \quad (2.9)$$

$$P(\text{R-handed}) = \frac{1+v}{2}, \quad P(\text{L-handed}) = \frac{1-v}{2} \quad \text{for an anti-fermion} \quad (2.10)$$

There is therefore a suppression for decays to light leptons, which would have $v \approx 1$.

Because leptonic decay to a τ is very difficult to detect, experiments typically measure the decay to the μ instead.

The CLEO collaboration has recently measured [16] the branching fraction¹²:

$$\mathcal{B}(D^+ \rightarrow \mu^+ \nu_\mu) = (4.45 \pm 0.67_{-0.36}^{+0.29}) \times 10^{-4}. \quad (2.12)$$

They combine this with the lifetime $\tau_{D^+} = 1.040$ ps and CKM matrix element $|V_{cd}| = 0.225$ to find the decay constant

$$f_{D^+} = (223 \pm 16_{-9}^{+7}) \text{ MeV}. \quad (2.13)$$

This improves on the 2004 world average [8] of

$$f_D = (300_{-150}^{+180+80}) \text{ MeV} \quad (2.14)$$

and the previous CLEO-c measurement [17] of

$$f_D = (202 \pm 41 \pm 17) \text{ MeV}. \quad (2.15)$$

The 2004 world average for the D_s decay constant is

$$f_{D_s} = (267 \pm 33) \text{ MeV}, \quad (2.16)$$

We will calculate the decay constants f_{D_s} and f_D on the lattice. In addition, one may combine the lattice calculations of f_{D_s} and f_D with experimental measurements of the branching fractions to obtain $|V_{cs}|$ and $|V_{cd}|$.

¹²The branching fraction is the partial decay width as a fraction of the total decay width:

$$\mathcal{B} \equiv \Gamma_i / \Gamma_{\text{tot}}. \quad (2.11)$$

2.3.2 Current Constraints on the CKM Matrix

Assuming 3-generation unitarity, the magnitudes of the CKM matrix elements are [8]

$$\mathbf{V} = \begin{pmatrix} 0.9745(6) & 0.224(3) & 0.0037(8) \\ 0.224(3) & 0.9737(7) & 0.0415(25) \\ 0.0094(46) & 0.040(3) & 0.9991(1) \end{pmatrix}. \quad (2.17)$$

Knowledge of the matrix elements requires the combination of theory input and experimental measurements of specific processes. In particular, the individual elements are determined in the following ways:

- $|V_{ud}|$ — The usual experiment looks at nuclear β decays that proceed through a vector current to muon decay. The theoretical uncertainty is small due to isospin symmetry.
- $|V_{us}|$ — The experimental evidence comes primarily from $K \rightarrow \pi \ell \nu_\ell$. Theoretical constraints can take advantage of the $SU(3)$ flavor symmetry.
- $|V_{ub}|$ — One method is to look at the semileptonic decay of B mesons, which contains the process $b \rightarrow u \ell \bar{\nu}_\ell$. Unfortunately, the process $b \rightarrow c \ell \bar{\nu}_\ell$ gives a background, and therefore must be well understood. An alternative method is to use exclusive decays, such as $B \rightarrow \pi \ell \nu_\ell$ or $B \rightarrow \rho \ell \nu_\ell$, however that method depends on theoretical models. Either method would benefit from a tighter constraint from theory.
- $|V_{cd}|$ — This can be determined from deep inelastic scattering of neutrinos incident on atomic nuclei. The quark-level process $\nu_\mu d \rightarrow \mu^- c$ is combined with perturbative QCD to obtain the matrix element. A tighter constraint from theory would be beneficial here.
- $|V_{cs}|$ — Direct measurements of charm-tagged W decays are used, with the best result from the ratio of hadronic W decays to leptonic decays. The result may be improved by taking advantage of our knowledge of the u and c rows of the CKM matrix and applying 3-generation unitarity.
- $|V_{cb}|$ — Heavy quark effective theory (HQET) takes advantage of the heavy quark symmetry of the b and c quarks. One can combine HQET with measurements of the exclusive

decay $B \rightarrow \overline{D}^* \ell^+ \nu_\ell$ (or $B \rightarrow \overline{D} \ell^+ \nu_\ell$). Alternatively, one can analyze inclusive decays.

- $|V_{td}|$ — $B - \overline{B}$ mixing can be used to determine this matrix element, which appears in the loop diagram. Large hadronic uncertainties limit the precision of the result.
- $|V_{ts}|$ — $B_s - \overline{B}_s$ mixing or $K - \overline{K}$ mixing can be used to determine this matrix element. Again, large hadronic uncertainties limit the precision of the result.
- $|V_{tb}|$ — This matrix element is unique in that it can be measured perturbatively (without hadronic uncertainties). That is because the timescale of t quark decay is shorter than the hadronization time. The particular process observed is semileptonic decays of $t \rightarrow b \ell^+ \nu_\ell$ as a ratio of all semileptonic decays. Because it is measured as a ratio, and not directly, the associated error is $\sim 30\%$.

Lattice calculations can improve our constraints on $|V_{ub}|$ by calculating the process $B \rightarrow \pi \ell \nu$. Similarly, $|V_{cd}|$ and $|V_{cs}|$ can be obtained through lattice calculations of $D \rightarrow \pi \ell \nu$ and $D \rightarrow K \ell \nu$, respectively.

One goal of this work is to improve our understanding of $|V_{cd}|$ and $|V_{cs}|$ through the study of decays of charmed mesons. These matrix elements pose an interesting challenge, in part because the symmetries which allow improved determinations of several of the other elements do not apply.

Chapter 3

Lattice Theory

3.1 QCD in the Continuum

Quantum Chromodynamics (QCD), the theory of the strong interactions, describes interactions between two classes of fundamental particles: quarks and gluons. The QCD Lagrangian is given by

$$\mathcal{L}_{\text{QCD}} = \bar{\psi}(i\not{D} - m)\psi - \frac{1}{4}(F_{\mu\nu}^a)^2 \quad (3.1)$$

where $D_\mu \equiv \partial_\mu - igA_\mu^a t^a$ is the covariant derivative and $F_{\mu\nu}^a \equiv \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf^{abc}A_\mu^b A_\nu^c$ is the field tensor. Note that the covariant derivative produces a quark-quark-gluon interaction term.

As explained in Chapter 2, QCD is a theory of quarks and gluons, while we observe hadrons. We would like to compute hadron masses and meson decay constants to confirm our understanding of the Standard Model. Although it is concise to write the QCD action, computations can be very complicated. The complexity arises from the ability of the gluons to interact with each other. In addition, the theory is confining — there are no free quarks. The coupling “constant” varies, allowing for asymptotic freedom at high energies. In this regime, the coupling becomes small enough to use as an expansion parameter in a perturbative series. At lower energies, however, the coupling constant increases to be of $\mathcal{O}(1)$, making the perturbative expansion converge slowly, if at all.

3.2 Discretization of Space-Time

We must use a non-perturbative method to calculate low-energy quantities in QCD. One non-perturbative approach to quantum field theory is to discretize space-time on a lattice. The finite lattice spacing introduces an ultraviolet cutoff, while the finite lattice size introduces an infrared cutoff. Building this cutoff into our effective field theory gives us a natural protection from the divergences found in continuum field theories. In addition, solving the resulting equations on a computer enables accurate calculations without depending on perturbation theory which would require α_s to be a small parameter.

Consider the task of discretizing the QCD action. Let ϕ represent both the fermion $(\bar{\psi}, \psi)$ and boson (A_μ^a) degrees of freedom. Then QCD is defined in terms of the path integral

$$\langle 0 | \mathbf{T} \{ \phi(x_1) \cdots \phi(x_n) \} | 0 \rangle \equiv \frac{1}{\mathcal{Z}} \int \mathcal{D}\phi \phi(x_1) \cdots \phi(x_n) e^{i \int d^4x \mathcal{L}_{\text{QCD}}} \quad (3.2)$$

$$\mathcal{Z} \equiv \int \mathcal{D}\phi e^{i \int d^4x \mathcal{L}_{\text{QCD}}} \quad (3.3)$$

where \mathbf{T} denotes a time-ordered product, and $\mathcal{D}\phi$ is the measure of integration. A numerical evaluation of this integral would be impossible because the imaginary exponent causes the integrand to oscillate wildly. We therefore perform a Wick rotation to imaginary time ($t \rightarrow \tau \equiv it$). Here the expectation value does not oscillate. Using the Euclidean metric¹ the QCD action becomes

$$S = \int d^4x \bar{\psi}(x) (\not{D} + m) \psi(x) + \frac{1}{4g^2} \int d^4x F_{\mu\nu}^a F_{\mu\nu}^a. \quad (3.4)$$

The Wick rotation comes at the cost of restricting us to the computation of static properties, such as masses and matrix elements of hadrons. We can't compute properties that depend on the Minkowski metric $g_{\mu\nu}$, which includes the parton distribution function in the proton, or any time-dependent quantity such as jet formation.

We may discretize spacetime by defining our fields $\psi(x)$ to exist only on lattice *sites*. But our action requires a derivative, which will necessitate connecting these lattice sites together. We may move from one position in spacetime to another through the use of a gauge-covariant

¹The Euclidean metric defines $g_{\mu\nu} \equiv \text{diag}(1, 1, 1, 1)$. Therefore we may simplify notation by writing all Lorentz indices as subscripts. For example, on the lattice it is customary to write $\not{D} = \gamma_\mu D_\mu$.

connection, which we define along an infinitesimal path as

$$U(x, x + dx) \equiv \mathbb{1} + igA_\mu(x)dx_\mu. \quad (3.5)$$

Over a finite interval, such infinitesimals would be multiplied together to get

$$\begin{aligned} U(x = x_1, x' = x_n) &\equiv U(x_1, x_2) \cdot U(x_2, x_3) \cdots U(x_{n-1}, x_n) \\ &= (\mathbb{1} + igA_\mu(x_1)dx_\mu) \cdot (\mathbb{1} + igA_\mu(x_2)dx_\mu) \cdots (\mathbb{1} + igA_\mu(x_{n-1})dx_\mu) \\ &= \mathbf{P} \exp \left[ig \int_x^{x'} dy_\mu A_\mu(y) \right]. \end{aligned} \quad (3.6)$$

where \mathbf{P} denotes path ordering. QCD is a non-Abelian $SU(3)$ gauge theory with generators t^a . Under the gauge transformation

$$\begin{aligned} \psi(x) &\rightarrow e^{i\alpha^a(x)t^a} \psi(x) \\ \bar{\psi}(x) &\rightarrow \bar{\psi}(x) e^{-i\alpha^a(x)t^a} \\ U(x, x') &\rightarrow e^{i\alpha^a t^a(x)} U(x, x') e^{-i\alpha^a t^a(x')} \end{aligned} \quad (3.7)$$

terms of the form

$$\bar{\psi}(x_1) U(x_1, x_2) U(x_2, x_3) \cdots U(x_{N-1}, x_N) \psi(x_N) \quad (3.8)$$

are clearly gauge invariant.

Because the lattice has a finite, but small, spacing a , we may define the connection as

$$U_\mu(x) \equiv U(x, x + \hat{\mu}) = \exp(igA_\mu(x)) \quad (3.9)$$

where $\hat{\mu}$ is the vector of length a in the μ direction.² This connection defines the gauge field on the *link*, where it is associated with both a site and a direction. Note that the connection is defined in the positive μ direction only; to connect to a point in the negative μ direction,

²By absorbing the lattice spacing a into the ‘unit’ vector, we make an implicit decision to work in “lattice units”.

one must take the adjoint:

$$U_\mu^\dagger(x) \equiv U(x + \hat{\mu}, x) = \exp(-iagA_\mu(x)). \quad (3.10)$$

We can organize our approach to discretizing the QCD action by taking advantage of its gauge symmetry. According to gauge symmetry, all operators must be gauge invariant — all color indices must contract. In the continuum, the Wilson loop

$$U_P(y, y) = \exp \left[-ie \oint_P dx^\mu A_\mu(x) \right] \quad (3.11)$$

is a nontrivial function of $A_\mu(x)$ that is locally gauge invariant [18, §15.3] (for the non-Abelian case we must take the color-trace to obtain a gauge-invariant quantity). It is possible to write all gauge-invariant functions of $A_\mu(x)$ as combinations of Wilson loops for various choices of the path P . Our goal should be to find a lattice representation of the Wilson loop, which could then be used to construct higher-order lattice operators.

The smallest possible Wilson loop is obtained by tracing a path around a single *plaquette*, which we define as

$$P_{\mu\nu}(x) \equiv \frac{1}{3} \Re \text{Tr} \left(U_\mu(x) U_\nu(x + \hat{\mu}) U_\mu^\dagger(x + \hat{\nu}) U_\nu^\dagger(x) \right). \quad (3.12)$$

For $a \ll 1$ we can substitute Eq. (3.9) to find

$$\begin{aligned} P_{\mu\nu} &= \frac{1}{3} \Re \text{Tr} \mathbf{P} \exp \left[-i \oint_{\square} gA \cdot dx \right] \\ &= \frac{1}{3} \Re \text{Tr} \left[1 - ig \oint_{\square} A \cdot dx + \frac{1}{2} \left(ig \oint_{\square} A \cdot dx \right)^2 + \mathcal{O}(A^3) \right]. \end{aligned} \quad (3.13)$$

Because $a \ll 1$, we can assume A is smooth and use a four-dimensional version of Stokes' theorem to evaluate the integral as

$$\begin{aligned} \oint_{\square} A \cdot dx &= \int_0^a dx_\mu \int_0^a dx_\nu [\partial_\mu A_\nu(x_0 + x) - \partial_\nu A_\mu(x_0 + x)] \\ &= a^2 F_{\mu\nu}(x_0) + \frac{a^3}{2} (D_\mu + D_\nu) F_{\mu\nu}(x_0) \end{aligned}$$

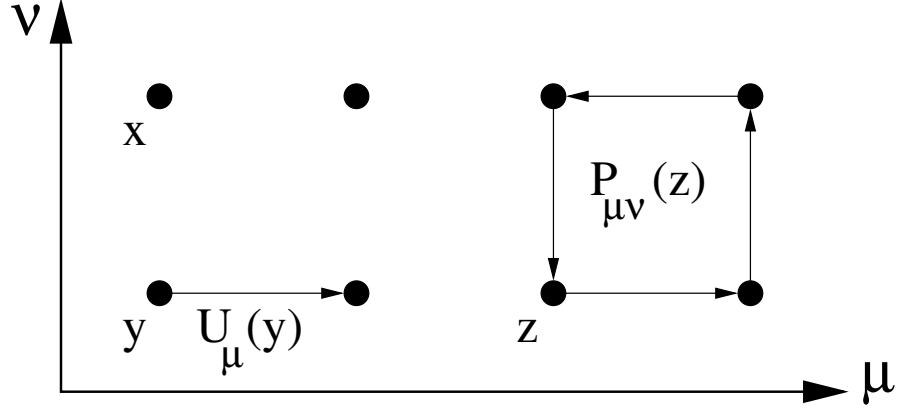


Figure 3.1: A small segment of the hypercubic lattice, showing a site x , a link $U_\mu(y)$, and a plaquette $P_{\mu\nu}(z)$.

$$+ \frac{1}{2} \left(\frac{a^4}{3} D_\mu^2 + 2 \frac{a^4}{4} D_\mu D_\nu + \frac{a^4}{3} D_\nu^2 \right) F_{\mu\nu}(x_0) + \mathcal{O}(a^5, A^2) \quad (3.14)$$

and therefore find the plaquette to be

$$P_{\mu\nu} = 1 - \frac{1}{6} a^4 \text{Tr}(gF_{\mu\nu}(x_0))^2 - \frac{1}{6} a^5 \text{Tr}(gF_{\mu\nu}(x_0)(D_\mu + D_\nu)gF_{\mu\nu}(x_0)) - \mathcal{O}(a^6, A^2) \quad (3.15)$$

The mathematical complexity lends itself well to a pictorial representation of the site, link, and plaquette, as seen in Figure 3.1. When constructing more complicated objects we will depend heavily on the pictorial notation.

Now that we have obtained a lattice representation of the Wilson loop, we can more easily attack the problem of discretizing the QCD action. Because the QCD action (Eq. (3.4)) is separable³ we can treat our discretization of the gauge and fermion parts separately.

3.2.1 Lattice Gauge Action

Wilson's gauge action is

$$S_g[P] = -\frac{1}{2g^2} \sum_p \left(1 - \frac{1}{2\text{Tr}_{\text{color}} \mathbb{1}} (\text{Tr}_{\text{color}} P_{\mu\nu}(x) + \text{Tr}_{\text{color}} P_{\mu\nu}^{-1}(x)) \right) \quad (3.16)$$

³The covariant derivative requires knowledge of the gauge field, but this is accounted for in our link variables.

where the sum only includes one orientation for each plaquette:

$$\sum'_p \equiv \sum_x \sum_{1 \leq \mu < \nu \leq 4} . \quad (3.17)$$

It is straightforward to show that

$$S_g[P] = -\frac{1}{4} \sum_x a^4 \text{Tr} (F_{\mu\nu}^a(x) F_{\mu\nu}^a(x)) + \mathcal{O}(a^2), \quad (3.18)$$

which reduces to the continuum gauge action in the limit $a \rightarrow 0$.

3.2.2 Naïve Fermion Action

A naïve discretization of the continuum fermion action leads to the lattice fermion action:

$$S_f^L = \sum_x a^4 \bar{\psi}(x) \left(\not{D}^L + m \right) \psi(x) + \mathcal{O}(a^2) \quad (3.19)$$

where D^L is the centered gauge-covariant lattice derivative

$$D_\mu^L \psi(x) \equiv \frac{U_\mu(x) \psi(x + \hat{\mu}) - U_\mu^\dagger(x - \hat{\mu}) \psi(x - \hat{\mu})}{2a}, \quad (3.20)$$

which is correct to $\mathcal{O}(a^2)$.

3.3 The Doubling Problem

Unfortunately, the naïve lattice fermion action suffers from an interesting, but serious, pathology. The propagator for a fermion of mass m is

$$C_0 = \left[m + \frac{i}{a} \gamma_\mu \sin(ap_\mu) \right]^{-1}. \quad (3.21)$$

Consider the $m \rightarrow 0$ limit. Taylor expanding $\sin(ap_\mu) \Rightarrow ap_\mu + \mathcal{O}((ap_\mu)^3)$ we see the continuum Dirac propagator $C_c = [i\not{p}]^{-1}$ is recovered. But the naïve lattice propagator has poles *both* at $ap_\mu = 0$ and at $ap_\mu = \pi$. The duplication of the fermion occurs in each space-time dimension

μ for a total of $2^4 = 16$ “tastes” of each fermion [19]. The doublers can also be understood as an extra symmetry under the transformation

$$\psi(x) \rightarrow \tilde{\psi}(x) \equiv i\gamma_5\gamma_\mu(-1)^{x_\mu}\psi(x) \quad (3.22)$$

$$= i\gamma_5\gamma_\mu \exp(ix_\mu\pi)\psi(x). \quad (3.23)$$

The so-called “doubling problem” is an unavoidable consequence of the Nielsen-Ninomiya theorem [20, 21], which states that there are an equal number of right- and left-handed particles for every set of quantum numbers [22]. In general, no lattice action can be chiral, undoubled, and ultralocal.⁴ One of the three must be sacrificed. Given three conditions, it is clear that there are three methods for dealing with the problem. Methods which sacrifice the ultralocal constraint (and instead use a local action) include “domain wall” and “overlap” fermions. This work considers the other two methods (sacrificing chiral symmetry, and allowing the doublers).

3.3.1 Wilson Fermions

Wilson proposed [19] to add an irrelevant term, of dimension five, to the fermion Lagrangian:

$$\mathcal{L}_W = \bar{\psi}_x \left(\not{D}^L + m - \frac{r}{2}a(D_\mu^2)^L \right) \psi_x \quad (3.24)$$

where r is an arbitrary value (usually we set $r = 1$ for simplicity) called the Wilson parameter.

The form of the action given in Eq. (3.24) is called the “mass form”. For reasons of computational efficiency, lattice calculations split the Wilson term:

$$-\frac{r}{2}a\bar{\psi}_x(D_\mu^2)^L\psi_x \equiv -\frac{r}{2a}\sum_{\mu=1}^4 \left[\bar{\psi}_{x+\hat{\mu}}U_\mu(x)\psi_x - 2\bar{\psi}_x\psi_x + \bar{\psi}_{x-\hat{\mu}}U_\mu^\dagger(x-\hat{\mu})\psi_x \right] \quad (3.25)$$

$$\rightarrow \frac{4r}{a}\bar{\psi}_x\psi_x - \frac{r}{2a}\sum_{\mu=1}^4 \left[\bar{\psi}_{x+\hat{\mu}}U_\mu(x)\psi_x + \bar{\psi}_{x-\hat{\mu}}U_\mu^\dagger(x-\hat{\mu})\psi_x \right]. \quad (3.26)$$

⁴Ultralocality means the effects fall off to exactly zero over a finite distance [23]. This is a tighter constraint than locality, which only requires effects to fall off exponentially.

We add these terms to the QED action to get the Wilson action

$$S_W = \sum_x a^4 \left\{ \bar{\psi}_x \left(m + \frac{4r}{a} \right) \psi_x - \frac{1}{2a} \sum_{\mu=\pm 1}^{\pm 4} \bar{\psi}_{x+\hat{\mu}} (r + \gamma_\mu) U_\mu(x) \psi_x \right\}. \quad (3.27)$$

In this form, we can clearly see that we have not evaded the Nielsen-Ninomiya theorem — the Wilson action leads to an additive mass renormalization, and therefore violates chiral symmetry.⁵ We can obtain a unit coefficient of the $\bar{\psi}\psi$ term if we absorb the current coefficient into the field definition:

$$a^{3/2} \sqrt{am + 4r} \psi \rightarrow \psi. \quad (3.32)$$

We then obtain

$$S_W = \sum_x \left\{ \bar{\psi}_x \psi_x - \kappa \sum_{\mu=\pm 1}^{\pm 4} \bar{\psi}_{x+\hat{\mu}} (r + \gamma_\mu) U_\mu(x) \psi_x \right\} \quad (3.33)$$

where $\kappa \equiv (2am + 8r)^{-1}$ is the “hopping parameter”. In the continuum limit, the lattice Lagrangian for massless quarks becomes

$$2\kappa \bar{\psi} \gamma_\mu \partial^\mu \psi + \mathcal{O}(a), \quad (3.34)$$

which tells us that $\sqrt{2\kappa}\psi$ is the lattice equivalent of the continuum gauge field — in the hopping-parameter form, the wavefunction rescaling is

$$\psi \rightarrow \frac{1}{\sqrt{2\kappa}} \psi. \quad (3.35)$$

⁵The chiral transformation is

$$\psi \rightarrow e^{\frac{i}{2}\gamma_5\theta} \psi. \quad (3.28)$$

Because $\{\gamma_5, \gamma_\mu\} = 0$ we find

$$\bar{\psi} \equiv \psi^\dagger \gamma_0 \rightarrow \left(e^{\frac{i}{2}\gamma_5\theta} \psi \right)^\dagger \gamma_0 = \psi^\dagger e^{-\frac{i}{2}\gamma_5\theta} \gamma_0 = \bar{\psi} e^{\frac{i}{2}\gamma_5\theta}. \quad (3.29)$$

The massless QCD Lagrangian $\bar{\psi} \not{D} \psi$ therefore transforms under a chiral transformation as

$$\bar{\psi} \not{D} \psi \rightarrow \bar{\psi} e^{\frac{i}{2}\gamma_5\theta} D_\mu \gamma_\mu e^{\frac{i}{2}\gamma_5\theta} \psi = \bar{\psi} e^{\frac{i}{2}\gamma_5\theta} e^{-\frac{i}{2}\gamma_5\theta} D_\mu \gamma_\mu \psi = \bar{\psi} \not{D} \psi, \quad (3.30)$$

i.e., it is invariant. The mass term, on the other hand, transforms as

$$\bar{\psi} m \psi \rightarrow \bar{\psi} e^{\frac{i}{2}\gamma_5\theta} m e^{\frac{i}{2}\gamma_5\theta} \psi = \bar{\psi} m e^{i\gamma_5\theta} \psi \quad (3.31)$$

and is not invariant.

At tree level, massless quarks correspond to $\kappa = \frac{1}{8r}$. Because of the additive mass renormalization, the value of κ will vary from its tree-level value. The “critical” value, κ_{crit} , is determined nonperturbatively as the value where the pion mass $m_\pi \rightarrow 0$. Bare masses can then be determined as offsets of κ_{crit} as

$$am_0 = \frac{1}{2\kappa} - \frac{1}{2\kappa_{\text{crit}}}. \quad (3.36)$$

With Wilson’s Lagrangian (and a choice of $(D_\mu^2)^L$ that includes one-hop differences only), the propagator becomes

$$C_W(t, \mathbf{p}) = \int_{-\pi}^{\pi} \frac{dp_0}{2\pi} \frac{e^{ip_0 t}}{m + \frac{i}{a}\gamma_\mu \sin(ap_\mu) + \frac{2r}{a^2} \sin^2(\frac{1}{2}ap_\mu)}. \quad (3.37)$$

In the $m \rightarrow 0$ limit we still recover the expected pole at $ap_\mu = 0$, but the additional pole is evaded through the $\sin^2(\frac{1}{2}ap_\mu)$ term. We therefore avoid the doubling problem at the cost of introducing $\mathcal{O}(a)$ errors into our action.

Energy-Momentum Relation

We use Wilson’s method for heavy quarks — it doesn’t work as well for light fermions because the Wilson term explicitly violates chiral symmetry. For heavy-quark physics, one may expand the continuum energy-momentum relation about the meson mass:

$$E = \sqrt{\mathbf{p}^2 + m^2} = m\sqrt{1 + (\mathbf{p}/m)^2} = m + \frac{\mathbf{p}^2}{2m} - \frac{\mathbf{p}^4}{8m^3} + \dots \quad (3.38)$$

The above expression is valid for non-relativistic mesons ($\mathbf{p}^2/m^2 \ll 1$), but assumes rotational symmetry. To get an expression that doesn’t depend on this assumption, and incorporates the rotational symmetry breaking of the lattice, it is helpful to take a Taylor expansion of $E(\mathbf{p})$:

$$E(\mathbf{p}) = E(\mathbf{0}) + \frac{1}{2!} \sum_{i,j} \frac{\partial^2 E}{\partial p_i \partial p_j} \Big|_{\mathbf{p}=\mathbf{0}} p_i p_j + \frac{1}{4!} \sum_{i,j,k,l} \frac{\partial^4 E}{\partial p_i \partial p_j \partial p_k \partial p_l} \Big|_{\mathbf{p}=\mathbf{0}} p_i p_j p_k p_l + \dots \quad (3.39)$$

Here the odd terms vanish since parity requires E to be an even function of the p_i .

Furthermore, we split the \mathbf{p}^4 term into two components, for pairwise contractions of the

indices, and for the case $i = j = k = l$. Consider first the $i = j = k = l$ case, which gives

$$\frac{1}{4!} \sum_i p_i^4 \left. \frac{\partial^4 E}{\partial p_i^4} \right|_{\mathbf{p}=\mathbf{0}}. \quad (3.40)$$

Now consider the pairwise case. There are 3 ways to contract the indices, leading to the expression

$$\frac{3}{4!} \sum_{i \neq j} p_i^2 p_j^2 \left. \frac{\partial^4 E}{\partial p_i^2 \partial p_j^2} \right|_{\mathbf{p}=\mathbf{0}}. \quad (3.41)$$

Splitting this term with the relation $\sum_{i \neq j} = \sum_{i,j} - \sum_{i=j}$ we obtain

$$\frac{3}{4!} \left[\sum_{i,j} p_i^2 p_j^2 - \sum_i p_i^4 \right] \left. \frac{\partial^4 E}{\partial p_i^2 \partial p_j^2} \right|_{\mathbf{p}=\mathbf{0}}. \quad (3.42)$$

Combining these cases together, and grouping terms by whether they sum over i only, or over both i and j , we find

$$\frac{3}{4!} \sum_{i,j} p_i^2 p_j^2 \left. \frac{\partial^4 E}{\partial p_i^2 \partial p_j^2} \right|_{\mathbf{p}=\mathbf{0}} + \frac{1}{4!} \sum_i p_i^4 \left[\left. \frac{\partial^4 E}{\partial p_i^4} \right|_{\mathbf{p}=\mathbf{0}} - 3 \left. \frac{\partial^4 E}{\partial p_i^2 \partial p_j^2} \right|_{\mathbf{p}=\mathbf{0}} \right]. \quad (3.43)$$

We recognize that $\sum_{i,j} p_i^2 p_j^2 = (\mathbf{p}^2)^2$, and write the dispersion relation in the recognizable form:

$$E(\mathbf{p}) = M_1 + \frac{\mathbf{p}^2}{2M_2} - \frac{(\mathbf{p}^2)^2}{8M_4^3} - \frac{1}{6} w_4 \sum_i p_i^4 + \dots \quad (3.44)$$

where the rest mass is

$$M_1 \equiv E(0) \quad (3.45)$$

and

$$\frac{1}{M_2} \equiv \left. \frac{\partial^2 E}{\partial p_i^2} \right|_{\mathbf{p}=\mathbf{0}} \quad (3.46)$$

defines the kinetic mass.⁶

⁶At tree level,

$$M_1 a = \ln(1 + m_0 a) \quad (3.47)$$

and

$$\frac{1}{M_2 a} = \frac{2}{m_0 a (2 + m_0 a)} + \frac{1}{1 + m_0 a}. \quad (3.48)$$

We also have defined another mass and introduced a term to parameterize the violation of rotational symmetry:

$$\frac{1}{M_4^3} \equiv - \left. \frac{\partial^4 E}{\partial p_i^2 \partial p_j^2} \right|_{\mathbf{p}=\mathbf{0}}, \quad i \neq j \quad (3.49)$$

$$w_4 \equiv -\frac{1}{4} \left. \frac{\partial^4 E}{\partial p_i^4} \right|_{\mathbf{p}=\mathbf{0}} - \frac{3}{4M_4^3}. \quad (3.50)$$

The three masses defined above (M_1 , M_2 , and M_4) are not necessarily equal on the lattice. The rest mass, M_1 , includes a non-physical offset; we use this mass only when computing mass splittings. We identify the kinetic mass, M_2 , as the relevant mass, and use it when tuning the parameters of our calculation. The mass difference $M_4 - M_2$ gives some $\mathcal{O}(a^2)$ discretization errors. This artifact, along with the rotational symmetry violations parameterized by w_4 , may be removed through the use of improved actions.

Fermilab Interpretation for Heavy Quarks

Although we would like to use the Wilson action for heavy quarks, one might worry that the relevant scale, am , could lead to large discretization effects for $m \gtrsim 1/a$. To study this possibility, it helps to calculate the Wilson quark propagator following Ref. [5]. To evaluate the integral in Eq. (3.37), first rationalize the denominator. Then, substitute $z = e^{ip_0 \text{sign}t}$, yielding a contour integral over the circle $|z| = 1$. Applying the residue theorem [24, §54], one finds

$$C(t, \mathbf{p}) = \mathcal{Z}_2 e^{-E|t|} \frac{\gamma_0 \text{sign}t \sinh Ea - i\boldsymbol{\gamma} \cdot \sin a\mathbf{p} + am_0 + 1 - \cosh Ea + 2 \left(\sin \frac{a\mathbf{p}}{2}\right)^2}{2 \sinh Ea} \quad (3.51)$$

where the residue is

$$\mathcal{Z}_2(\mathbf{p}) = \left(1 + m_0 a + 2 \left(\sin \frac{a\mathbf{p}}{2} \right)^2 \right)^{-1}. \quad (3.52)$$

The residue can be identified with a normalization of the fermion field. Ignoring the $\mathcal{O}(\mathbf{p}^2)$ term, one finds

$$\mathcal{Z}_2 = (1 + m_0 a)^{-1} = e^{-M_1 a}. \quad (3.53)$$

Therefore $\mathcal{Z}_2^{-1/2} \psi_x = e^{M_1 a/2} \psi_x$ has the canonical normalization. In hopping-parameter form,

the normalization absorbs the $\sqrt{2\kappa}$ of Eq. (3.35) and becomes $\sqrt{1-6r\kappa}\psi_x$

It has been shown [5] that taking M_2 as the physical mass and using this wavefunction normalization ensures there will be no $(am_0)^\gamma$ discretization effects for $\gamma > 0$.

3.3.2 Staggered Fermions

Susskind, along with Kogut, suggested [25] that we might reduce the number of degrees of freedom by using a single component of the Dirac spinor on each site of the lattice, rather than all four components. The simplest construction of this method is through comparison to the case of the naïve fermion action. We begin by making the association:

$$\psi(x) = \Omega(x)\chi(x) \quad \bar{\psi}(x) = \bar{\chi}(x)\Omega^\dagger(x) \quad (3.54)$$

where

$$\Omega(x) \equiv \prod_{\mu=0}^3 (\gamma_\mu)^{x_\mu} = \gamma_0^{x_0} \gamma_1^{x_1} \gamma_2^{x_2} \gamma_3^{x_3}. \quad (3.55)$$

Because the square of any gamma-matrix is $\mathbb{1}$, there are only 16 different Ω matrices. We can therefore rewrite our fermion Lagrangian in the form

$$\bar{\psi}(x)(\gamma \cdot D^L + m)\psi(x) = \bar{\chi}(x)(\alpha(x) \cdot D^L + m)\chi(x) \quad (3.56)$$

where

$$\alpha_\mu(x) \equiv \Omega^\dagger(x)\gamma_\mu\Omega(x \pm \hat{\mu}) = (-\mathbb{1})^{x_0+x_1+\dots+x_{\mu-1}} \quad (3.57)$$

defines a position-dependent change of basis in spinor space. Because $\alpha(x)$ is diagonal at each point x , the action written in this form is diagonal in spinor space. Therefore the four spinor components of the χ fields are decoupled! Furthermore, because $\alpha(x) = \pm\mathbb{1}$, these four components behave identically. We are therefore free to eliminate three of them from our calculation (thereby reducing the computational complexity by a factor of four). Rather than computing 16 tastes of fermions, each with four spinor components, we are left with 16 tastes, each with a single spinor component. The continuum, however, has four spinor components for each flavor. We therefore group the 16 tastes to form four tastes each with four spinor

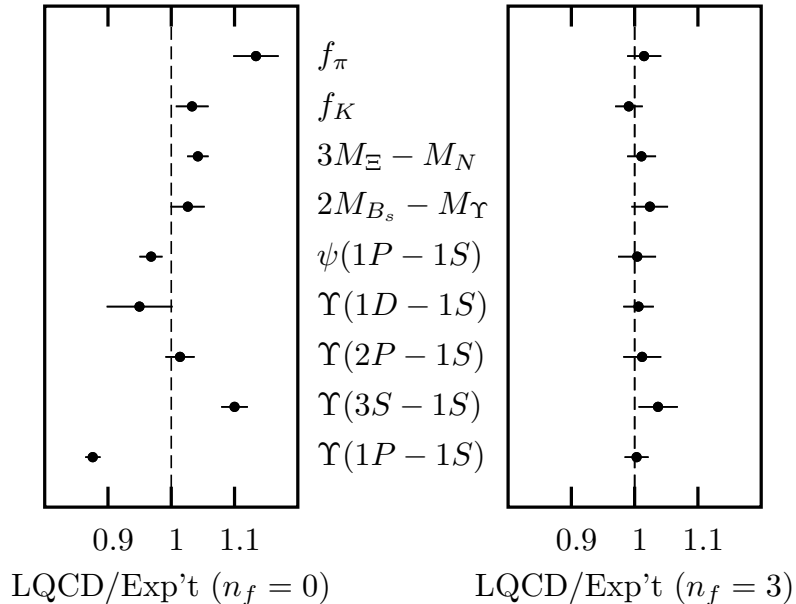


Figure 3.2: Effects of unquenching on several observables. The left panel results from quenched calculations, the right panel is unquenched. Details of the analysis are given in Ref. [32].

components.

Of course, even four tastes is too many, so an additional trick is required. The effects of the fermions are consolidated into the determinant of Eq. (3.109). Because the determinant of a matrix is the product of its eigenvalues, we can approximate the reduction of four tastes to one by taking the fourth-root of the fermion determinant (for the case of quarks with degenerate mass (the u and d) we simply take the square root of the determinant).⁷ Currently we do not have a mathematical proof that the fourth-root trick leaves us with an ultralocal, or even exponentially localized, action [30]. A greater concern is that any taste-symmetry violations in the action mean the reduction of tastes might not be well defined. However, the technique should be comparable to a theory with four flavors in which two have been quenched [31] so it is widely believed to be valid. This work is part of a community effort to test the fourth-root trick for some “gold-plated” quantities (see Section 3.6), and our results have been very encouraging (see Figure 3.2).

⁷In practice there are various ways of taking the root [26]: through the Lanczos method [27, §4.3], via a noisy Metropolis step [28], or by treating it as a polynomial [29].

Naïve Fermions

The staggered action is implemented on the computer in terms of the χ fields. Those fields are non-physical, however. When developing new operators, it is helpful to transition back to the ψ fields through the use of Eqs. (3.54) and (3.55). When this is done, the resulting fermions are similar to the naïve fermions of Section 3.2.2 (with the distinction that there is 1 taste rather than 16. To distinguish between the two methods of obtaining naïve fermions, we will continue to refer to the naïve fermions resulting from a staggered action as staggered fermions, but with the reminder that the staggering has been removed.

3.4 Improvement

Consider discretizing the covariant derivative

$$D_\mu \psi(x) \rightarrow D_\mu^L \psi(x) \equiv \frac{\psi_{x+a\hat{\mu}} - \psi_{x-a\hat{\mu}}}{2a} = \frac{U_\mu(x) - U_\mu^\dagger(x - \hat{\mu})}{2a} \psi(x) + \mathcal{O}(a^2). \quad (3.58)$$

It is evident that truncating an expansion in a (while keeping a finite) introduces discretization errors. Such errors will introduce lattice effects into our QCD action. These effects can be reduced by using the smallest lattice spacing possible, or studied by examining the calculated value of various observables as a function of a . More importantly, these effects can be reduced through a technique called *improvement*.

We can understand improvement by thinking of lattice QCD as an effective theory of continuum QCD. As an effective theory, lattice QCD is not correct at all energy scales, but rather has a built-in cutoff (a result of the lattice spacing, a). Like other effective theories, we can better approximate the correct theory by adding additional terms to the action of our effective theory. The operators we add correspond to interactions; their expectation values describe the long-distance physics. The coefficients of these terms indicate the strength of the new interactions, and therefore describe the short-distance physics. Finding which new operators to add, and their coefficients, is a complex task that is an area of active research. Symanzik proposed an improvement program [33] which provides a systematic method for doing this. The primary steps are:

1. determine all possible terms up to a given (engineering) dimension
2. eliminate redundant operators (terms which do not affect the physics)
3. determine coefficients for the remaining terms (usually from continuum field theory)

Each coefficient must be calculated using lattice perturbation theory. But such calculations are very difficult, so it is desirable to only do the ones that are absolutely necessary. As an alternative to this full-scale improvement, Lüscher and Weisz proposed [34, 35] in 1984 that we improve only on-shell quantities. This on-shell improvement provides a parameterization that will give correct results for experimentally-observable quantities. It will be instructive to take a brief tour of the improvements we make to construct the modern gauge, staggered, and Wilson actions we use in our analysis.

3.4.1 Symanzik Improvement

Even without improvement, lattice QCD has produced many useful results. However, it is limited in its accuracy due to lattice artifacts, the errors resulting from a finite lattice spacing. If we are to beat Moore’s law (the observation that computing power doubles roughly every 18 months), we must improve the actions themselves. This section details the construction of improved Wilson actions, following the approach of Sheikholeslami and Wohlert [36]. We begin with the determination of all possible terms at each order.

Before beginning, let us outline the requirements for each term. We know that the action is dimensionless so the Lagrangian must be of dimension four. In addition, it should satisfy some basic symmetries:

discrete rotational symmetry: all lorentz indices must contract⁸

gauge symmetry: all color indices must contract

positive under parity: the Lagrangian must be invariant under the transformation

$$\psi(x) \rightarrow \psi'(x) = S(\mathcal{P})\psi(\mathcal{P}x)$$

⁸We only require a discrete rotational symmetry, but can check that continuous rotational symmetry is restored in the continuum limit by comparing on-axis results to off-axis results

$$\begin{aligned}
\bar{\psi}(x) &\rightarrow \bar{\psi}'(x) = \bar{\psi}(\mathcal{P}x)S(\mathcal{P})^{-1} \\
D_\mu &\rightarrow D'_\mu = (\mathcal{P}^{-1})_{\mu\nu}D_\nu
\end{aligned}
\tag{3.59}$$

where $S(\mathcal{P}) = \gamma_\tau = S(\mathcal{P})^{-1}$ and $\mathcal{P}_{\mu\nu} = \mathcal{P}_{\mu\nu}^{-1} = -(-1)^{\delta_{\mu\tau}}\delta_{\mu\nu} = \text{diag}(-, -, -, +)$.

positive under charge conjugation: the Lagrangian must be invariant under the transformation

$$\begin{aligned}
\psi(x) &\rightarrow C\bar{\psi}^T(x) \\
\bar{\psi}(x) &\rightarrow \psi^T(x)C
\end{aligned}
\tag{3.60}$$

where $C\gamma_\mu^T C^{-1} = -\gamma_\mu$ and $C^{-1} = -C$.

We know from continuum QCD that $\psi(x)$ is of dimension 3/2, so the lowest-dimension operator is a bilinear of dimension 3; all higher-dimension operators are possible. Let our dimension 3 bilinear operator be written as

$$O(x) = \bar{\psi}(x)O^b\psi(x) \tag{3.61}$$

where $O^b = \Gamma$ must be chosen to satisfy gauge symmetry and be invariant under discrete rotations. Using the basis

$$\Gamma = \left\{ \mathbb{1}, \gamma_\mu, \sigma_{\mu\nu} \equiv \frac{i}{2}[\gamma_\mu, \gamma_\nu], \gamma_5\gamma_\mu, \gamma_5 \right\} \tag{3.62}$$

we can already restrict O^b to $\mathbb{1}$ and γ_5 . Parity requires the equality

$$\bar{\psi}(x)\Gamma\psi(x) = \bar{\psi}(\mathcal{P}x)\gamma_\tau\Gamma\gamma_\tau\psi(\mathcal{P}x) \tag{3.63}$$

so we must have $\Gamma = \gamma_\tau\Gamma\gamma_\tau$, which is satisfied for $\Gamma = \mathbb{1}$ but not for $\Gamma = \gamma_5$. Considering charge conjugation symmetry we find

$$\bar{\psi}(x)\psi(x) \rightarrow \psi^T(x)CC\bar{\psi}^T(x) = +(\bar{\psi}(x)\psi(x))^T = \bar{\psi}(x)\psi(x) \tag{3.64}$$

Table 3.1: Transformation of various operators under parity conjugation.

Operator	Transformed operator
$\mathbb{1}$	$\mathbb{1}$ (and similarly for $\delta_{\mu\nu}$)
γ_μ	$-(-1)^{\delta_{\mu\tau}} \gamma_\mu$
$\sigma_{\mu\nu}$	$(-1)^{\delta_{\mu\tau} + \delta_{\nu\tau}} \sigma_{\mu\nu}$
$\gamma_5 \gamma_\mu$	$(-1)^{\delta_{\mu\tau}} \gamma_5 \gamma_\mu$
γ_5	$-\gamma_5$

so this provides no further constraint. Therefore at dimension 3 the only valid operator is

$$O_0(x) = O_0^L(x) = \bar{\psi}(x)\psi(x). \quad (3.65)$$

Here I introduce the notation $O_i(x)$ to refer to the continuum operator, and $O_i^L(x)$ to refer to the discrete representation of that operator. In this case, they are identical.

At dimension 4 we consider operators of the form $O(x) = \bar{\psi}(x)O^b\psi(x)$ where $O^b = \Gamma_\mu D_\mu$ is the most general operator that satisfies gauge symmetry and discrete rotations. It will be helpful (now and for future use) to refer to Table 3.1 to see how various operators transform under parity. We see that parity restrictions allow only the case $\Gamma_\mu = \gamma_\mu$. Checking charge conjugation symmetry we find

$$\bar{\psi}\gamma_\mu D_\mu\psi = -\psi^T \gamma_\mu^T D_\mu \bar{\psi}^T = \bar{\psi}\gamma_\mu D_\mu\psi \quad (3.66)$$

so it is invariant. Therefore at dimension 4 the only valid operator is

$$O_1(x) = \bar{\psi}(x)\not{D}\psi(x) \quad (3.67)$$

$$O_1^L(x) = \bar{\psi}(x)\gamma_\mu D_\mu^L\psi(x) = \bar{\psi}(x) \sum_\mu \left(\gamma_\mu \frac{U_\mu(x)\psi(x + \hat{\mu}) - U_\mu^\dagger(x - \hat{\mu})\psi(x - \hat{\mu})}{2a} \right) \quad (3.68)$$

This process can be repeated at dimension 5, where we find the allowed bilinear operators

$$\Gamma_{\mu\nu} D_{\mu\nu} = \begin{cases} \delta_{\mu\nu} D_{\mu\nu} = D^2 \\ \sigma_{\mu\nu} D_{\mu\nu} = \sigma_{\mu\nu} F_{\mu\nu} \end{cases} . \quad (3.69)$$

For reasons that will become clear later (in Section 3.4.2) we choose as our basis for dimension

5 operators the linear combinations:

$$\begin{aligned} O_2(x) &= \bar{\psi}(x) \left(D^2 - \frac{1}{2} i \sigma_{\mu\nu} F_{\mu\nu} \right) \psi(x) \\ O_3(x) &= \frac{1}{2} i \bar{\psi}(x) \sigma_{\mu\nu} F_{\mu\nu} \psi(x). \end{aligned} \tag{3.70}$$

Improved through $\mathcal{O}(a)$, our fermion action is

$$S_L^{F_1} = -\frac{a^4}{g_0^2} \sum_x \sum_{i=0}^3 a^{\dim O_i^L(x)-4} b_i(g_0^2, ma) O_i^L(x) \tag{3.71}$$

where the coefficients $b_i(g_0^2, ma)$ have to be determined order by order in perturbation theory. Improvement through $\mathcal{O}(a^2)$ would require operators up to $O_{18}(x)$ (see Appendix F), rather than truncating the sum at operator $O_3(x)$.

3.4.2 On-Shell Improvement

Any time we make a change to the Lagrangian, we might expect it to affect the physics, so we might ask what effect the addition of new terms will have. We already know that some changes will not have any effect on physical observables. For example, rescaling a field by a constant leaves observables invariant. In general, any redefinition of a field that leaves the path integral invariant may be performed. These field redefinitions may induce additional terms in the Lagrangian. Any terms in the Lagrangian which can be obtained through a field redefinition can not effect the physics, and are considered *redundant*. Such terms can therefore be added to our Lagrangian with a coefficient chosen for convenience.

Consider, for example, the most general local covariant field redefinition up to $\mathcal{O}(a)$ that leaves the action invariant:

$$\begin{aligned} \psi &\rightarrow e^{\epsilon a(\not{D}+m)} \psi \\ \bar{\psi} &\rightarrow \bar{\psi} e^{\bar{\epsilon} a(\not{D}+m)} \end{aligned} \tag{3.72}$$

where $\epsilon, \bar{\epsilon}$ are free parameters and a is chosen so that $a(\not{D}+m)$ is small. Performing this transformation on the naïve lattice action Eq. (3.19) and then expanding (keeping terms to

$\mathcal{O}(\epsilon, \bar{\epsilon})$ yields

$$\begin{aligned}
\bar{\psi}(\not{D} + m)\psi &\rightarrow \bar{\psi}e^{\bar{\epsilon}a(\not{D}+m)}(\not{D} + m)e^{\epsilon a(\not{D}+m)}\psi \\
&= \bar{\psi}[1 + \bar{\epsilon}a(\not{D} + m)](\not{D} + m)[1 + \epsilon a(\not{D} + m)]\psi \\
&= \bar{\psi}(\not{D} + m + a(\epsilon + \bar{\epsilon})(\not{D} + m)^2 + \mathcal{O}(\epsilon^2))\psi \\
&= \bar{\psi} \left[(1 + am(\epsilon + \bar{\epsilon}))m + (1 + 2am(\epsilon + \bar{\epsilon}))\not{D} + a(\epsilon + \bar{\epsilon})\not{D}^2 \right] \psi. \quad (3.73)
\end{aligned}$$

Note that if we simply rescale the fields and mass under the transformation

$$\begin{aligned}
\sqrt{1 - 2am(\epsilon + \bar{\epsilon})}\psi &\rightarrow \psi' \\
\sqrt{1 - 2am(\epsilon + \bar{\epsilon})}\bar{\psi} &\rightarrow \bar{\psi}' \\
(1 + am(\epsilon + \bar{\epsilon}))m &\rightarrow m' \quad (3.74)
\end{aligned}$$

then Eq. (3.73) can be written as

$$m\bar{\psi}\psi + \bar{\psi}\not{D}\psi \rightarrow m'\bar{\psi}'\psi' + \bar{\psi}'\not{D}\psi' + a(\epsilon + \bar{\epsilon})\bar{\psi}'\not{D}^2\psi' \quad (3.75)$$

to terms of $\mathcal{O}(a)$. This new term is redundant, since it corresponds only to a field redefinition [36].

We may rewrite

$$\begin{aligned}
\not{D}^2 &= \gamma_\mu\gamma_\nu D_\mu D_\nu \\
&= \frac{1}{2}(\{\gamma_\mu, \gamma_\nu\} + [\gamma_\mu, \gamma_\nu])\frac{1}{2}(\{D_\mu, D_\nu\} + [D_\mu, D_\nu]) \\
&= \frac{1}{4}(2\delta_{\mu\nu} - 2i\sigma_{\mu\nu})(\{D_\mu, D_\nu\} + [D_\mu, D_\nu]) \\
&= \frac{1}{2}(2D_\mu^2 - i\sigma_{\mu\nu}F_{\mu\nu}) \\
&= D_\mu^2 - \frac{i}{2}\sigma_{\mu\nu}F_{\mu\nu} \equiv O_2 \quad (3.76)
\end{aligned}$$

which explains our earlier choice of basis for the operator O_2 in Eq. (3.70). Note that the action of Eq. (3.75) breaks chiral symmetry like the Wilson action, but does so with the

addition of an extra term. The $F_{\mu\nu}$ term can be approximated on the lattice as

$$\begin{aligned}
F_{\mu\nu}(x) &= \frac{1}{4} \left[U_\mu(x)U_\nu(x + \hat{\mu})U_\mu^\dagger(x + \hat{\nu})U_\nu^\dagger(x) \right. \\
&\quad + U_\mu^\dagger(x - \hat{\mu})U_\nu^\dagger(x - \hat{\mu} - \hat{\nu})U_\mu(x - \hat{\mu} - \hat{\nu})U_\nu(x - \hat{\nu}) \\
&\quad + U_\nu(x)U_\mu^\dagger(x - \hat{\mu} + \hat{\nu})U_\nu^\dagger(x - \hat{\mu})U_\mu(x - \hat{\mu}) \\
&\quad \left. + U_\nu^\dagger(x - \hat{\nu})U_\mu(x - \hat{\nu})U_\nu(x + \hat{\mu} - \hat{\nu})U_\mu^\dagger(x) \right] \tag{3.77}
\end{aligned}$$

$$= \begin{array}{cc} \overleftrightarrow{\square} & \overleftrightarrow{\square} \\ \overleftrightarrow{\square} & \overleftrightarrow{\square} \end{array} \tag{3.78}$$

The extension of the naïve fermion action (Eq. (3.19)) to dimension 5 has added both O_2 and O_3 . Because O_2 is redundant, its coefficient can be chosen to be any nonzero value to fix the doubling problem. The O_3 operator, however, is not redundant, and its coefficient must be determined perturbatively (by comparison of lattice and non-lattice short distance physics). One obtains the Sheikholeslami-Wohlert action

$$S_{\text{SW}} = \sum_x \left\{ \bar{\psi}_x \psi_x - \kappa \sum_{\mu=\pm 1}^{\pm 4} \bar{\psi}_{x+\hat{\mu}} (r + \gamma_\mu) \psi_x + \frac{i}{2} \kappa_{\text{CSW}} \sum_{\mu,\nu} \bar{\psi}_x \sigma_{\mu\nu} F_{\mu\nu}(x) \psi_x \right\} + \mathcal{O}(a^2). \tag{3.79}$$

This is commonly called the “clover” action because of the shape of the lattice representation of the $F_{\mu\nu}$ operator (see Eq. (3.78)).

We can now understand more clearly what has been accomplished with the Wilson action and the clover action. The Wilson term, $\bar{\psi} D^2 \psi$ added *part* of a redundant operator, O_2 . Because it added only part, it changed the physics (by removing the doublers). The removal of the doublers was, of course, intentional, but other changes to the physics, such as introducing $\mathcal{O}(a)$ errors, are undesirable. To remove those errors, we must include the other part of operator O_2 , which is accomplished through the clover term. Because Wilson’s discretization of D^2 is different from the Sheikholeslami-Wohlert discretization of the $\sigma_{\mu\nu} F_{\mu\nu}$ term, the inclusion of both halves of O_2 does not re-introduce the doublers.

This process can, of course, be repeated at higher order. This is an area of active research — a brief introduction is given in Appendix F.

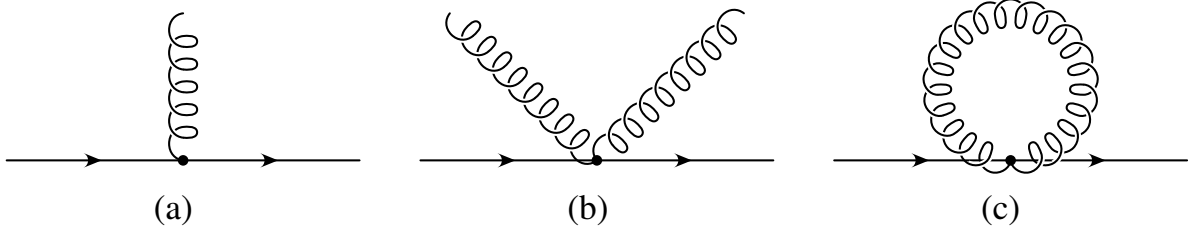


Figure 3.3: Lattice Feynman diagrams: (a) Fermion-fermion-gluon vertex; (b) Fermion-fermion-gluon-gluon vertex, unique to lattice QCD; (c) Tadpole diagram

3.4.3 Tadpole Improvement

One particularly interesting correction to the coefficients comes from *tadpole improvement* [37, 38]. In the continuum, the only interaction between quarks and gluons comes from the term $\bar{\psi} \not{D} \psi$, which corresponds to the diagram shown in Figure 3.3a. We map this onto the analogous lattice operator

$$U_\mu(x) \equiv e^{iagA_\mu(x)} \rightarrow 1 + iagA_\mu(x) - \frac{1}{2}a^2g^2A_\mu^2(x) + \mathcal{O}(a^3). \quad (3.80)$$

On the lattice, all higher-order operators are allowed including, for example, the term $\bar{\psi} \not{D}^2 \psi$, which corresponds to the diagram shown in Figure 3.3b. If the A_μ 's contract with each other, as shown in Figure 3.3c, they generate an ultraviolet divergence that precisely cancels the a^2 , leaving this term suppressed only by g^2 . The result is that the bare coupling α_{lat} is a poor choice for an expansion parameter. The corrections to a perturbative expansion will be large at any order.

Recall that our definition of the link variable did not take these large renormalizations into account. Our goal is to eliminate the undesirable ultraviolet divergence resulting from Figure 3.3c. To see how this can be accomplished, imagine if we could treat the ultraviolet (UV) and infrared (IR) parts separately. If one were to integrate out the UV modes, the link operator would reduce to its IR part

$$U_\mu \rightarrow u_0 e^{iagA_\mu^{\text{IR}}} \approx u_0(1 + iagA_\mu^{\text{IR}}) \quad (3.81)$$

where the Taylor expansion of the exponential now converges more rapidly. Here the parameter

$0 < u_0 < 1$ contains the averaged UV contribution, which will deviate from unity due to tadpole effects. We see, therefore, that we can cancel the UV tadpole contribution by simply dividing our link operator U_μ by u_0 . In practice, the value of u_0 is measured in a simulation using, for example, the gauge-invariant definition

$$u_0 \equiv \left\langle \frac{1}{3} \text{Tr} U_{\text{plaq}} \right\rangle^{1/4} \quad (3.82)$$

where we take advantage of the plaquette as a short-distance quantity to estimate the average link in a gauge-invariant way. Essentially, we have accomplished a nonperturbative resummation of the tadpoles.

When constructing a lattice action, it is helpful to think of this resummation by rewriting

$$U_\mu(x) \rightarrow u_0 \left[\frac{U_\mu(x)}{u_0} \right] \quad (3.83)$$

After we absorb the first factor u_0 into the couplings, the second factor $[U_\mu(x)/u_0]$ leads to an improved expansion parameter. The factors of u_0 absorbed into the couplings depend on the number of link variables in their corresponding operators. For example, the hopping parameter absorbs a single power of u_0 :

$$\kappa \rightarrow \tilde{\kappa} = u_0 \kappa \quad (3.84)$$

and the clover term coefficient absorbs three factors⁹ of u_0 :

$$c_{\text{SW}} \rightarrow \tilde{c}_{\text{SW}} = u_0^3 c_{\text{SW}}. \quad (3.85)$$

We obtain the tadpole-improved clover action:

$$S_{\text{SW}} = \sum_x \left\{ \bar{\psi}_x \psi_x - \tilde{\kappa} \sum_{\mu=\pm 1}^{\pm 4} \bar{\psi}_{x+\hat{\mu}} (r + \gamma_\mu) \psi_x + \frac{i}{2} \tilde{\kappa} \tilde{c}_{\text{SW}} \sum_{\mu,\nu} \bar{\psi}_x \sigma_{\mu\nu} F_{\mu\nu}(x) \psi_x \right\} + \mathcal{O}(\alpha_s a, a^2). \quad (3.86)$$

⁹The fourth factor of u_0 has already been absorbed into a factor of κ that also accompanies the clover term.

3.4.4 Classical Improvement of the Staggered Action: The Naik Term

Classical improvement demands that all corrections to the leading term $(\bar{\psi}\not{D}\psi + m\bar{\psi}\psi)$ vanish in the small a expansion to the desired order. Note that if we expand $\psi(x)$ in a Taylor series

$$\psi(x+a) = \psi(x) + a\psi'(x) + \frac{a^2}{2!}\psi''(x) + \frac{a^3}{3!}\psi'''(x) + \dots \quad (3.87)$$

we can expand a centered derivative as

$$\frac{\psi(x+a) - \psi(x-a)}{2a} = \psi'(x) + \frac{a^2}{6}\psi'''(x) + \dots \quad (3.88)$$

Because we are only considering operators of dimension 3 through 6 (and because the dimension 3 term is exact) we only need to improve the dimension 4 operator $O_1^L(x)$, which has the expansion

$$\begin{aligned} O_1^L(x) &\equiv \bar{\psi}(x)\not{D}^L\psi \\ &= \bar{\psi}(x)\gamma_\mu \left[D_\mu + \frac{1}{6}a^2 D_\mu^3 + \mathcal{O}(a^4) \right] \psi(x) \\ &= O_1(x) + \frac{a^2}{6}O_4(x) + \mathcal{O}(a^4). \end{aligned} \quad (3.89)$$

We must therefore get a better approximation of the continuum operator with the replacement $O_1^L(x) \rightarrow O_1^L(x) - \frac{a^2}{6}O_4^L(x)$. At $\mathcal{O}(a^2)$ the classically improved naïve action is

$$S^{\text{Cl}_2} = -\frac{a^4}{g_0^2} \sum_x \left[mO_0^L(x) + O_1^L(x) - \frac{1}{6}a^2O_4^L(x) \right] + \mathcal{O}(a^4) \quad (3.90)$$

where O_0^L , O_1^L , and O_4^L are defined in Eqs. (3.65), (3.68), and (F.7). The additional operator takes a derivative relative to a site three links away (in the same direction) and is called the Naik term [39].

3.4.5 Improving the Taste Symmetry of the Staggered Action

In the naïve action, there are 16 equivalent tastes for each quark flavor. One of these tastes is the expected low-energy mode, while the other 15 have high energies. Because they are

equivalent, however, there is the possibility of an additional interaction, whereby a high-energy gluon can change a quark from one taste to another. In particular, a low-energy quark can emit a gluon of momentum $\zeta\pi/a$, where ζ is a 4-vector with one or more components equal to 1 and all the others 0. (The vector ζ is used to identify one of the 15 doublers.) Rather than being driven far off-shell, it is transformed into a low-momentum mode of one of the other tastes. Of course, the gluon is highly virtual, and will soon be absorbed by another quark (causing a second taste change). This is a non-physical process, and therefore we seek to eliminate its effects from our simulations.

Because the taste-changing interactions involve gluon exchange, a simple method for reducing the effect is to suppress gluon momenta near $\zeta\pi/a$. This can be accomplished by “smearing out” the gauge fields [40, 41]:

$$A_\mu(x) \rightarrow A'_\mu(x) = \left(\prod_\mu \frac{D_{-\mu} + 2 + D_\mu}{4} \right) A_\mu(x). \quad (3.91)$$

In momentum space, this is equivalent to

$$A_\mu(k) \rightarrow A'_\mu(k) = \left(\prod_\mu \frac{1 + \cos k_\mu}{2} \right) A_\mu(k), \quad (3.92)$$

which approaches $A_\mu(k)$ as $k \rightarrow 0$ and vanishes when any component of k equals π .

The procedure can be formalized as follows Ref. [42]: we make the replacement

$$U_\mu(x) \rightarrow \left(1 + \sum_{\rho \neq \mu} \frac{a^2 \Delta_\rho^{(2)}}{4} \right) U_\mu(x) \quad (3.93)$$

where

$$\begin{aligned} \Delta_\rho^{(2)} &\equiv \frac{1}{u_0^2 a^2} \left[U_\rho(x) U_\mu(x + \hat{\rho}) U_\rho^\dagger(x + \hat{\mu}) - 2u_0^2 U_\mu(x) + U_\rho^\dagger(x - \hat{\rho}) U_\mu(x - \hat{\rho}) U_\rho(x - \hat{\rho} + \hat{\mu}) \right] \\ &= \frac{1}{u_0^2 a^2} \left[\begin{array}{c} \uparrow \rightarrow \downarrow \\ \uparrow \end{array} - 2u_0^2 \rightarrow + \begin{array}{c} \downarrow \leftarrow \\ \downarrow \end{array} \right] \end{aligned} \quad (3.94)$$

is a covariant second derivative that acts on link operators. This replacement suppresses

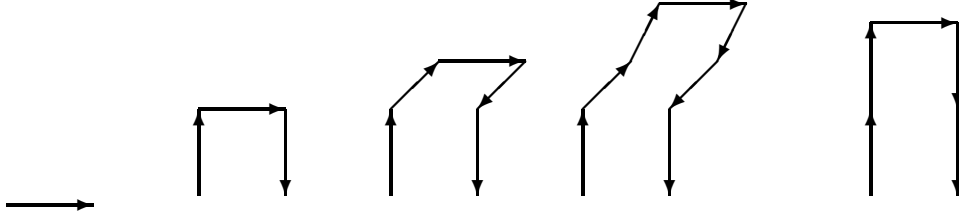


Figure 3.4: The simple link is shown along with the 3-, 5-, and 7-staple terms used to suppress taste-changing interactions. Also shown is the 5-link Lepage term used to suppress taste-conserving errors due to double taste change. The 3-link Naik term is not shown. Together with the coefficients of Table 3.2, these form the asqtad action shown in Eq. (3.97).

gluons with momentum $p_\rho = \pi/a$, for the three cases of $\rho \neq \mu$, by including “staples”.¹⁰

We can extend this to include all 15 doublers by multiplying the terms to include the combined effects of changing taste in multiple directions:

$$U_\mu(x) \rightarrow V_\mu(x) = \prod_{\rho \neq \mu} \left(1 + \frac{a^2 \Delta_\rho^{(2)}}{4} \right) \Big|_{\text{symm.}} U_\mu(x). \quad (3.95)$$

It should be noted that this equation adds 5-staples and 7-staples to our lattice action (see Figure 3.4).

We have one remaining issue. There is the possibility of a *taste-conserving* interaction, wherein the taste changes twice: first to a doubler mode, and then back. This leads to $\mathcal{O}(a^2 p^2)$ errors. Elimination of these effects requires one extra term:

$$V_\mu(x) \rightarrow V'_\mu(x) \equiv V_\mu(x) - \sum_{\rho \neq \mu} \frac{a^2 (\Delta_\rho)^2}{4} U_\mu(x). \quad (3.96)$$

The $(\Delta_\rho)^2$ is the square of the first derivative, leading to the Lepage term shown in Figure 3.4. The combination of the Naik term and the taste-symmetry restoring terms produces the a^2 tadpole-improved (*asqtad*) action

$$S_{\text{asqtad}} = \bar{\psi} \gamma_\mu \left(V'_\mu(x) - \frac{1}{6} (D_\mu^3)^L \right) \psi + \mathcal{O}(\alpha_s a^2, a^4). \quad (3.97)$$

Furthermore, tadpole improvement ensures the α_s coupling in the error is small. It is usually

¹⁰The MILC collaboration introduced “fat links” as the simplest change to the link fields, and discovered including them greatly reduces taste-symmetry violations [43]. Their success wasn’t explained until later.

Table 3.2: The coefficients we use for the asqtad derivative. The coefficients are chosen such that the coupling to a momentum-zero gluon is unity, while the coupling to gluons with momentum components $\pm\pi$ is zero. The factor of $(1/u_0)^{L-1}$, where L is the length of the path, accounts for the tadpole improvement (one power of u_0 has been absorbed into the quark mass). See Ref. [44] for additional details.

Term	Coefficient
Link	$\frac{1}{8} + \frac{3}{8} + \frac{1}{8}$
3-staple	$\frac{1}{8} \frac{1}{2} u_0^{-2}$
5-staple	$\frac{1}{8} \frac{1}{8} u_0^{-4}$
7-staple	$\frac{1}{8} \frac{1}{48} u_0^{-6}$
Lepage	$-\frac{1}{16} u_0^{-4}$
Naik	$-\frac{1}{24} u_0^{-2}$

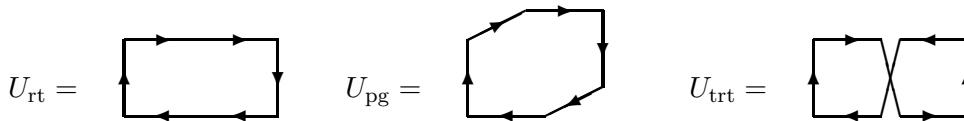


Figure 3.5: The rectangle, parallelogram, and twisted rectangle operators which can be used to improve the gauge action. Not shown are the bent rectangle and twisted bent rectangle. Coefficients for tree-level improvement are $c_{\text{pl}} = \frac{5}{3}$, $c_{\text{rt}} = -\frac{1}{12}$, $c_{\text{pg}} = 0$, and $c_{\text{trt}} = \frac{5}{36}$.

simplest to think of the asqtad action pictorially, as in Figure 3.4, in which case the terms have the coefficients listed in Table 3.2.

3.4.6 Improving the Gauge Action

Now that we have highly-improved fermion actions (the clover action improved to $\mathcal{O}(\alpha_s a, a^2)$ and the asqtad action improved to $\mathcal{O}(\alpha_s a^2, a^4)$), it is worthwhile to investigate improving our gauge action. The standard gauge action has errors of $\mathcal{O}(a^2)$. These can be removed by some combination of additional terms, following the procedure of Lüscher and Weisz [34, 35]. Choices of these terms include those shown in Figure 3.5. Only two of the terms are needed to improve the gauge action to $\mathcal{O}(\alpha_s a^2)$. For example, using rectangles gives a tree-level tadpole-improved gauge action of

$$S_G = -\beta \sum_{x, \mu > \nu} \left\{ \frac{5}{3} \frac{P_{\mu\nu}(x)}{u_\mu^2 u_\nu^2} - \frac{1}{12} \frac{R_{\mu\nu}(x)}{u_\mu^4 u_\nu^2} - \frac{1}{12} \frac{R_{\nu\mu}(x)}{u_\nu^4 u_\mu^2} \right\} + \mathcal{O}(\alpha_s a^2, a^4). \quad (3.98)$$

The action can be improved further by including some¹¹ one-loop effects in the coefficients [45]:¹²

$$c_{\text{rt}} \rightarrow -\frac{1}{12}(1 + 0.4805\alpha_s), \quad (3.99)$$

$$c_{\text{pg}} \rightarrow -\frac{3}{5}(0.03325\alpha_s). \quad (3.100)$$

Using the rectangle and parallelogram terms¹³ gives

$$S_G = \beta_{\text{pl}} \sum_{\text{pl}} \frac{1}{3} \Re \text{Tr} (1 - U_{\text{pl}}) + \beta_{\text{rt}} \sum_{\text{rt}} \frac{1}{3} \Re \text{Tr} (1 - U_{\text{rt}}) + \beta_{\text{pg}} \sum_{\text{pg}} \frac{1}{3} \Re \text{Tr} (1 - U_{\text{pg}}) + \mathcal{O}(\alpha_s a^2, a^4) \quad (3.101)$$

where U_{pl} and β_{pl} are the standard plaquette term and its coefficient, and

$$\beta_{\text{rt}} = -\frac{\beta_{\text{pl}}}{20u_0^2}(1 + 0.4805\alpha_s), \quad \beta_{\text{pg}} = -\frac{\beta_{\text{pl}}}{u_0^2}(0.03325\alpha_s), \quad \beta_{\text{trt}} = \frac{\beta_{\text{pl}}}{12u_0^4}. \quad (3.102)$$

Because the calculation did not take fermion loops into account, errors are still introduced at $\mathcal{O}(\alpha_s a^2, a^4)$, though some improvement over the tree-level tadpole-improved gauge action of Eq. (3.98) is expected.

3.4.7 Putting it All Together

Before continuing, let us summarize the discretization of the QCD action used in this work. Our action is divided into three parts, representing the light quarks, the heavy quarks, and the gauge fields. The light quarks use the asqtad action of Eq. (3.97), which has $\mathcal{O}(\alpha_s a^2, a^4)$ errors. Heavy quarks use tadpole-improved clover action of Eq. (3.86) in the Fermilab formalism, which has $\mathcal{O}(\alpha_s a, a^2)$ discretization errors. Appendix A estimates these errors through the use of a potential model calculation. The gauge fields use the tadpole-improved action of Eq. (3.101). This action accounts for some one-loop effects¹⁴, and therefore has $\mathcal{O}(\alpha_s a^2, a^4)$ errors.

¹¹The one-loop coefficients have only been calculated for the gauge field, not for fermion loops.

¹²The one-loop improvement of the twisted rectangle was not computed as that term is more sensitive to tadpole improvement than any of the others.

¹³The twisted rectangle is not used, as that term is more sensitive to tadpole improvement than the others.

¹⁴See footnote 11.

3.5 Calculating Observable Quantities

Once a consistent lattice field theory has been defined, we must implement it on a computer to extract information about physical observables. In doing so, we must confront the practical aspects of a computer implementation, and then design a method to extract the information we desire.

Recall that in Eq. (3.2) we defined QCD via a path integral. Wick rotating to Euclidean space and substituting $S[\phi] \rightarrow S_f[U, \bar{\psi}, \psi] + S_g[U]$, we find the lattice equivalent:

$$\langle 0 | \Phi(U, \bar{\psi}, \psi) | 0 \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}U \mathcal{D}\bar{\psi} \mathcal{D}\psi \Phi(U, \bar{\psi}, \psi) e^{-S_f[U, \bar{\psi}, \psi] - S_g[U]}. \quad (3.103)$$

Consider writing the fermion action in the matrix form

$$S_f = \sum_{x,y} \bar{\psi}_x M_{xy} \psi_y. \quad (3.104)$$

We represent fermions using Grassmann variables, which are described in Appendix B. In our current notation, Eq. (B.7) becomes

$$\int \mathcal{D}\bar{\psi} \mathcal{D}\psi e^{-\bar{\psi} \mathbf{M} \psi} = \det \mathbf{M}. \quad (3.105)$$

Integrating out the fermion degrees of freedom therefore gives

$$\langle 0 | \Phi(U, \bar{\psi}, \psi) | 0 \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}U \Phi[U] \det \mathbf{M}[U] e^{-S_g[U]}. \quad (3.106)$$

Due to the high dimensionality of the integral over $\mathcal{D}U$ we must use Monte Carlo methods. We generate a set of statistically independent gauge configurations, U_i . For each one, we calculate the integrand, and sum them to obtain a reliable estimate of the integral. The available space of the integral is very large, and only a small subset provides a significant contribution. If we were to randomly sample the entire space, the integral would not converge with an accessible number of gauge configurations. Through a technique known as *importance*

sampling, we choose the gauge configurations according to the probability distribution

$$P[U] \propto \det \mathbf{M}[U] e^{-S_g[U]}. \quad (3.107)$$

Importance sampling allows the Monte Carlo integral to converge significantly faster, which is critical to the viability of the method. Additional details of how the gauge configurations are generated according to Eq. (3.107) are given in Section 4.2.

Due to the large size of \mathbf{M} , which has rank $12L^4$ (4 spins \times 3 colors \times L^4 lattice sites), calculating the determinant quickly becomes prohibitively expensive as the size L of the lattice increases. It is therefore tempting to make the approximation $\det \mathbf{M} = 1$. This approximation, called the *quenched approximation*, effectively ignores the effects of sea quarks by preventing gluons from splitting into quark-antiquark pairs. In previous calculations, quenching was required due to computational limitations. However, the approximation is an uncontrolled one (it cannot be systematically improved). It was observed that quenched calculations has $\sim 20\%$ errors when compared to experiment. This work does not use the quenched approximation — we use the techniques described in Section 3.4 and gauge configurations produced using advanced algorithms [28, 46] to speed up the calculation. Additionally, advances in computing following Moore’s Law [47] have allowed us to dramatically increase our computational resources.

For the case of the quark propagator, we must use Eq. (B.8) instead:

$$\int \mathcal{D}\bar{\psi} \mathcal{D}\psi \bar{\psi} \psi e^{-\bar{\psi} \mathbf{M} \psi} = \mathbf{M}^{-1} \det \mathbf{M}. \quad (3.108)$$

The quark propagator is then

$$\langle \bar{\psi}(x) \psi(y) \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}U (\det \mathbf{M}) M_{xy}^{-1}[U] e^{-S_g[U]}. \quad (3.109)$$

Ignoring the difficulties of computing the fermion determinant, our problem is reduced to the form $\boldsymbol{\xi} = \mathbf{M}^{-1} \boldsymbol{\eta}$, or simply the matrix equation $\mathbf{M} \boldsymbol{\xi} = \boldsymbol{\eta}$. Directly taking the inverse would require $\mathcal{O}(N^3)$ operations, but our matrix, by virtue of the fact that our action is ultralocal,¹⁵

¹⁵Ultralocality means the action falls off to exactly zero after a finite distance. This is a stronger requirement

Table 3.3: Gamma-matrix structure for various quantum states.

State	J^{PC}	Operator
scalar	0^{++}	$\bar{\psi}\psi'$ or $\bar{\psi}\gamma_0\psi'$
pseudoscalar	0^{-+}	$\bar{\psi}\gamma_5\psi'$ or $\bar{\psi}\gamma_0\gamma_5\psi'$
vector	1^{--}	$\bar{\psi}\gamma_\mu\psi'$ or $\bar{\psi}\gamma_\mu\gamma_0\psi'$
axial-vector	1^{+-}	$\bar{\psi}\gamma_\mu\gamma_5\psi'$

is sparse. Therefore iterative matrix inversion methods become much more efficient. We use the minimal residual algorithm described in Appendix C.2.

Once we have constructed the quark propagator we can begin to construct hadrons. This is done by building operators that contain the proper quantum numbers. A state with $J = 0$ or 1 can be built using the structure

$$\chi(x) = \bar{\psi}(x)\Gamma\psi'(x), \quad (3.110)$$

where ψ and ψ' could represent different quark flavors. Table 3.3 lists some possible states.

3.5.1 Hadron Spectra

Finally, we are prepared to extract physical quantities. In general, the hadron propagators we build will contain information not only of the ground state, but also of all excited states that satisfy the specified quantum numbers and quark content. It is therefore necessary to extract only the information wanted [48]. Consider the example of extracting the η_c mass. We may write the correlation function as

$$\langle G_{\eta_c}(x, y) \rangle = \langle \mathbf{T} \chi_{\eta_c}(x) \chi_{\eta_c}^\dagger(y) \rangle \quad (3.111)$$

where \mathbf{T} denotes time ordering. We can write the η_c propagator in terms of quark propagators as

$$\langle G_{\eta_c}(x, y) \rangle = \frac{1}{Z} \int \mathcal{D}U \det \mathbf{M}[U] M_{yx}^{-1}[U] \gamma^5 M_{xy}^{-1}[U] \gamma^5 e^{-S_g[U]}. \quad (3.112)$$

than locality, which only requires exponential falloff.

In general, any hadron propagator can be written as a product of quark propagators in a gauge-field background. If we Fourier transform to momentum space and project out a specific 3-momentum \mathbf{p} , we can show that

$$\begin{aligned} G_{\eta_c}(\mathbf{p}, t) &= \sum_{\mathbf{x}} e^{i\mathbf{p}\cdot\mathbf{x}} \langle G_{\eta_c}(x, 0) \rangle \\ &= \sum_n Z_n^2 e^{-E_n t}, \end{aligned} \tag{3.113}$$

where $Z_n = \langle 0 | \chi_{\eta_c} | n \rangle$ is the matrix element giving the overlap of the operator $\chi_{\eta_c}(x)$ with state n . Here the sum is over all states n that have the same quantum numbers as the hadron in question (in this case, the η_c). At large times t , the higher energy states will be exponentially suppressed, so the ground state will dominate. The rate of exponential falloff of $G_{\eta_c}(\mathbf{p} = \mathbf{0}, t)$ at large times therefore gives us $E_0 = am_{\eta_c}$, the η_c mass in lattice-spacing units.

Excited states can be found as well, by fitting a correlation function to multiple states and projecting out the desired state [49, 50]. Because the propagator is dominated by the ground state exponential, the determination of excited states requires more configurations in the Monte Carlo calculation for statistical significance. The above method can also be extended to the use of three-point functions.

The above assumes that the two quarks in the meson are both generated using the clover action for heavy quarks. When we create mesons that include a light quark, we use the naïve action for the light quark. As explained in Section 3.3, the naïve action suffers from the doubling problem. The result is 16 different tastes of a fermion, and the possibility of taste-changing interactions. This taste symmetry is $\psi \rightarrow (i\gamma_\mu \gamma_5)(-1)^\mu \psi$. The spatial components lead to mesons which are much higher in energy, and therefore can be ignored. When the temporal component is applied to the light quark of a meson we find a second meson:

$$\Psi_h \gamma_5 \psi \rightarrow \Psi_h \gamma_5 (i\gamma_0 \gamma_5)(-1)^t \psi = i\Psi_h \gamma_0 \psi (-1)^t. \tag{3.114}$$

In the end, this means that each meson involving a light quark is accompanied by its parity

partner, which appears as an oscillating state in the meson propagator:

$$G_{\eta_c}(\mathbf{p}, t) = \sum_n Z_n^2 e^{-E_n t} + \sum_m \mathcal{Z}_m^2 (-1)^t e^{-\mathcal{E}_m t} \quad (3.115)$$

where I have used scripted letters to indicate the parity partner energy and matrix element. The spatial doublers don't contribute, however, as such a state would have high physical momentum, and therefore be very energetic. The contribution from such states enters the meson correlation functions with an energy splitting $\Delta E \sim \mathcal{O}(1/(Ma^2))$ [51, Appendix A], so we do not need to take them into account. They could also affect low-energy states through loops, but the effect is comparable to the $\mathcal{O}(\alpha_s a^2)$ error of the asqtad action, and can also be ignored.

3.5.2 Decay Constants

In the case of the D_s or D meson, there is the possibility of leptonic decay, where the quarks annihilate to a W which subsequently decays to leptons (see Figure 2.6). This decay is determined, in part, by the relevant CKM and hadronic matrix element (V_{cs} for D_s decay; V_{cd} for D decay). We would therefore like to determine the decay constants, f_{D_s} and f_D .

To extract the decay constant, we annihilate the meson propagator with an axial-vector current, as in the correlation function:

$$\langle G_{A_4-\eta_c}(x, y) \rangle = \langle \mathbf{T} \chi_{A_4}(x) \chi_{\eta_c}^\dagger(y) \rangle. \quad (3.116)$$

The A_4 current is the lattice representation of a coupling to the axial-vector W boson. The resulting propagator describes a particle with the energy of the meson, but with a coefficient related to the decay constant. We then fit

$$G_{A_4-\eta_c}(\mathbf{p}, t) = \sum_n Z_{A_4}^{(n)} Z_{D_s}^{(n)} e^{-E_n t}. \quad (3.117)$$

Here $Z_{A_4} = \langle 0 | A_4 | D_s \rangle$ and $Z_{D_s} = \langle D_s | \bar{s} \gamma_5 c | 0 \rangle$. Using our knowledge of the Z_{D_s} coefficient from Eq. (3.113), we can isolate Z_{A_4} , which is proportional to the decay constant, f_{D_s} .

3.6 Gold-Plated Quantities

Some quantities are “easy” to compute on the lattice. In particular, lattice QCD should be able to provide accurate predictions for heavy quark physics, because the non-relativistic quarks can be rapidly calculated. The Fermilab and UKQCD collaborations are using different methods to simulate the heavy quarks, which provides an internal cross-check on any results. In addition to the spectra for charmonium and bottomonium, we should be able to calculate the B and D meson decay constants, mixing amplitudes, and semileptonic form factors. These quantities test the validity of the fourth-root trick used by the staggered action to simulate light quarks. In addition, they test our ability to perform chiral extrapolations. We refer to this set of quantities as “gold-plated”, as they are the easiest, and most important, quantities. Such quantities provide a critical first test of lattice QCD — demonstration of our ability to successfully predict these quantities is a requirement if we are to have any confidence in future lattice predictions. This work studies charmonium and the D_s and D systems as a proving-ground of lattice techniques. We can combine our results with those of others who study the b -quark systems to obtain a more general test [32].

Chapter 4

Methodology

4.1 Computational Resources

The branch of computational physics lies between theory and experiment. Our measurements come from the computers described in this section.

4.1.1 MILC Resources

Our collaboration works closely with the MIMD¹ Lattice Computation (MILC) collaboration. They have extensive parallel programming experience and resources at many large computing centers, including NCSA. As a result, they provide gauge configurations to a large subset of the lattice community. This is because generation of gauge configurations takes an immense amount of processor time [52, 53]. These gauge configurations must then be analyzed for the particular physics that interests each group, typically on smaller clusters of computers. Parameters of the lattice gauge configurations generated by the MILC collaboration are given in Table 4.1.

4.1.2 FNAL Resources

The Fermilab lattice gauge theory group consists of several researchers at Fermilab and smaller groups at UIUC and other universities. Funding through DOE grants, the SciDAC² program, and other sources has allowed us to purchase several computing clusters to use in the analysis

¹Multiple Instruction/Multiple Data, a type of computing where many processors work on different sets of data in parallel

²Scientific Discovery through Advanced Computing

Table 4.1: Parameters of the lattices generated by the MILC collaboration. This analysis uses the “coarse” ensembles.

Name	a (fm)	Lattice Size	β $10/g^2$	$am_{u,d}/am_s$	κ_{crit}	u_0	# of configs (analyzed)
xcoarse	0.172	$16^3 \times 48$	6.503	0.0492/0.082			401
		$16^3 \times 48$	6.485	0.0328/0.082			331
		$16^3 \times 48$	6.467	0.0164/0.082			645
		$16^3 \times 48$	6.458	0.0082/0.0820			400
coarse	0.123	$20^3 \times 64$	6.85	0.050/0.05	0.13765	0.8707	493
		$20^3 \times 64$	6.81	0.030/0.05	0.13772	0.8696	564 (549)
		$20^3 \times 64$	6.79	0.020/0.05	0.13781	0.8688	484 (460)
		$20^3 \times 64$	6.76	0.010/0.05	0.13788	0.8677	658 (593)
		$20^3 \times 64$	6.76	0.007/0.05			493 (403)
		$24^3 \times 64$	6.76	0.005/0.05			197 (136)
fine	0.089	$28^3 \times 96$	7.11	0.0124/0.031	0.13717	0.8788	527 (261)
		$28^3 \times 96$	7.09	0.0062/0.031	0.13720	0.8782	592 (472)

of the MILC gauge configurations.

ACPMAPS

The first major undertaking, in 1989, was ACPMAPS.³ This was a home-built supercomputer designed and constructed at Fermilab. The foundation was the Intel i860 processor [55]. In total, the machine contained several hundred processors. In addition, a parallel processing language, Canopy, was written as a layer on top of C to take advantage of the hardware. This machine was decommissioned in 2003, when it was deemed no longer worth the space and power it consumed, having been obsoleted by PC clusters.

QCD80 Cluster

Beginning in late 1993, the concept of a beowulf cluster was introduced. In essence, the goal was to combine Commodity Off-The-Shelf (COTS) computers and network them together. It was hoped that this would provide a cheap alternative to groups that require the power of a supercomputer, but don’t have the funds to purchase one. Over the next few years, the idea gained support from the computing community, and matured to the point that it has replaced supercomputers for applications that scale well with limited communication. (Supercomputers

³The Fermilab Advanced Computer Program MultiArray Processor System [54].

are still required for shared-memory applications.)

After experimenting with an 8-node cluster in 1999, we obtained an 80-node cluster in 2001. The machines were dual Pentium 3 processors running at 700MHz. They were interconnected via Myrinet, a low-latency transport designed by Myricom specifically for clustered computing.

LQCD Clusters

The success of the 80-node cluster encouraged us to purchase a 128-node cluster in late 2002. This cluster contained dual Pentium 4 processors running at 2.0GHz on a 400MHz FSB. An upgrade in 2003 added another 128 nodes of dual P4's, these running at 2.4GHz. We have just purchased a 256-node cluster of single-processor P4's, each with 1GB of RAM. These machines are currently being added to the cluster using an Infiniband interconnect.

Other Projects

Despite the successes of beowulf clusters, it is important to note that specialty machines are still being developed. In the lattice community, the most prominent example is the QCD On a Chip (QCDOC) project [56, 57]. Researchers at Columbia University are working with IBM to design a small processing unit that can be interconnected on a multidimensional lattice. It is expected that this system will produce extremely fast computations for lattice QCD, and could also be used for other problems that are easy to define on multidimensional lattices. The primary advantage of this approach is speed; the disadvantages are the development time and the inability to upgrade as newer technologies become available.

4.2 Generating the Gauge Configurations

The gauge fields are generated first, with probabilities determined by Eq. (3.107). Generating gauge fields with that distribution is extremely time-consuming, because it is a nontrivial task to find a gauge field that produces a minimum of the action. For a practical calculation, we require several hundred such gauge fields. We use hybrid molecular-dynamics algorithms [28, 46] to generate the gauge fields. These algorithms work by approximating the fermions as random noise. The fermion fields are evolved through several timesteps using a molecular

dynamics algorithm (which is similar to a random walk, except that it allows a more rapid movement through configuration space by increasing the likelihood of multiple steps in the same direction). Then they are subjected to a heat bath step, which drives them towards the minimum. By alternating between molecular dynamics steps and heat bath steps, we can move towards, and then explore, the configuration space of gauge fields that minimizes the action.

The simulation may be started from a “cold start” where all fields are set to unity, or from a “warm start” where they are set randomly. Either way, the fields must be evolved for several timesteps (a period called “warming”) to bring them into the region that minimizes the action. Once there, we can start selecting gauge-field configurations for use in our Monte Carlo integral. Because each configuration is generated by modifying the previous one the resulting configurations are highly correlated. An analysis of the correlation length indicates that we should only save every sixth configuration that is generated. (We will later perform a more careful analysis of the correlations to ensure they do not influence our results.) A simulation will be run for a particular value of the lattice size, gauge coupling, and sea quark masses until we have collected ~ 500 gauge configurations. Parameters of the gauge configurations produced by the MILC collaboration are listed in Table 4.1.

4.3 Obtaining Physical Results

Section 3.5 showed how we compute the lattice quark propagators, and combine them together to form hadrons. Once we have our meson propagators we must extract physical results. This is done by recognizing that a propagator falls off exponentially, with the exponential determined by the energy of the meson. Recall that the our meson propagators only specify the quark content and whether the meson is a vector, pseudoscalar, etc., *not* the specific energy level. The meson propagators, therefore, contain all possible energy levels (cf. Eq. (3.113)):

$$G(t) = \sum_{n=0}^{\infty} Z_n^2 e^{-E_n t}. \quad (4.1)$$

Of course, on the lattice we have periodic boundary conditions, so this must be generalized to

$$G(t) = \sum_{n=0}^{\infty} Z_n^2 \left(e^{-E_n t} + e^{-E_n(T-t)} \right). \quad (4.2)$$

When attempting to extract information about a specific state, it would be helpful to enhance the wavefunction contributing to that state. To do so, we use a *smearing* of the quark content over several lattice sites, weighted by the expected wavefunction of the state, as predicted by potential models. The trivial smearing is not a smearing at all, but rather a local operator implemented by the δ -function. That operator has no preference for any particular state — all states are included with approximately equal weight. If we wish to select a $1S$ state, we can do so with the smearing

$$\psi_{1S}(x, t) = \sum_{x, t} P_{1S}(x, t) \psi(x, t) \quad (4.3)$$

where $P_{1S}(x, t)$ is the $1S$ wavefunction shown in Figure A.2. This will enhance the contribution from the $1S$ state, while suppressing the contribution from other states [58]. We can use a similar procedure to enhance the $2S$ state. In this work we did not use a $1P$ smearing — such a smearing is being developed for future analyses. In a given analysis, we might use the δ , $1S$, and $2S$ smearings to enhance the signal-to-noise ratio for the state we're trying to observe. In the most general case, we can use different smearings for the source and sink (the two ends of the hadron propagator). Therefore, we modify the form of our propagator to:

$$G_{\text{src,snk}}(t) = \sum_n Z_{\text{src}}^{(n)} Z_{\text{snk}}^{(n)} \left(e^{-E_n t} + e^{-E_n(T-t)} \right). \quad (4.4)$$

For the case of heavy-light mesons, we must consider the oscillating states in our fits, as explained in Section 3.5.1:

$$G_{\text{src,snk}}(t) = \sum_n Z_{\text{src}}^{(n)} Z_{\text{snk}}^{(n)} (-1)^{n(t+1)} \left(e^{-E_n t} + e^{-E_n(T-t)} \right). \quad (4.5)$$

Our task is now to use this model function to fit to the data we have obtained for the meson propagator. The E_n of the fit will correspond to meson energies (masses), while the

Z_n contain information about matrix elements. The remaining sections of this chapter will review fitting procedures, including multidimensional fitting and constrained curvefitting. The chapter will conclude with a summary of our planned analysis. Details of the analysis will be given in Chapter 5.

4.4 Fitting to Non-Linear Multi-Dimensional Functions

The model function to which we fit the meson correlators is both non-linear and multi-dimensional, making fitting difficult. A review of multi-dimensional fitting algorithms is given in Appendix C.3; the remainder of this section will focus on changes we made to the Marquardt-Levenberg method.

The defining characteristic of the Marquardt-Levenberg method is that it introduces a dimensionless parameter, λ , that carries us smoothly from the region of linear convergence ($\lambda \gg 1$) to a region of quadratic convergence ($\lambda \ll 1$). The astute reader will notice that the scale parameter λ is dimensionless, and therefore does not have a natural scale. In particular, the optimal value of λ may differ between the variables parametrizing our fit. To remedy the situation, we can convert λ to a vector, $\boldsymbol{\lambda}$. Rather than adjusting all of the λ_i in the same way, we can increase/decrease them according to the relative weight of how much their respective parameter had varied for that fit step. In this way, a parameter j that changes significantly will have a greater impact on its λ_j than a different parameter k that changed by an insignificant amount during that minimization step. My weighting for a parameter change Δ is $\Delta/\sqrt{\Delta \cdot \Delta}$. Using this weighting allows the fit to converge much faster (in fewer steps) than the standard method of keeping all λ_i equal. The improvement appears to be related to the number of parameters in the fit: for 6 parameters the minimizer required 1/6 as many steps.

Normally, one assumes convergence when the χ^2 stops changing. In addition to requiring the χ^2 to stop changing, we introduce an additional fit requirement, $\lambda \ll 1$. This ensures that we are in the region of quadratic convergence, and did not simply take a step that did not change the χ^2 . We also require the magnitude of the gradient at the minimum to be small. While this is not required to obtain a result within the Hessian errors of the method, it is

helpful to obtain more precise minima for use in a bootstrap analysis.⁴

4.5 Constrained Curvefitting

In our case, we don't know how many energy levels are needed in the fit. The required balance between a general model function without introducing too much freedom makes it difficult to obtain good fits.

In the past, it was commonplace to vary the minimum and maximum timeslices of the data, and the number of energy levels in the model function, and look for a region of stability in the results. This was a tricky procedure, because of the following constraints:

- N_{states} must be large enough to include excited states in the data, but not large enough to allow the fit parameters to wander to unphysical values
- t_{min} must be large enough to exclude excited states, but not reduce the signal
- t_{max} must be small enough to exclude the noise, but not reduce the signal

This work overcomes these limitations by using constrained curvefitting.

Traditional fitting methods utilize a few parameters that are allowed to take on any value. This can cause problems when the data does not sufficiently constrain the parameters to physical values. As an extreme example, consider the case where we expect our data to fall on the sum of two lines:

$$y = (ax + b) + (a'x + b') \tag{4.6}$$

It is easy to see that the data can only constrain the sums $a + a'$ and $b + b'$, and not the individual parameters. A fitter would get lost trying to find the “correct” solution. While this is an extreme case, it is not significantly different from a case where the model function may be analytically well defined (no redundant directions) but have redundant directions to machine precision. Guarding against such runaway solutions is therefore a general problem.

Because we have some prior knowledge of what the physical values can be, it is possible to guide the fitter to the correct solution. For example, consider the case where we expect that

⁴The bootstrap analysis is explained in Section 5.1.2.

some set of parameters a_n will be centered about \tilde{a}_n with width $\tilde{\sigma}_{a_n}$. We could then augment our χ^2 function to contain our prior knowledge:

$$\chi^2 \rightarrow \chi_{\text{aug}}^2 \equiv \chi^2 + \chi_{\text{prior}}^2 \quad (4.7)$$

where

$$\chi_{\text{prior}}^2 \equiv \sum_n \frac{(a_n - \tilde{a}_n)^2}{\tilde{\sigma}_{a_n}^2}. \quad (4.8)$$

The additions to the χ^2 cause the fit to favor values with $a_n = \tilde{a}_n \pm \tilde{\sigma}_{a_n}$.

Our current goal is to fit a meson correlator to a sum of several states. If we use a constrained fitting technique, the data will provide strong constraints over the lowest states, while the priors will prevent the fitter from getting lost in the parameter-space allowed by the highly-excited states. In this way, we can perform successful fits regardless of the quality of the data, or of the number of states we allow.

One can think of traditional fitting (without priors) as a special case of constrained fitting. Unconstrained fitting is equivalent to having infinitely wide priors for the first few parameters and infinitely narrow priors for all other parameters. In our case, it would correspond to performing a fit to only one or two energy levels, and setting all higher excitations to exactly zero. Because one cannot justify setting infinitely narrow priors for some parameters, traditional fitting has an infinite systematic error. The use of constrained fitting therefore helps us keep our systematic errors under control [59, 60].

4.5.1 Testing the Priors

In the case of our model function, it is easy to see that the first few energy levels are likely to be well-determined by the data, while higher energy levels could be lost in the noise. Constrained curvefitting allows us to obtain sensible results for all parameters, but it is important to determine exactly what part of those results is determined from the data, and what part is determined by the priors. A straightforward way to see the effects of the priors is to modify them, and see how the fit results change. If, for example, we double the prior width, and the error on the fit doubles, then it was likely only constrained by the prior. If, on the other hand,

the fit stays constant, then it must be independent of the prior choice.

We can conduct a similar test to see how many energy levels to include in the fit. In the case of too-few energy levels, adding an additional level will affect previous results, since they suffered from excited-state contamination. Once enough energy levels have been included in the fit, however, adding more will not change the fit results.

4.6 Extracting Physics

At the end of the day, our goal is to obtain measurements that can be compared to experiment. Our results take the form of splittings, kinetic masses, and decay constants.

4.6.1 Splittings

Splittings between meson masses are the simplest of our results. As described above, we calculate spectra of mesons satisfying the given quantum numbers. These spectra are in lattice units ($a = 1$) — hadron masses are calculated as aM . Conversion to physical units requires us to determine the lattice spacing, a . One way to do this is to use some of our results as inputs, and use experimental results to set the scale. A concrete example will be given in Chapter 5.

This is not adding an extra parameter to QCD — lattice QCD has exactly the same number of free parameters as QCD. All we are doing here is determining one of the available free parameters, just as must be done in QCD. In this case, the lattice spacing a corresponds to the strong coupling constant, α_s .

4.6.2 Kinetic Masses

Experimentalists don't usually measure mass splittings, but rather absolute masses. It is therefore helpful for us to calculate the masses of various mesons. It is the kinetic mass, M_2 , that we compare to experiment (see §3.3.1); we obtain it by varying the momentum and performing a fit to Eq. (3.44). For mass splittings, we can obtain the most accurate results by using the M_1 values. The higher-order terms are used to allow us to carry out the fits to high momentum, and to measure any violation of rotational symmetry.

4.7 Summary of Fitting Procedure

Before delving into the details of the fits, it is helpful to outline our goals, and the path we plan to take to achieve them.

4.7.1 Goals

For this project, we will be investigating stable mesons containing a charm quark. The quark content of the study will therefore be charm-charm, charm-strange, or charm-light (where light can refer to either an up or down quark). We will be interested in the meson spectroscopy and, in the case of the heavy-light mesons, the decay constants. In order to provide the most accurate results, it will be useful to describe the spectroscopy both in terms of splittings and in terms of an absolute scale. Our primary focus will be to estimate (and control) all sources of systematic error.

4.7.2 Plan of Attack

Because high accuracy is our goal, we will need to take a meandering path through the calculation. This is because different calculations are expected to give the smallest uncertainties for particular quantities. Combining results from the different meson systems also helps ensure that our results are self-consistent across all stated quantities.

Initial Conditions

We use lattice gauge-field configurations provided by the MILC collaboration. The configurations vary the lattice spacing, size of the lattice, and masses of the light quarks (see Table 4.1). The mass of the strange quark on these lattices is known (from kaon fits), and we use their rough estimate of the lattice spacings as a starting guess for our own measurements. Building from these basic blocks it is possible to perform an analysis of all systematic errors.

Charmonium Gives a^{-1}

The first step of the calculation will involve determining the charmonium spectrum in terms of mass splittings. Because the $c\bar{c}$ splittings depend mostly on light-quark physics, the tuning

of the charm-quark mass is not critical here. We can obtain results for the rest masses (M_1 in Eq. (3.44)) which, though incorrect⁵, share a constant offset within each system. The mass splittings are therefore expected to be quite accurate. The dominant error of our heavy quark action is caused by the clover term, which introduces $\mathcal{O}(\alpha_s a)$ discretization errors. The spatial component of the clover term $\sigma_{\mu\nu} F_{\mu\nu}$ corresponds to $\boldsymbol{\sigma} \cdot \mathbf{B}$ in the continuum, which gives rise to the hyperfine splitting. We can therefore eliminate the $\mathcal{O}(\alpha_s a)$ errors introduced by this term by spin-averaging the S -wave states as

$$m_{\overline{1S}} \equiv \frac{1}{4} (m_{\eta_c} + 3m_{J/\psi}). \quad (4.9)$$

Furthermore, some $\mathcal{O}(a^2)$ discretization errors also cancel in the spin-average.

A determination of the lattice spacing (necessary to calibrate our results) requires a single mass splitting (to set the scale). There are two obvious candidates:

- Use the splitting between the $m_{\overline{2S}} \equiv \frac{1}{4} (m_{\eta'_c} + 3m_{\psi'})$ and the $\overline{1S}$. A potential concern is that the η'_c and the ψ' are very close to the $D\bar{D}$ threshold (see Figure 2.1), and therefore we expect that threshold effects might contaminate our results for those particles.
- Use the splitting between the $m_{\overline{1P}} \equiv \frac{1}{9} (m_{\chi_{c0}} + 3m_{\chi_{c1}} + 5m_{\chi_{c2}})$ and the $\overline{1S}$. Because we don't currently have a lattice calculation of the χ_{c2} state⁶, we instead use the h_c alone to determine $\overline{1P}$. To date, the h_c has only been observed by one experiment. Its “unconfirmed” status makes it somewhat risky to use as a calibration. However, the observed mass $m_{h_c} = 3526$ MeV is consistent with expectations:

$$\begin{aligned} m_{\overline{1P}} &\equiv \frac{1}{9} (m_{\chi_{c0}} + 3m_{\chi_{c1}} + 5m_{\chi_{c2}}) \\ &\approx \frac{1}{9} (3415 \text{ MeV} + 3 \cdot 3511 \text{ MeV} + 5 \cdot 3556 \text{ MeV}) \\ &\approx 3525 \text{ MeV}. \end{aligned} \quad (4.10)$$

Because the h_c mass is consistent with expectations, we cautiously use it to set the scale of our calculation. Comparisons with lattice spacings determined from the $\overline{2S} - \overline{1S}$ splitting,

⁵See the discussion on our usage of M_1 and M_2 on page 27.

⁶The calculation of the χ_{c2} is in progress.

as well as lattice spacings determined by other groups, will be used as a cross-check on our results.

Because simulations at the physical light-quark mass would be prohibitively expensive, all simulations are performed at heavier masses in the range $m_s/10 \leq m_l < m_s$. To compare with experiment, it is therefore necessary to perform an extrapolation down to the physical light quark mass, $m_l \approx m_s/25$. The dependence on m_{sea} is mild, and we therefore simply fit to a constant or linear function.

D_s Gives m_c

At this point, we still do not have a properly-tuned charm-quark mass. This is intentional, as the charm quark mass can be tuned more accurately using the D_s than with charmonium. We can compute the kinetic mass (M_2 in Eq. (3.44)), which can be compared directly to the experimentally-measured mass of the D_s . Because we already know the strange-quark mass in lattice units (from MILC's kaon results) and the lattice spacing (from the charmonium $1P-1S$ splitting), this comparison allows us to tune the only remaining parameter: the charm-quark mass m_c . Again, the dependence on m_{sea} is mild.

In practice, we adjust the parameter κ , which is related to m_c through Eq. (3.36):

$$am_0 = \frac{1}{u_0} \left[\frac{1}{2\kappa} - \frac{1}{2\kappa_{\text{crit}}} \right] \quad (4.11)$$

where the $1/u_0$ provides tadpole improvement. Values of κ_{crit} for each lattice are listed in Table 4.1. The tuning selects $0 < \kappa_{\text{charm}} < \kappa_{\text{crit}}$ where $\kappa \approx 0$ corresponds to an infinite bare mass and $\kappa \approx \kappa_{\text{crit}}$ corresponds to a massless particle.

We can use this charm-quark mass to re-investigate the charmonium spectrum to obtain an absolute offset for comparison with experiment. Also, we can use the D_s results to obtain information about the D_s decay constant f_{D_s} .

Chiral Extrapolation

Once the lattice spacing a and the charm quark mass m_c have been determined, all other quantities are predictions. The D system, however, introduces an additional complication: we

must extrapolate in both m_{sea} and m_{val} . In the limit $m_{u,d} \rightarrow 0$ QCD has a chiral symmetry. An effective theory which describes this symmetry is chiral perturbation theory (χ PT). It is a well-tested theory, so we may use it to guide our extrapolation to the light quark mass. We will also use it when interpolating to the physical strange quark mass (the strange quark mass used in our calculations has some error, which we want to remove). Lattice χ PT has some additional complications. For example, we will work in a “partially quenched” theory, where the mass of the sea quarks does not necessarily equal the mass of the valence quarks. This is possible on the lattice because the sea quarks correspond to the $\det M_{\text{sea}}$ term while the valence quarks correspond to the M_{val}^{-1} term. By using $M_{\text{sea}} \neq M_{\text{val}}$ we can easily achieve partial quenching. Additional complications arise from the staggering of the sea quarks, and the taste doubling of the naïve valence quark. These issues are discussed in Appendix E.

Investigation of the D decay constant f_D is also possible, but likewise will require a complicated chiral extrapolation.

4.8 Systematic Errors: Estimation or Removal

As with any measurement, we must estimate, and hopefully remove, any systematic errors. Previous work in lattice QCD was performed under the quenched approximation, where sea quark effects were ignored. It was estimated that the quenched approximation led to 20% errors. This work does not make that approximation, but we still have to consider the effects of several other approximations.

4.8.1 Finite Lattice-Spacing Errors

The most obvious approximation is that we discretize space-time on a lattice with a finite lattice spacing, a . It has been estimated that the computational complexity increases as a^{-7} [61] (a^{-4} comes from the size of the lattice, a^{-2} from the fermion determinant, and a power of a^{-1} that accompanies fermion masses). If we are to obtain results valid in the continuum, we must have control over these lattice spacing errors. As described in Section 3.4 this control is most efficiently obtained by using improved actions. Even with the improved actions, however, we must test our results on several lattice-spacings to guard against lattice-spacing effects.

4.8.2 Finite Volume Effects

Another problem is that we are working in a finite box. We use periodic boundary conditions, but if the box isn't large enough (relative to the size of the hadron we are investigating) then there will be some effect. The solution to this is fairly simple: we increase the size of the box. Of course, working in four space-time dimensions means any increase in size gets raised to the fourth power (in reality the computational complexity scales as L^5 [61]), so we can't completely eliminate these effects. The goal, therefore, is to reduce them until they are comparable to the statistical uncertainties. It is also helpful to take the periodicity into account when performing our analysis, hence the usage of a hyperbolic above, rather than a simple exponential.

4.8.3 Violation of the Chiral Limit

Recall that our algorithm requires us to compute the inverse of the fermion matrix for valence quarks, and the determinant of the fermion matrix for sea quarks. As the quarks become lighter, the condition number of the matrix increases, indicating it is closer to being singular. As explained in Appendix C.1, the computational effort required to calculate the inverse or determinant of a matrix increases with its condition number. It is therefore computationally expensive to simulate light quarks, with a complexity that increases as $\left(\frac{m_\pi}{m_\rho}\right)^{-6}$. We are forced to use light quarks that are heavier than their physical values, but light enough to allow for an accurate chiral perturbation theory (χ PT) to be used. In practice, we use light quark masses in the range $m_s/10 \leq m \leq m_s$ and extrapolate to the physical light quark mass.

Chapter 5

Fitting Procedure in Detail

5.1 Preparation: Statistical and Numerical Techniques

Before analyzing data, it is helpful to review some of the statistical techniques we will be using.

5.1.1 Jackknife Method

Because it is a sum of exponentials, our model function implicitly assumes that the meson propagator is positive at every timeslice. Statistical fluctuations in the data, however, can cause individual propagators to fluctuate negative on occasion. This is real, if nonphysical, aspect of the raw data. To prevent situations like this from affecting our results, we use a technique called the jackknife method. The jackknife method also gives a better estimate of errors for correlated input.

In standard statistics, the mean and variance are computed using the data taken one point at a time. Under the jackknife method, the inverse is done: the mean and variance are computed using *all* of the data *except* one point at a time. By doing this, we avoid situations where an individual point may fluctuate negative while the average is positive.

It is immediately obvious that the calculation of the mean need not change using this method. The correlation matrix, however, changes from being computed as

$$\sigma_{i,j}^2 = \sum_{n=1}^N \frac{(\bar{x}_i - x_{i,n})(\bar{x}_j - x_{j,n})}{N(N-1)} \quad (5.1)$$

to

$$\sigma_{i,j}^2 = \frac{N-1}{N} \sum_{n=1}^N \left(\bar{x}_i - \frac{1}{N-1} \sum_{m \neq n} x_{i,m} \right) \left(\bar{x}_j - \frac{1}{N-1} \sum_{m \neq n} x_{j,m} \right) \quad (5.2)$$

$$= \frac{N-1}{N} \sum_{n=1}^N \left(\bar{x}_i - \frac{N\bar{x}_i - x_{i,n}}{N-1} \right) \left(\bar{x}_j - \frac{N\bar{x}_j - x_{j,n}}{N-1} \right) \quad (5.3)$$

5.1.2 Statistical Uncertainties

Errorbars are supposed to give a confidence interval. Assuming Gaussian errors, an errorbar of 1σ gives a 68% confidence bound. In our case, however, the errors may not be Gaussian, which complicates our statement of statistical uncertainties. There are three methods we can use to determine these uncertainties: chi-squared, hessian, and bootstrap.

Chi-Squared Errorbars

Most scientific work specifies errorbars to indicate the 1σ , or 68% confidence bound, of the result. The most precise way to determine this bound is to map out the locus of all points which increase the χ^2 of the fit by unity from the minimum. This is typically easy in a one-dimensional fit, but becomes computationally expensive in a multidimensional fit. As a result, this method is rarely used to determine statistical uncertainties.

Hessian Errorbars

If Gaussian errors can be assumed, then a nice trick is to quote errorbars based on the Hessian matrix of the χ^2 function. The Hessian matrix provides second derivatives of the χ^2 , which means its diagonal provides the curvature of the χ^2 function in the direction of each variable. Approximating the χ^2 as a parabola in each direction, we can therefore construct the appropriate bounds for our errorbars. This method has the advantage of being very fast and easy, but the disadvantage of requiring the assumption of Gaussian uncertainties.

Bootstrap Errorbars

We are fitting to a high-dimensional model function, and do not have the guarantee of Gaussian uncertainties, so we use the bootstrap method. To construct the errorbars, we repeat the fit multiple times (on slightly different data sets) and quote an error based on the variation in fitted parameters.

Obviously the fits need to be done to different datasets, but we have a very limited set of data. The solution is to generate new datasets from the existing data, with the distribution of the existing data. Consider the case where we have N data points, x_i . We construct a set of data by randomly selecting N of those points, with replacement. This new “bootstrap” dataset has all the statistical properties of the original dataset.

Once we have all the bootstrap datasets we calculate their best-fit parameters. The parameters are then sorted to construct an errorbar. The central value is taken to be the fit result from the bootstrap dataset which includes each data point exactly once. The errors come from the inclusion of 68% of the fit results.

When performing a constrained fit, the priors are an input, just like the data is. When constructing our bootstrap samples it therefore makes sense to adjust the priors as well. This is done by adjusting them within each prior’s width, using Gaussian statistics.

A Real-World Example

It is instructive to look at the differences in errorbars for a real-world situation. I show a slice through the parameter-space in Figure 5.1. One can see from the figure that hessian errors can be quite unreliable, especially in a case where the minimum is asymmetric. The figure also shows how priors can be useful to assist the fitter in locating the minimum for cases where there may be a fairly large plateau in the χ^2 landscape. Without the priors, the fitter would not have been able to determine which direction was downhill if it wandered into the region $\log Z < -4$. With the addition of the priors, however, that is not a problem. Of course, once the true minimum is located, the priors have negligible effect on the fit result.

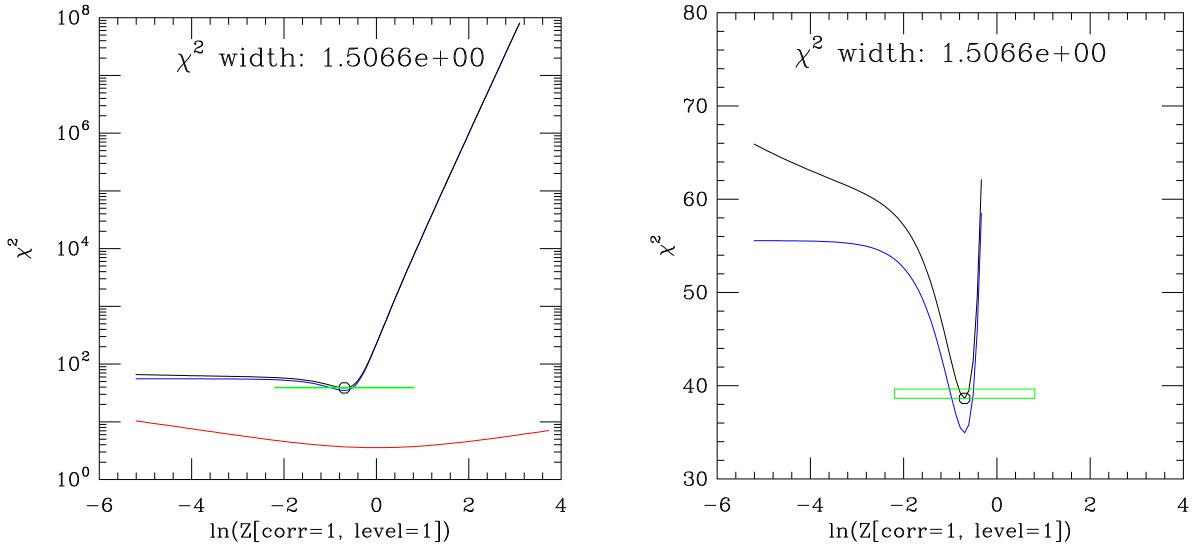


Figure 5.1: Example of the effect of priors and different types of errorbars, for the case of a coefficient. The red curve indicates the contribution to the χ^2 from the priors, while the blue curve gives the contribution from data. The black curve is their sum, and is the function that is minimized. The green square has height 1 (to show where a true χ^2 errorbar should lie) and width determined by the hessian errorbar.

5.1.3 Correlations and Blocking

One concern is that of correlations in the data skewing our results. It is necessary to check for, and eliminate, any correlations. A simple way to do this is by “blocking” the data. This is done by taking consecutive data points and averaging them in blocks of 2, 3, etc. Increasing the block size reduces the effect of correlations from one block to the next, but reduces the number of blocks, and therefore the statistical significance of the final result. If the data is highly correlated, doubling the block size will increase the error bars by roughly a factor of $\sqrt{2}$. We can therefore try a few block sizes to eliminate correlations (and confirm that they are gone).

5.1.4 Automatic Differentiation

As previously mentioned, we use the Levenberg-Marquardt method for our fits. As a tradeoff for the efficiency of this method, it is necessary to have the derivative of the model function with respect to each parameter. For simple model functions this is not an issue, but in our case the model functions are sufficiently complex that obtaining the vector of derivatives is a non-

trivial task. The difficulty of programming a derivative functional is further complicated by the fact that we do not know until runtime the dimension of our problem — the number of energy levels and coefficients. The fitting process requires us to vary the number of parameters to watch for bias. Computer routines defining the derivative of our model function can therefore be extremely complicated, and therefore highly error-prone. As a result, initial testing of a model function proved very tedious, as can be seen in Appendix D.

As the project progressed, we found it useful to try new model functions to see if they could provide tighter constraints on the fit parameters. For the purposes of rapidly testing new model functions, it was therefore useful to find an automated way to obtain the derivative functions. The chosen method is based on that of Wengert [62]. In this method, each variable is taken to consist of two components: a value and a derivative. The number 5 might be represented as $(5, 0)$, and the variable x as $(x, 1)$. More complicated expressions can be built up by teaching the computer about the chain rule. For example, we teach the computer that

$$\text{der}(xy) \equiv \text{val}(x)\text{der}(y) + \text{der}(x)\text{val}(y), \quad (5.4)$$

and that

$$\text{der}(\cosh(x)) \equiv \text{val}(\sinh(\text{val}(x)))\text{der}(x). \quad (5.5)$$

The computer then can already handle something simple, such as the derivative of $5 \cosh(x)$. It proceeds as follows:

$$\text{der}(5 \cosh(x)) = \text{val}(5)\text{der}(\cosh(x)) + \text{der}(5)\text{val}(\cosh(x)) \quad (5.6)$$

$$= 5\text{val}(\sinh(\text{val}(x)))\text{der}(x) \quad (5.7)$$

$$= 5 \sinh(x) \quad (5.8)$$

It is easy to see that by defining just a few expressions we can teach the computer to automatically differentiate any model function we might require.

While useful for prototyping new model functions, the automatic differentiation does not completely replace the need to program up explicit derivative functions. Some simple tests

showed that automatic differentiation runs about 12 times slower than a pre-programmed derivative. For production running, therefore, hand-coded derivatives are still best. Those functions can be tested against the results of the automated derivatives to ensure correctness.

5.2 Problems and Solutions

Constrained fitting, while providing better control over systematic errors, is not without its drawbacks. We still have to specify a model function (and its derivative) and estimate errors. The constrained fits make error estimation slightly more complicated, since the possibility of over-constrained priors must be considered. Furthermore, we had trouble with statistical fluctuations causing bad fits, and had to find ways to eliminate them. Specific examples of these problems, and how we overcame them, will be given in the following sections.

5.3 Charmonium Gives a^{-1}

The first step of the plan outlined in Section 4.7.2 is to study charmonium. We tie together two charm quark propagators to obtain charmonium propagators. From these propagators we hope to extract mass splittings in the charmonium spectrum. We can also use propagators at $\mathbf{p} \neq \mathbf{0}$, combined with the energy-momentum relation, to determine the kinetic masses of these mesons. That step, however, will need to wait until we have tuned the charm-quark mass.

5.3.1 The Model Function

The basic equation we want to model is Eq. (4.4):

$$G_{\text{src,snk}}(t) = \sum_n Z_{\text{src}}^{(n)} Z_{\text{snk}}^{(n)} \left(e^{-E_n t} + e^{-E_n(T-t)} \right). \quad (5.9)$$

Of course, we know various things about the energies, such as the obvious fact that energies are ordered. We can eliminate a lot of redundancy in the fit parameters by forcing the energies

to take on a specific order by specifying not the energies, but the splittings. We therefore write

$$\begin{aligned}
 E_1 &= E_0 + e^{\epsilon_1} \\
 E_2 &= E_1 + e^{\epsilon_2} \\
 E_3 &= E_2 + e^{\epsilon_3} \\
 &\vdots
 \end{aligned}
 \tag{5.10}$$

where our fit parameters are E_0 and the ϵ_i .¹

Some other changes can be made to the model function as well. The sum of exponentials is difficult for the computer to handle, since for $t \neq T/2$ they will be of vastly different scales, and therefore subject to roundoff errors due to numeric underflow. To reduce such errors it is helpful to recognize that

$$\exp(a + b) + \exp(a - b) = 2 \exp(a) \cosh(b)
 \tag{5.11}$$

which has the advantage that it avoids roundoff errors.

5.3.2 Effective Mass

Many of the fit parameters can be estimated in advance, and used to set priors for the fits. For example, we know from other analyses that our lattice spacings are $a^{-1} \approx 1.6$ GeV. We can combine our knowledge of the approximate lattice spacing with our knowledge of the scale of mass splittings (using experimental inputs for those which have been observed, or assuming $\mathcal{O}(\Lambda_{\text{QCD}})$) to construct priors for the energy splittings. We can also expect that coefficients are $\mathcal{O}(1)$. One parameter, the ground state energy, cannot be guessed. This is because the rest mass M_1 in the dispersion relation (Eq. (3.44)) has an offset that is not directly related to any physical parameter.

To get a starting guess for the ground state, we take advantage of the fact that the fits fall off exponentially, with the ground state having the slowest decay. We can therefore look

¹ Note that an alternative method to specify the splittings is via ϵ_i^2 . Our method of using an exponential has the advantage of allowing the range in the positive direction to be larger than the range in the negative direction, which gives more meaning to specifying an energy splitting to within 50%.

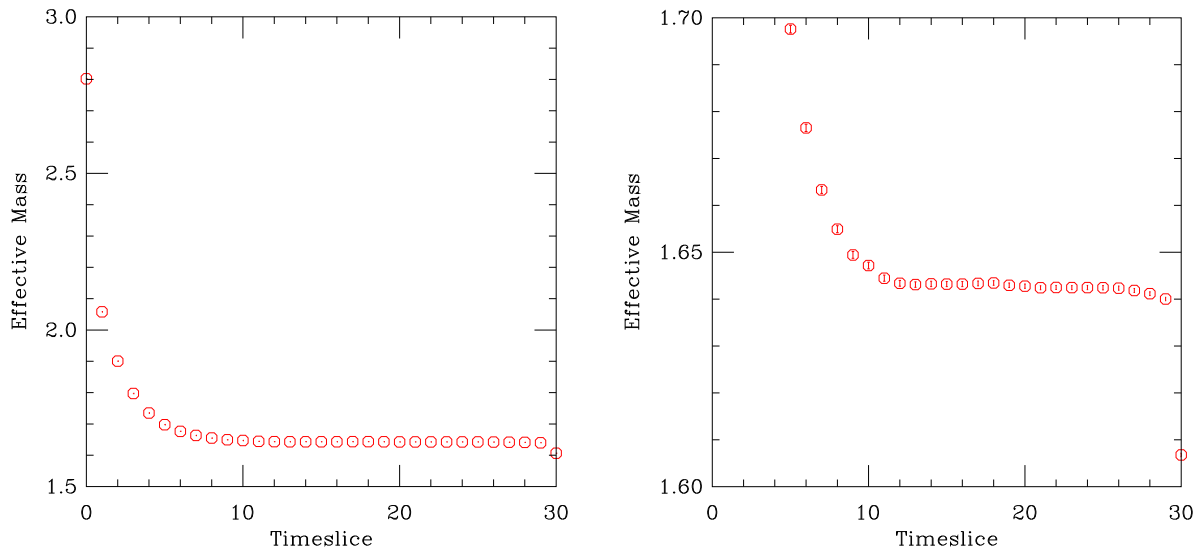


Figure 5.2: (a) Effective mass plot for the local-local smearing of the pion state of charmonium on a .01/.05 lattice. (b) A magnification of the same plot.

at the exponential decay and pick out the leading exponential. A convenient way to do this is plot the “effective mass” of consecutive timeslices. Assume the decay falls off as a single exponential. Then

$$G(t) = \exp(-M_{\text{eff}}t). \quad (5.12)$$

Comparing consecutive timeslices, we find

$$G(t+1) = \exp(-M_{\text{eff}}t) \exp(-M_{\text{eff}}) \quad (5.13)$$

and therefore we can solve:

$$M_{\text{eff}} = \log(G(t)/G(t+1)). \quad (5.14)$$

I show a sample effective mass plot in Figure 5.2. The important characteristic of this plot is that it levels off after the first few timeslices, since the excited-state exponentials have died off. The plateau therefore gives us a reasonable guess of the ground state energy — in this case 1.64. We repeat this process for each dataset in order to obtain a reasonable estimate of the ground state energy.

It is worth noting that this is essentially the same as performing a fit to a single exponential, with the advantage that it allows us to quickly identify both a reasonable timeslice range and

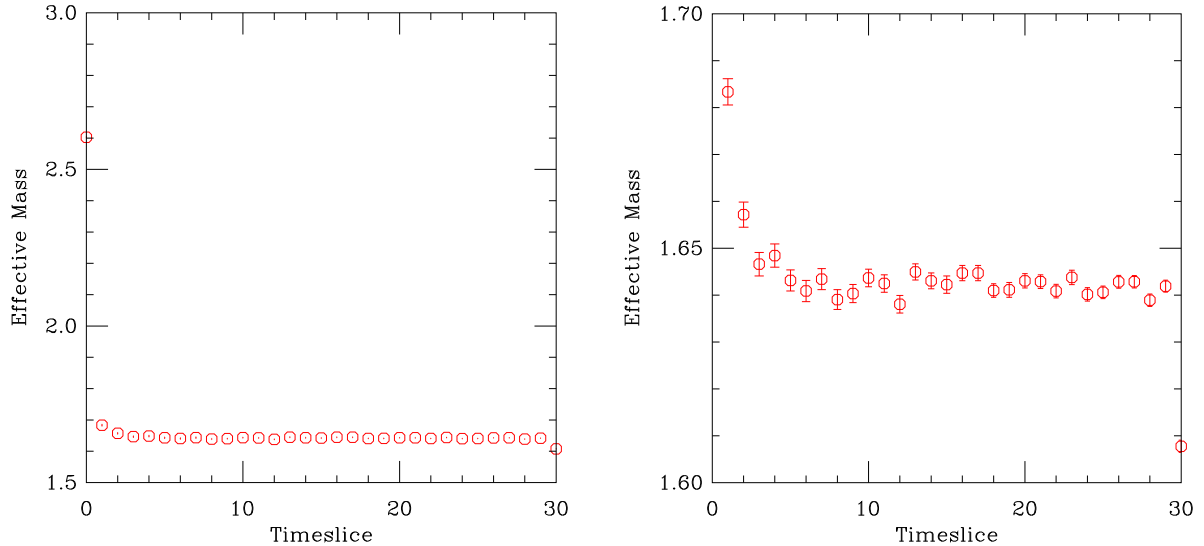


Figure 5.3: (a) Effective mass plot for the $1S$ - $1S$ smearing of the pion state of charmonium on a $.01/.05$ lattice. (b) A magnification of the same plot. Note the plateau is much longer than that of Figure 5.2.

the fit result.

5.3.3 Local, $1S$, and $2S$ Smearings

In the previous section, we identified the plateau of the effective mass plot as the energy of the ground state. But it is clear from Figure 5.2 that the plateau is contaminated, at early timeslices, by the excited states. The usable portion is therefore reduced to later timeslices, where the signal-to-noise ratio is diminished. In the standard δ -function smearing, there is a contribution from all energy levels. By spatially smearing the meson propagator wavefunction (using the wavefunctions calculated in Figure A.2) we can enhance the contribution from the desired state. For example, using a $1S$ - $1S$ smearing changes Figure 5.2 to have a much larger plateau, as seen in Figure 5.3. The tradeoff is that there is a bit more noise. Similarly, one can enhance the excited-state contribution by using a $2S$ smearing, or a P -wave state using the $1P$ smearing.

5.3.4 Chaining Procedure for Starting Values

One of the problems with multidimensional fitting is that the final fit result can depend on the starting guess. This is a direct result of the fact that the parameter phase space is large,

and often too complex to be guaranteed to find the global minimum simply by following a method of steepest descent. It is helpful, therefore, to start as close as possible to the correct solution.

Obviously, if we knew the correct solution in advance, we wouldn't need a fitter, and experience shows that even setting priors does not always sufficiently constrain the starting guess. We recognize, however, that it is much easier to find the true minimum when we perform fits with fewer parameters. We can therefore run a chain of fits where we add a single energy level at each link in the chain, and use the results of the previous fit as the starting guess. The starting guess for the first fit simply comes from our priors. It is important to note that the priors of subsequent fits are not adjusted — the priors encode our bias before beginning the fitting procedure, and do not change once we have seen the data.

This method has proven to work fairly well, but can lead to problems if one of the fits in the chain happens to be a poor fit. Subsequent fits will then be given a poor starting guess, and may find a local minimum, rather than the global minimum. It is therefore important to keep an eye on the fitting procedure, rather than fully automating it.

5.3.5 Setting the Priors

The priors must be treated very carefully to avoid introducing any bias into the fit results. Ideally, the priors will have negligible effect on the fit at the true minimum — their purpose is only to prevent the fitter from wandering into a non-physical region of parameter space. However, the priors for highly excited states may (and should) directly affect the results. It is therefore important to learn to recognize when the priors are affecting the data so we can determine which results depend only on the data, and which are affected by the priors.

A simple way to observe the influence is to modify the prior widths. If the width can be changed without affecting the fit results (either central value or errorbar), then we can be confident that the results are independent of the priors. As an example, consider a fit to the $1S$ - $1S$ smearing of the η_c , shown in Figure 5.4. As the figure shows, if either the energy or coefficient prior is too tightly constrained, the fit results will be biased. Once the constraints are sufficiently relaxed, however, the fit will adjust to its correct value. This is highlighted by

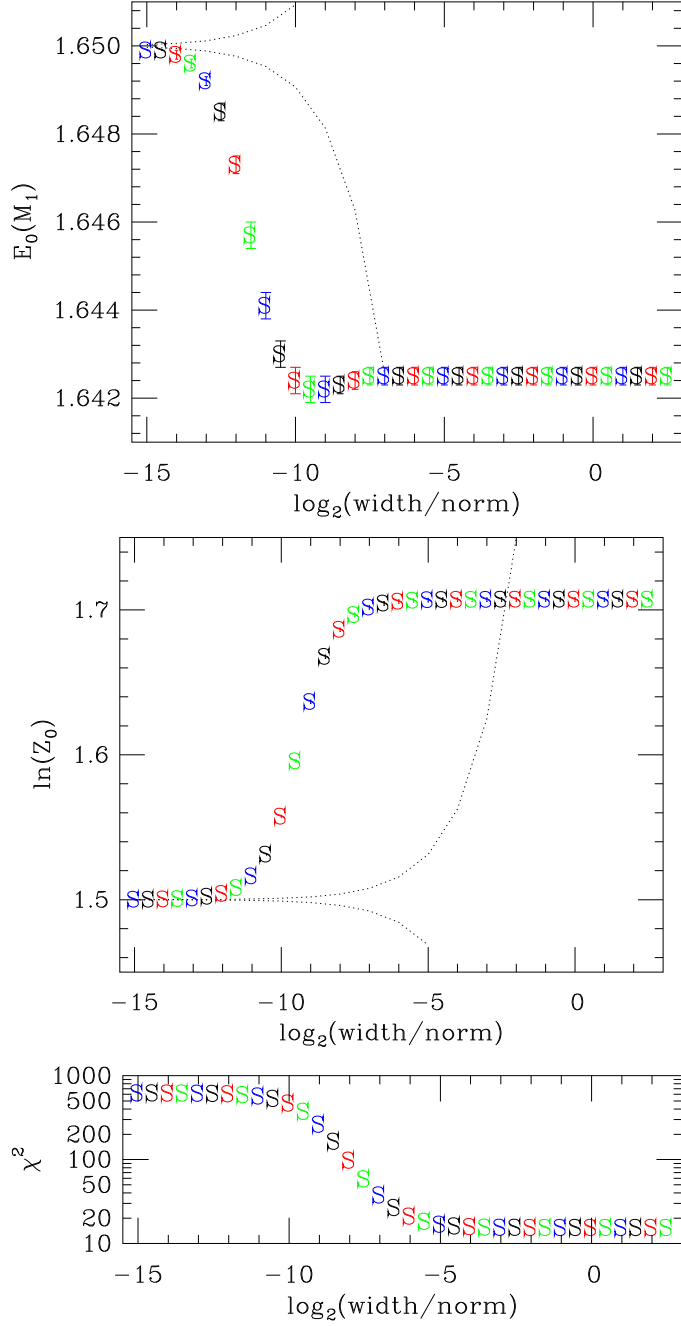


Figure 5.4: A test of the effect of the priors on a one-state fit to the $1S$ - $1S$ smearing of the η_c . The horizontal axis shows $\log_2(\text{prior width}/\text{standard width})$. Figures (a) and (b) show the energy and coefficient, respectively. In these figures, the dotted line shows the 1σ constraint introduced by the priors, which reaches our standard value at a value of 0 on the horizontal axis. Figure (c) shows the χ^2 of each fit. The colors serve only to guide the eye between the plots.

Table 5.1: The priors used in the charmonium analysis. Here $\overline{1S} \equiv \frac{1}{4}\eta_c + \frac{3}{4}J/\psi$ and $\overline{2S} \equiv \frac{1}{4}\eta'_c + \frac{3}{4}\psi'$.

State	E_0	ΔE_0	$E_1 - E_0$	$\Delta(E_1 - E_0)$
pi	eff. mass	$\frac{1}{2}\eta_c \approx 1.49$ GeV	$\eta'_c - \eta_c \approx 674$ MeV	$\frac{1}{2}(\eta'_c - \eta_c) \approx 337$ MeV
ro	eff. mass	$\frac{1}{2}J/\psi \approx 1.55$ GeV	$\psi' - J/\psi \approx 589$ MeV	$\frac{1}{2}(\psi' - J/\psi) \approx 295$ MeV
b1	eff. mass	$\frac{1}{2}h_c \approx 1.76$ GeV	$\overline{2S} - \overline{1S} \approx 610$ MeV	$\frac{1}{2}(\overline{2S} - \overline{1S}) \approx 305$ MeV
a0	eff. mass	$\frac{1}{2}\chi_{c0} \approx 1.71$ GeV	$\overline{2S} - \overline{1S} \approx 610$ MeV	$\frac{1}{2}(\overline{2S} - \overline{1S}) \approx 305$ MeV
a1	eff. mass	$\frac{1}{2}\chi_{c1} \approx 1.76$ GeV	$\overline{2S} - \overline{1S} \approx 610$ MeV	$\frac{1}{2}(\overline{2S} - \overline{1S}) \approx 305$ MeV

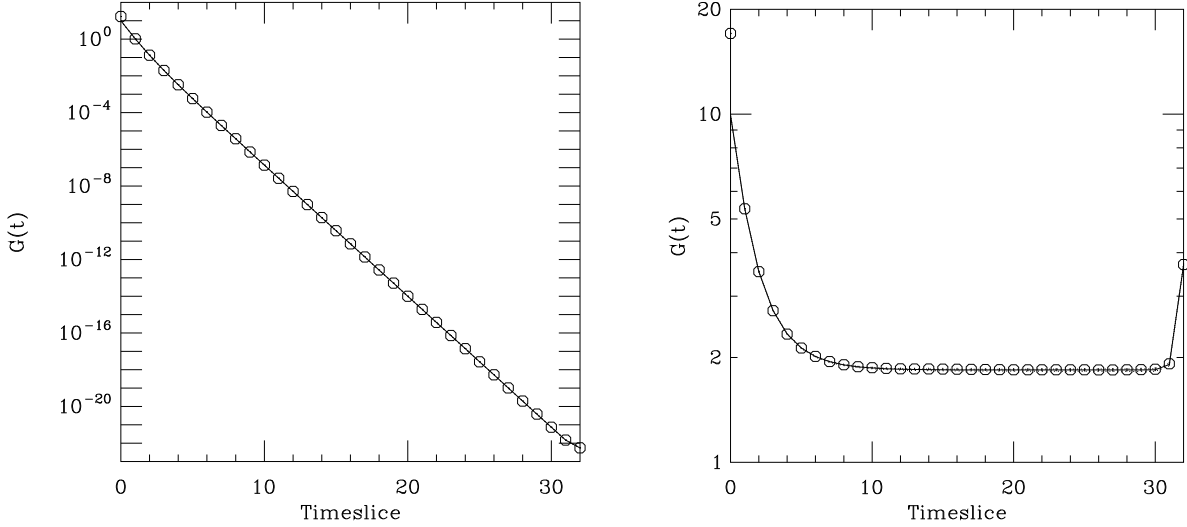


Figure 5.5: (a) A 4-state fit to the propagator of Figure 5.2. Errorbars for both the data and the fit are plotted, but are too small to resolve. (b) The same fit, but shown with the leading exponential removed to show the remaining terms. The fit goes up on the last timeslice due to the periodic boundary conditions of the lattice.

Figure 5.4(c), which shows how the χ^2 drops dramatically as the fit is allowed to reach the proper solution.

For our final results, we use priors as shown in Table 5.1.

5.3.6 Initial Results

Using the techniques already described, we can obtain some initial results and evaluate whether the fitter is working as expected. Figure 5.5 shows a sample fit. Using the chaining procedure and fitting to different numbers of states allows us to see very clearly the utility of using a smeared propagator. Figure 5.6 shows the fits to (a) the η_c and (b) the η'_c using the δ - δ smearing only, the $1S$ - $1S$ smearing only, or the $2S$ - $2S$ smearing only. Note that the $1S$

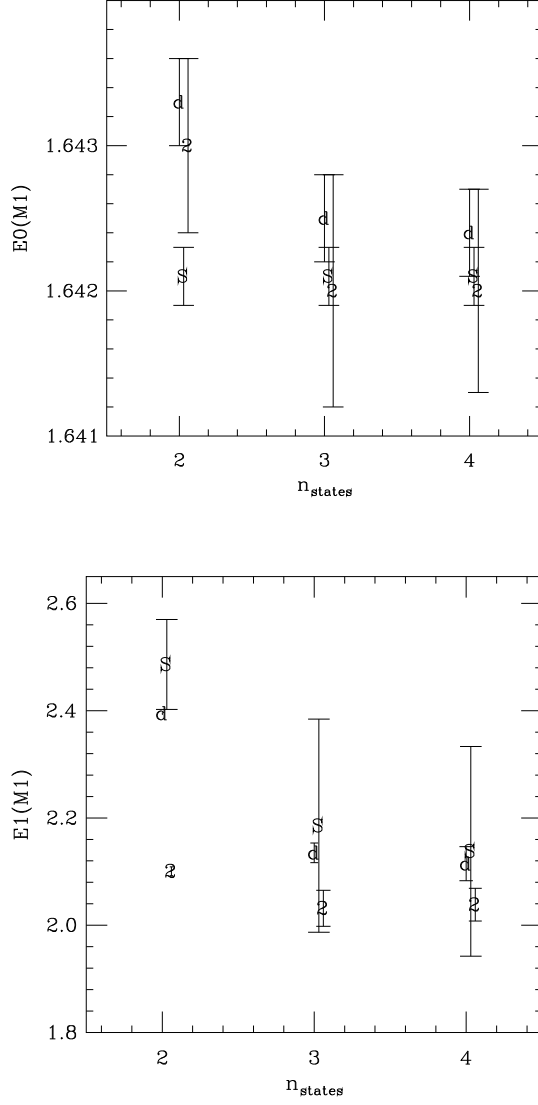


Figure 5.6: Fits to (a) the η_c and (b) the η'_c . Shown are fits that used the δ - δ smearing only (indicated by a d), the $1S$ - $1S$ smearing only (S), and the $2S$ - $2S$ smearing only (2).

smearing allows us to determine the ground state very accurately, even without including many excited states. Meanwhile, the $2S$ smearing accomplishes the same for the excited state.

We can also perform combined fits to multiple smearings, using any elements of the matrix

$$\begin{pmatrix} \delta\text{-}\delta & \delta\text{-}1S & \delta\text{-}2S \\ 1S\text{-}\delta & 1S\text{-}1S & 1S\text{-}2S \\ 2S\text{-}\delta & 2S\text{-}1S & 2S\text{-}2S \end{pmatrix}. \quad (5.15)$$

While it is tempting to simply use the entire matrix, we caution that doing so may not be the best choice. Some entries have a poor signal-to-noise ratio, as seen in Figure 5.7. Note that if we restrict ourselves to the diagonal of this matrix we obtain $Z_{\text{src}}^{(n)} Z_{\text{snk}}^{(n)} = (Z^{(n)})^2$, and can constrain the $Z^{(n)}$ to be positive. This turns out to be an important constraint.

5.3.7 Cutting Out Anomalous Bootstraps

The bootstrap method is supposed to be insensitive to individual fit results. When done properly, several fits are performed, and a histogram of the sorted fit parameters forms an approximate bell curve. Because one standard deviation isn't very far out into the tail, selecting the central 68% should be fairly insensitive to the individual fit results.² Sometimes, however, the parameter-space has two minima that are each fairly likely to be the true minimum. This can lead to a distribution which has a second (smaller) peak. Depending on the size of the second peak, it can cause us to overestimate one side of the bootstrap errorbar (and underestimate the other side). The occurrence of this problem is easy to detect: the bootstrap errorbar exhibits a strong asymmetry. By looking at which minimum occurs more frequently and has a smaller χ^2 we can determine which of the minima is the “true” one, and which is the result of a statistical fluctuation. In those cases, we manually eliminate the portion of the tail corresponding to the incorrect fits. The resulting data then has reasonable errorbars, which are insensitive to the precise location of our cut. Prior knowledge of physics may also help to eliminate one of the minima as being in a disallowed region of phase space,

²Asymmetric errors are allowed by this method also: we take 34% from each side of the “best fit” value.

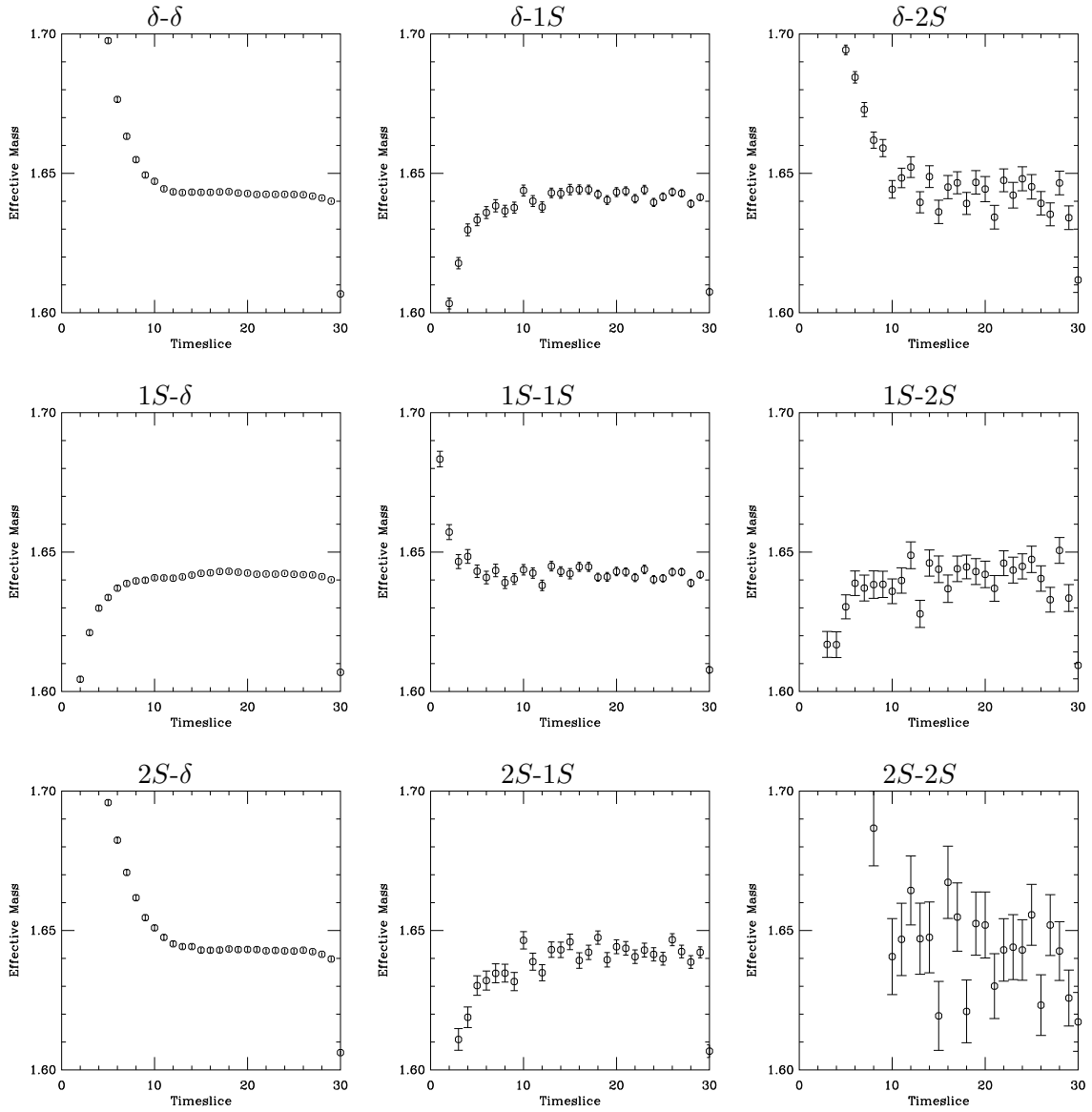


Figure 5.7: Effective mass plots corresponding to the matrix of Eq. (5.15). Note that the lower triangle of the matrix has a better signal-to-noise ratio, in general, than the upper triangle.

though we will see that incorporating that knowledge into the model function or priors is the best approach.

5.3.8 A More Robust Fitting Procedure

We found our naïve model function did not sufficiently constrain the parameters. Our conservative choice for the energy splitting prior widths allows the possibility of the minimizer squeezing an extra state in the gap between two energy levels, or even below the ground state. Such a state is nonphysical, but if the fitter adjusts its coefficient to 0, the state would not contribute to the fit. While this does not occur all the time, it occurred with sufficient frequency to distort the bootstrap errorbars, making it impossible to produce a quantitative estimate of the statistical uncertainty in the result.

If we could prevent the coefficient from becoming small, the insertion of a spurious “state” could not occur. Unfortunately we do not know the sign of the coefficients, and it is not possible to allow the sign to vary while excluding zero.³

When the source and sink smearings are the same, the model function reduces to:

$$G(t) = \sum_n Z_{(n)}^2 \left(e^{-E_n t} + e^{-E_n(T-t)} \right). \quad (5.16)$$

Here it is clear that we can constrain the $Z_{(n)}$ to be positive. This is the case whenever we do not allow mixed terms: $\text{src} \neq \text{snk}$. By constraining ourselves to data where the source and sink share the same smearing, we can force the coefficients to be positive using $Z = \exp(\zeta)$.

The constrained model function alone is not sufficient to cure the problem of nonphysical states (due to our lack of knowledge of the magnitude of Z). As an example, consider a two-state fit to the h_c 1S-1S smearing. Depending on the initial conditions, the fitter might find either of the solutions listed in Table 5.2.

We were unable to eliminate the statistical fluctuations through the use of the priors, as doing so would introduce a bias into our results. Figure 5.8 reveals that the principal trigger for settling on the incorrect minimum is the starting guess. When performing the bootstrap

³A functional form like $z \equiv 1/x$ contains a discontinuity which a minimizer would be unable to cross. One might imagine “reducing” the area near 0 with a function like $z \equiv \tan^{-1} x$, but that simply makes the fitter work harder to find the minimum where $z = 0$, and doesn’t prevent it.

Table 5.2: Two plausible local minima for the h_c $1S$ - $1S$ smearing. The first is believed to be the correct solution; the second contains a spurious ground “state”. Note that the first excited state of Solution 2 agrees with the ground state of Solution 1. Also note that the coefficient of the ground state of Solution 2 is comparatively small.

	Solution 1	Solution 2
E_0	1.97553 ± 0.00875223	1.78266 ± 0.0788273
$E_1 - E_0$	$0.339531^{+0.206815}_{-0.128527}$	$0.211059^{+0.0883556}_{-0.0622823}$
Z_0	$2.63319^{+0.0891991}_{-0.0862765}$	$0.379053^{+0.382813}_{-0.190462}$
Z_1	$0.756851^{+0.538772}_{-0.314729}$	$2.69700^{+0.0549813}_{-0.0345334}$
χ^2/dof	10.496 / 14	12.666 / 14

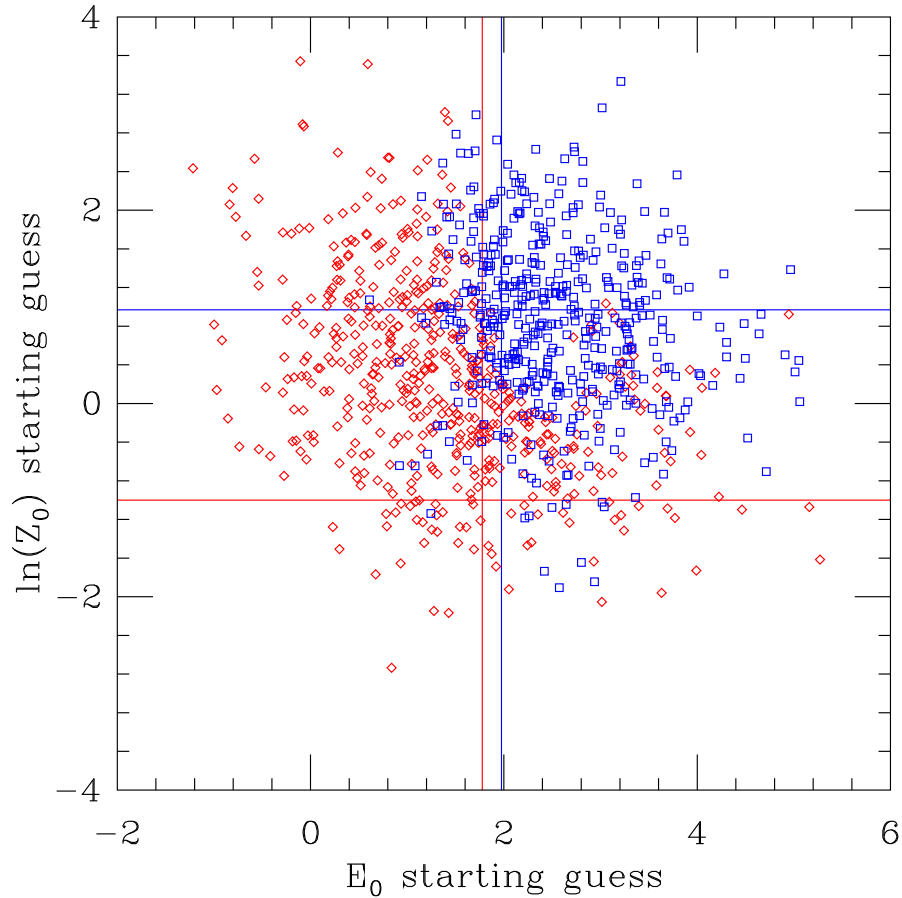


Figure 5.8: Fit results for the h_c $1S$ - $1S$ smearing are shown as a function of starting guess. Here the ensemble and priors were held fixed, and the starting guess allowed to vary. The blue squares correspond to Solution 1 of Table 5.2, while the red diamonds correspond to Solution 2. The vertical blue (red) line indicates the E_0 of Solution 1 (2), while the horizontal lines indicate the $\ln(Z_0)$.

fits, we vary the prior central values randomly (with Gaussian width determined by the prior width) to ensure the priors are not affecting the fit results. It is not necessary, however, to vary the starting guess with the prior central values. Keeping the starting guess fixed greatly reduces the problem of settling on an incorrect minimum.

This method, while an improvement, leaves open the question of how to ensure we aren't *always* using the wrong starting guess. There are three options for how to select a starting guess:

- set it to the original prior central value: This doesn't work well because the prior central values are often a poor guess at the final fit results.
- allow it to vary, but only by $\sigma/2$: This seems promising, but the effect is insufficient to solve the problem.
- fix it to the results of our "best fit": This works well, but suffers from catastrophic failure if the "best fit" exhibited the problem.

We choose the third option, and introduce some additional precautions to ensure that the "best fit" contains proper physics. Whenever performing a series of fits, we do not begin with the final number of energy levels, but rather use a "chaining" method to build up to it. For example, the first fit will be to two states, then we'll fit to three states, and finally to four states. We can use the fit results of the two-state fit as our starting guess for the three-state fit, since the parameters should be fairly close to their final values. Then we use the results of the three-state fit as the starting guess for the four-state fit.

Additionally, we preface each fit with a test run where we vary the starting guess randomly and observe the effect on the fit results. This is an effective way to map out the space of plausible local minima. We then ensure that the minimum chosen as the starting guess for future fits is one that "makes sense".

Figure 5.8 suggests an additional technique for finding the true ground state. Fits where the starting guess assumed a large E_0 and Z_0 are very likely to find the correct ground state. This is because the data (in either a δ - δ or $1S$ - $1S$ smearing, but not necessarily in a $2S$ - $2S$ smearing) will pull the fit down to the correct values. We might therefore intentionally set

our priors (and starting guesses) slightly high, thereby ensuring the fitter will find the true minimum.

It is important to note that the starting guess and the prior central value do not necessarily need to be the same. The information we retain from an n -state fit to an $(n+1)$ -state fit does not need to affect our priors. In theory, it should be acceptable to modify our prior central values (keeping the prior widths constant) based on results of previous fits. In practice, we found this is not necessary — the fit results are rarely affected by the prior central values if the prior widths are set to reasonable values. For simplicity, we only quote results where the prior central values were *not* modified.

5.3.9 Testing for Bias

Implementing constrained curvefitting allows us to include additional states in our fits, but it comes at the cost of requiring additional user inputs to set all the priors. In addition to prior central values and widths (for each parameter), we also have freedom to choose

- the number of states to include: n_{states}
- the range of timeslices used in the fit: $[t_{\text{min}}, t_{\text{max}}]$
- the blocksize (to avoid correlations between consecutive configurations)

We must adjust each of these parameters independently and check to confirm there is no bias in our results. The goal is to find a plateau — a range where adjusting the parameters does not greatly influence the fit result. I will demonstrate this procedure using the example of the local-local smearing of the pion propagator on the .01/.05 lattice (the same dataset shown in the effective mass plot of Figure 5.2).

Determining the Number of States to Include

Typical results showing a variation of n_{states} from 2 up to 7 states are shown in Figure 5.9. As the figure shows, the ground state is well determined, so even fits to only two or three states obtain the correct value. Unbiased determination of the first excited state, however, clearly requires $n_{\text{states}} \geq 4$. Further increasing the number of states does not improve the fit

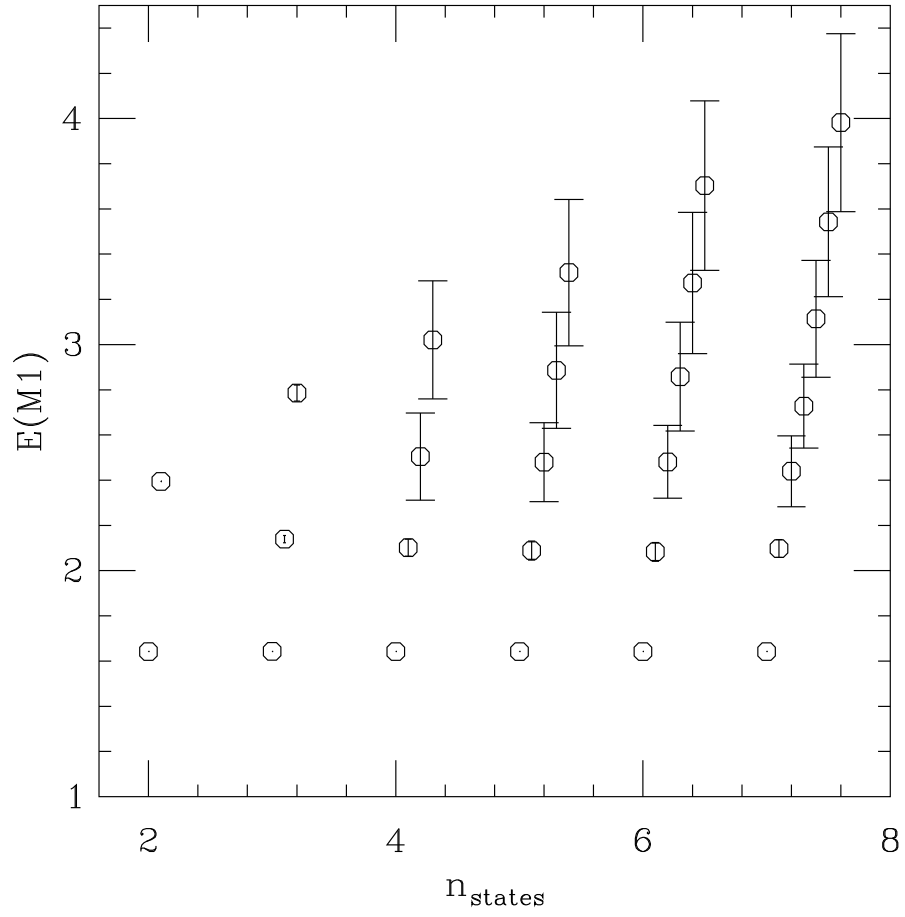


Figure 5.9: Energy levels vs. number of states for the local-local smearing of the pion propagator on the .01/.05 lattice. The ground state is stable for $n_{\text{states}} \geq 3$, while the first excited state requires $n_{\text{states}} \geq 4$ before it plateaus.

results, but dramatically increases the time taken for the fit (which increases as the square of the number of fit parameters), so we perform all fits to four energy levels.

Determining Which Timeslices to Use

As seen in the effective mass plot (Figure 5.2), the meson propagator includes information not only about the ground state, but also about many excited states. The propagator has the tightest errorbars (and therefore the most statistical significance) for small t ; as t gets larger the signal-to-noise ratio declines. At timeslice 0, however, *all* states contribute (since there is no exponential suppression of excited states). We refer to this condition as a “contact term”. Therefore there is no reason to include timeslice 0 in the fits — it cannot provide additional information. Because we want to extract the maximum information from the meson propagators, it is desirable to set t_{\min} as small as possible without introducing contact terms, so we expect that, in general, the best results will be obtained for $t_{\min} = 1$ (for S -wave states) or 2 (for P -wave states). Similarly, t_{\max} should be kept as large as possible, though if the effective mass plot shows excessive noise it may improve fitting speed to reduce t_{\max} to eliminate the noise. Plots showing fit results under a variation of t_{\min}, t_{\max} are not very illuminating — there is no discernible difference between fit results using a wide range of timeslices.

Determining the Blocksize to Avoid Correlations

Finally, it is necessary to ensure our results are uncorrelated. We accomplish this by blocking the results together, and then treating the blocks of averaged data as the new input dataset. If the correlation length fits within a block, then consecutive blocks will be statistically independent. Of course, when using a block size of N we have a factor of N fewer datapoints. Therefore it is expected that the statistical error will increase by \sqrt{N} . At the same time, however, the averaged data is closer to the mean than the raw data, and therefore the error decreases by a factor of \sqrt{N} . These two effects cancel for the case of uncorrelated data.⁴ If the underlying data were correlated, however, the averaged data would not be significantly

⁴This assumes that the number of datapoints is large compared to the block size, such that the number of blocks is much greater than unity.

closer to the mean than the raw data (because one high point is likely to be followed by another) and therefore we would see the errorbars grow as \sqrt{N} until all correlations have been taken into account. In our case, analysis shows the errorbars do not change in size, so we can safely assume our data is not correlated. Nevertheless, we standardize on a block size of 4 configurations just to be safe.

5.3.10 Splittings and Determination of a^{-1}

Repeating the process for each quantum state, we obtain energies of all particles in the spectrum. Using the strategy of Section 4.7.2 we determine the lattice spacing by comparing the $h_c - \overline{1S}$ splitting to experiment. A lattice spacing computed from the $\overline{2S} - \overline{1S}$ splitting is consistent. In addition, we compare to lattice spacings determined by the UKQCD collaboration from the analysis of the bottomonium spectrum on these lattices. They obtain very tight errorbars on the lattice spacing, as shown in Table 5.3, and are consistent with our findings. Because we can propagate errors through bootstrap only with lattice spacings we determine, we use our own lattice spacings for this analysis. A future goal is to standardize our methods so the lattice spacings from both collaborations can be interchangeably propagated through the bootstrap analysis.

We show mass splittings between charmonium mesons for a typical m_{sea} in Figure 5.10. The zero of the plot is the spin-averaged $1S$; the plot has been scaled such that the calculated result of the h_c matches experiment. By comparing the lattice $1P - \overline{1S}$ splitting with experiment we can set the scale of our calculation: the lattice spacing. For the case of the MILC coarse lattices we find $a^{-1} \approx 1.6 \text{ GeV}$, with no discernable dependence on m_{sea} (see Figure 6.1). See Table 5.3 for details on all lattices.

5.4 D_s Gives m_c

The next step in our procedure is to tune the charm-quark mass, m_c , through fits to the D_s spectrum. We choose the D_s for this tuning because the discretization effects are smaller for the D_s than for charmonium. This is a direct result of the $\mathcal{O}(\alpha_s a, a^2)$ errors present in our heavy-quark action.

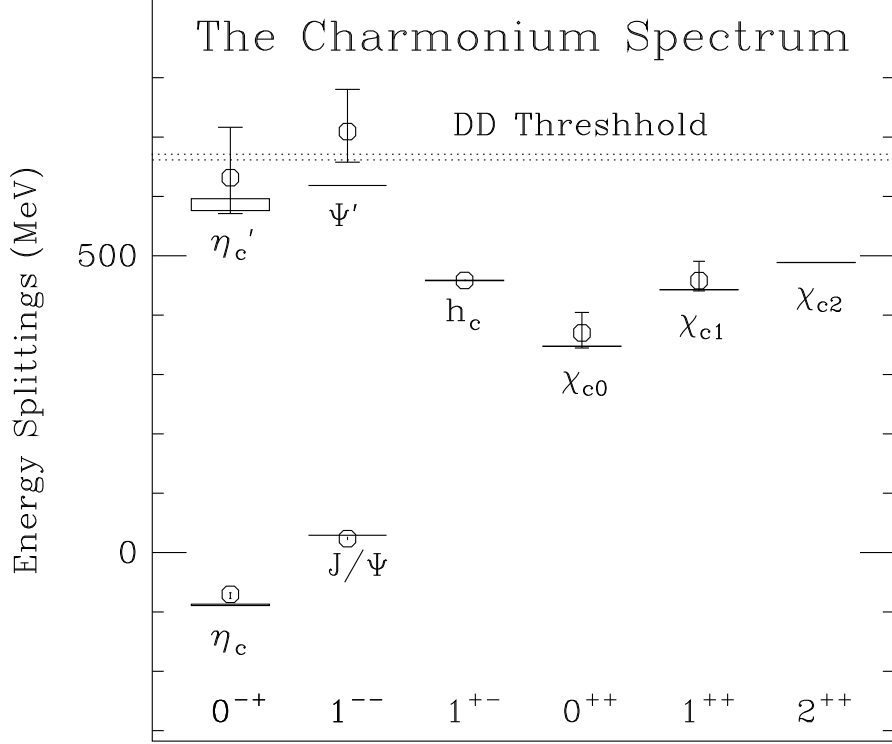


Figure 5.10: Charmonium splittings for the 01/05 lattice. The zero of the plot is the spin-averaged $1S$. The lines are experimental results, while the circles are the lattice result. The lattice results have been scaled so the h_c matches experiment.

Table 5.3: Lattice spacings measured on the MILC lattices. The lattice spacing determined from the $\psi(1P) - \psi(\overline{1S})$ splitting is part of this work; the spacings from the Υ spectrum courtesy of the UKQCD Collaboration [63]. For additional details see Ref. [64].

Lattice Size	$am_{u,d}/am_s$	a^{-1} (GeV) $\psi(1P) - \psi(\overline{1S})$	a^{-1} (GeV) $\Upsilon(2S) - \Upsilon(\overline{1S})$	a^{-1} (GeV) $\Upsilon(1P) - \Upsilon(\overline{1S})$
$16^3 \times 48$	0.0492/0.082			
$16^3 \times 48$	0.0328/0.082			
$16^3 \times 48$	0.0164/0.082			
$16^3 \times 48$	0.0082/0.0820		1.144(19)	1.128(19)
$20^3 \times 64$	0.050/0.05		1.645(46)	1.670(39)
$20^3 \times 64$	0.030/0.05	1.585(28)	1.714(52)	1.687(45)
$20^3 \times 64$	0.020/0.05	1.607(30)	1.605(20)	1.634(25)
$20^3 \times 64$	0.010/0.05	1.613(30)	1.596(25)	1.571(21)
$20^3 \times 64$	0.007/0.05	1.567(30)		
$24^3 \times 64$	0.005/0.05			
$28^3 \times 96$	0.0124/0.031		2.312(26)	2.390(38)
$28^3 \times 96$	0.0062/0.031		2.258(30)	2.305(45)

5.4.1 Introducing a Light Quark

When simulating charmonium, we had two charm valence quarks. For the D_s spectrum, one of the quarks is light. Simulating the light quark using the Wilson method is, as mentioned in Chapter 3, not very efficient. Instead, we use an improved staggered quark (the same treatment as the sea quarks). This has the advantage of having smaller discretization errors. Also, this is the same treatment we will use for the D , and therefore we can re-use much of the analysis.

5.4.2 The Heavy-Light Model Function

As explained in Section 3.5.1, the D_s meson propagators allow for both even- and odd-parity states. As a consequence, the model function we use to extract particle masses must be modified as in Eq. (4.5):

$$G_{\text{src,snk}}(t) = \sum_n Z_{\text{src}}^{(n)} Z_{\text{snk}}^{(n)} (-1)^{n(t+1)} \left(e^{-E_n t} + e^{-E_n(T-t)} \right). \quad (5.17)$$

Here, the parity of n forces the states to alternate between non-oscillating and oscillating.

Again, it is helpful to constrain the parameters in order to improve the fits. First, we note that there are two hierarchies of energies, one for the oscillating states, and one for the non-oscillating states. We know that the oscillating ground state will always have higher energy than the non-oscillating ground state.⁵ Using the notation that the non-oscillating ground state is E_0 and the oscillating ground state is E_1 , we enforce this hierarchy with the constraints:

$$\begin{aligned} E_1 &= E_0 + e^{\epsilon_1} \\ E_2 &= E_0 + e^{\epsilon_2} \\ E_3 &= E_1 + e^{\epsilon_3} \\ E_4 &= E_2 + e^{\epsilon_3} \\ &\vdots \end{aligned} \quad (5.18)$$

⁵The non-oscillating state has $J^P = 0^-$, and has orbital angular momentum $L = 0$ and spin angular momentum $S = 0$. The parity partner has $J^P = 0^+$, and has $L = 1$ and $S = 1$.

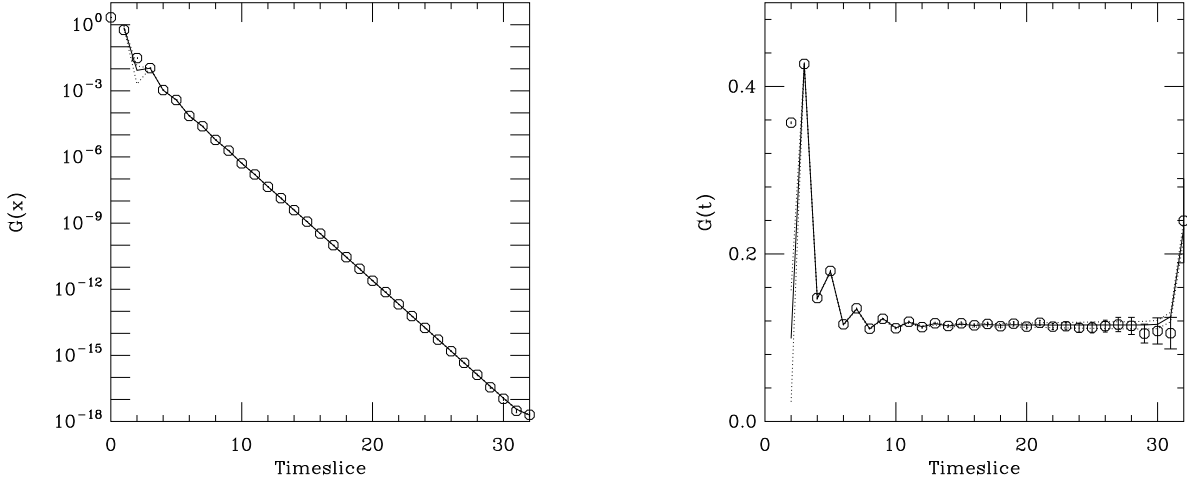


Figure 5.11: (a) A 4-state fit to the vector state of a charm-strange meson. Errorbars for both the data and the fit are plotted, but are too small to resolve. (b) The same fit, but shown with the leading exponential removed to show the remaining terms. Here the oscillations of the parity-partner state are visible.

$$E_i = E_{i-2} + e^{\epsilon_i}$$

where our fit parameters are E_0 and the ϵ_i . Note that while these constraints require the oscillating ground state to have higher energy than the non-oscillating ground state, the excited states may appear in any ordering. We also reuse Eq. (5.11) to reduce computational roundoff errors. Using this functional form, we obtain fits as seen in Figure 5.11. We obtain the D_s spectrum shown in Figure 5.12. Splittings that take into account both m_{sea} and m_{val} extrapolations are given in Table 6.2.

5.4.3 Kinetic Masses

Once we have the ability to fit the heavy-light propagators, our next step is to determine the kinetic mass of the D_s . This mass, compared with experiment, allows us to tune our lattice charm-quark mass. To accomplish this, we use the dispersion relation. Working in a finite box of side $L = na$, our momentum is discretized with Brillouin zones of width $2\pi/L$. As a result, it is simplest to refer to momenta in units of $2\pi/L$ (which provides an infrared (IR) cutoff) in each of the three spatial directions. A particle at rest would have a momentum of

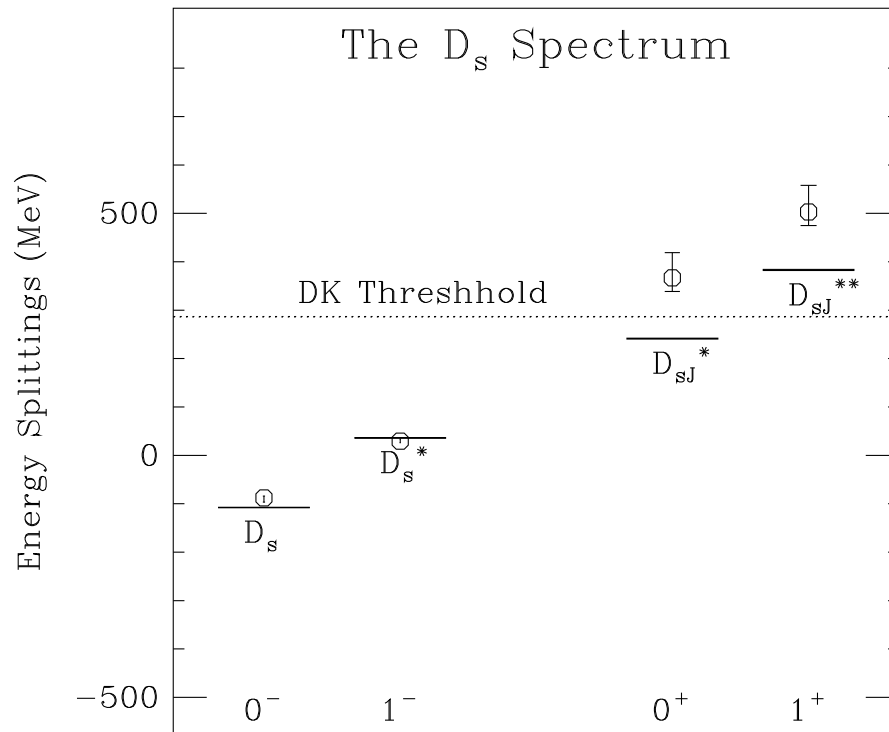


Figure 5.12: D_s splittings for the 01/05 lattice. The zero of the plot is the spin-average of the D_s and the D_s^* . The lines are experimental results, while the circles are the lattice result. Here the scale is determined by the lattice spacing found in Section 5.3.10.

(0,0,0), while the slowest-moving particle would have a momentum of (1,0,0).⁶ The highest momentum a particle can have is $2\pi/a$, which provides an ultraviolet (UV) cutoff.

We perform fits to the dispersion relation using momenta from the set shown in Table 5.4. The double-counting at a momentum of 3 gives us the opportunity to observe any effects of

Table 5.4: The momentum states used in the kinetic mass fits. The momentum $\mathbf{p} = (4, 0, 0)$ is only used when it improves the fit.

(p_x, p_y, p_z)	\mathbf{p}^2
(0 , 0 , 0)	0
(1 , 0 , 0)	1
(1 , 1 , 0)	2
(1 , 1 , 1)	3
(2 , 0 , 0)	4
(2 , 1 , 0)	5
(2 , 1 , 1)	6
(2 , 2 , 0)	8
(2 , 2 , 1)	9
(3 , 0 , 0)	9
(4 , 0 , 0)	16

our discrete rotational symmetry. We also use the momentum $\mathbf{p} = (4, 0, 0)$ when it improves the fit.⁷ In general, it will be informative to compare on-axis results with off-axis results, so they are plotted in different colors, as shown in Figure 5.13.

The fit to the dispersion relation of Eq. (3.44) gives us the kinetic mass of the particle. In this case, the fit gave a kinetic mass of 1.29(2), in lattice units. Using our determination of the lattice spacing, we find the lattice determination of the kinetic mass of the D_s is $1.91_{-0.08}^{+0.19}$ GeV. This is in good agreement with the experimental measurement of 1.968 GeV.

5.4.4 Tuning m_c

We use the $am_l/am_s = .01/.05$ configuration to tune the charm quark mass, as that configuration has the most lattices available (and therefore the best statistical significance). For $\kappa = 0.119$ our calculation of m_{D_s} was slightly low. We therefore reduce the parameter to $\kappa = 0.117$ and repeat the calculation of Section 5.4.3. As it turns out, the value $\kappa_{\text{charm}} = 0.119$

⁶The spatial dimensions are interchangeable, so we need not distinguish between elements of the set $\{(1,0,0), (0,1,0), (0,0,1)\}$.

⁷Errorbars increase at larger momentum, as seen in Figure 5.13, so this extra point has minimal effect on the fit results.

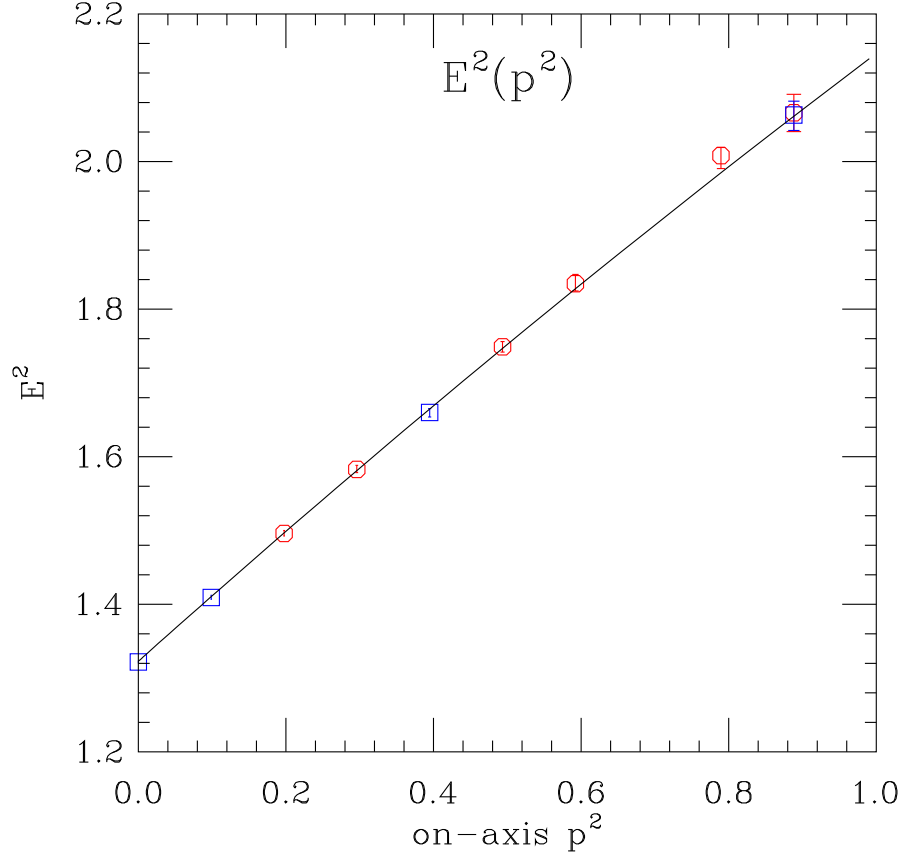


Figure 5.13: A fit to the dispersion relation of Eq. (3.44). Data for on-axis momenta are shown in blue squares, while off-axis momenta are shown in red circles. The alignment of on-axis momenta with off-axis momenta indicates the rotational symmetry violations are small. The fit had a χ^2 of 6.6 for 6 degrees of freedom (10 data points – 4 parameters in the fit).

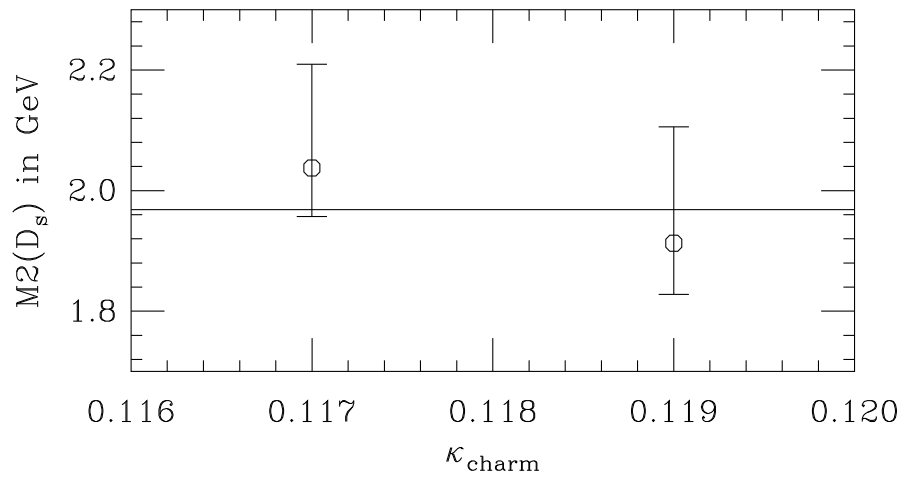


Figure 5.14: The mass of the D_s for two different values of κ_{charm} on the 01/05 lattice. The horizontal line indicates the experimental result.

Table 5.5: Rest and kinetic masses for heavy-heavy and heavy-light mesons on the 01/05 lattice.

	HH	HL
aM_1	1.6897(1)	1.2089(16)
aM_2	2.2626(150)	1.3689(200)
$a\delta M$	0.5729(150)	0.1600(201)

gave a better approximation to the physical mass of the D_s , and was selected as the final value.

An Inconsistency Between HH and HL Results

As explained in Section 4.7.2, we use the heavy-light (HL) D_s system to tune m_c rather than the heavy-heavy (HH) charmonium system because the HL system has smaller discretization effects. Although we can't expect an exact tuning of κ_{charm} from charmonium, we might expect the results to be similar. It therefore came as a surprise that the charmonium $M_2(\overline{1S})$ comes out 10-20% high⁸ (as compared to experiment) with only 1% errors. We know we should expect some discretization error from lattice artifacts, but how much?

The discrepancy between tunings for the heavy-quark mass was seen previously by a group studying the Υ and B systems in quenched lattice QCD [65]. They note that the “inconsistency”

$$I(\text{H}, \text{L}) = \frac{\delta M(\text{HL}) - [\delta M(\text{HH}) + \delta M(\text{LL})]/2}{M_2(\text{HL})} \quad (5.19)$$

where

$$\delta M \equiv M_2 - M_1 \quad (5.20)$$

is 0 for $M_2(\text{HL}) < 1$, and goes negative linearly for $M_2(\text{HL}) > 1$. In our case, fitting the kinetic mass on the 01/05 lattice using momenta up to $\mathbf{p} = (3, 0, 0)$ ⁹ gives the results of Table 5.5.

The dependence on am_{sea} is negligible. Ignoring the kaon (which is light) we find

⁸Results were 10% high using our lattice spacing from the charmonium $1P - \overline{1S}$ splitting, and 20% high using the lattice spacing from the UKQCD analysis of the Υ spectrum.

⁹Restricting the HL kinetic mass fits to momenta up to $\mathbf{p} = (2, 0, 0)$ gives a slightly worse quality-of-fit, and slightly larger errorbars, but the central value of the result is the same. Restricting the HH kinetic mass fits in the same way has negligible effect. In both cases, restricting the fit to only the linear term in \mathbf{p}^2 gives a poor quality of fit.

$$I = \frac{0.1600(201) - \frac{1}{2}0.5729(150)}{1.3689(200)} = \frac{-0.1265(215)}{1.3689(200)} = -0.0924(158) \text{ at } aM_2(cs) = 1.3689(200). \quad (5.21)$$

This is similar to the inconsistency seen in Ref. [65]. We therefore find that the inconsistency is independent of quenching effects.

A complete analysis of the expected inconsistency arising from discretization effects was given by Kronfeld [66]. It is expected that the inconsistency be approximately

$$I \approx -\frac{\delta B_{QQ}}{2M_{Qq}} \quad (5.22)$$

where δB is the difference in binding energy between the kinetic mass and the rest mass:

$$\begin{aligned} M_{1\bar{Q}q} &\equiv M_{1\bar{Q}} + M_{1q} + B_{1\bar{Q}q} \\ M_{2\bar{Q}q} &\equiv M_{2\bar{Q}} + M_{2q} + B_{2\bar{Q}q} \end{aligned} \quad (5.23)$$

$$\delta B_{\bar{Q}q} = B_{2\bar{Q}q} - B_{1\bar{Q}q} \quad (5.24)$$

For an S -wave state, the binding energy difference can be predicted from

$$\frac{\delta B_{Qq}}{\langle T \rangle} = \frac{1}{3} \left\{ 5 \left[\mu_2 \left(\frac{M_{2Q}^2}{M_{4Q}^3} + \frac{M_{2q}^2}{M_{4q}^3} \right) - 1 \right] + 4a^2 \mu_2 (M_{2Q}^2 w_{4Q} + M_{2q}^2 w_{4q}) \right\}, \quad (5.25)$$

where $\langle T \rangle$ is the kinetic energy of the meson, and $\mu_2 = \frac{1}{2} (M_{2Q} + M_{2q})$ is the kinetic reduced mass of the constituent quarks. In our case, $\delta B/\langle T \rangle = 1.049$. Using μ and \mathbf{p}^2 from our potential model calculation (Tables A.1 and A.2) we can estimate the kinetic energy

$$\langle T \rangle = \frac{\langle \mathbf{p}^2 \rangle}{2\mu} = \frac{0.5224 \text{ GeV}^2}{1.4935 \text{ GeV}} = 0.350 \text{ GeV} \quad (5.26)$$

so $\delta B = 0.367 \text{ GeV}$. The error on $\langle \mathbf{p}^2 \rangle$ (and on μ) is a few percent (it is only a potential model calculation, not a simulation of QCD). We therefore expect the inconsistency to be

$$I \approx -\frac{\delta B_{QQ}}{2M_{Qq}} = -\frac{\delta B_{QQ}}{2a^{-1}(aM_{Qq})} = -\frac{0.367 \text{ GeV}}{2(1.6 \text{ GeV})(1.3689)} = -0.084. \quad (5.27)$$

This expectation is slightly larger than the results found by others, but is in agreement with our measured result.

We expect the inconsistency I to decrease when a heavy-quark action that eliminates all $\mathcal{O}(a)$ discretization errors becomes available.

5.4.5 The D_s Decay Constant

The extraction of the decay constant was described in Section 3.5.2. After performing the fits of Eqs. (3.113) and (3.117) to obtain the coefficients $|Z_{D_s}|^2$ and $Z_{A_4}Z_{D_s}$, it is clear that we may extract

$$Z_{A_4} = \frac{Z_{A_4}Z_{D_s}}{\sqrt{|Z_{D_s}|^2}} = \frac{\langle 0|A_4|D_s\rangle\langle D_s|\bar{s}\gamma_5 c|0\rangle}{\sqrt{|\langle D_s|\bar{s}\gamma_5 c|0\rangle|^2}}. \quad (5.28)$$

Here Z_{A_4} is a lattice quantity. To compare to the continuum quantity $f\sqrt{m}$ we must account for the lattice action normalization conventions. We find

$$f\sqrt{m} = Z_A^{hl}\sqrt{2u_0}\sqrt{1-6\tilde{\kappa}}\sqrt{2}Z_{A_4} \quad (5.29)$$

where Z_A^{hl} is a current renormalization factor from perturbative matching, $\sqrt{2u_0}$ accounts for MILC's staggered quark normalization, and $\sqrt{1-6\tilde{\kappa}}$ is the normalization for the Wilson quark in the Fermilab interpretation (see text surrounding Eq. (3.53)). The final factor of $\sqrt{2}$ comes from the normalization convention

$$\langle 0|\bar{c}\gamma_5 s|D_s\rangle\langle D_s|\bar{s}\gamma_5 c|0\rangle = 2m_{D_s}. \quad (5.30)$$

We break the calculation of Z_A^{hl} into a flavor-conserving part and a nonperturbative correction as [67, 68]:

$$Z_A^{hl} = \rho_A\sqrt{Z_V^{hh}Z_V^{ll}}. \quad (5.31)$$

The flavor-conserving renormalization factors were computed nonperturbatively [69]:

$$Z_V^{hh} = 1.33(2); \quad Z_V^{ll} = 0.86(5). \quad (5.32)$$

The parameter ρ_A includes radiative corrections, and is expected to be close to unity [70]. Indeed, a one-loop calculation gives $\rho_A = 1.044$ on the coarse ensemble [71].

We find, after conversion to physical units, $f_{D_s} \approx 250$ MeV. Again, the sea quark extrapolation is insignificant. We will, however, improve upon this result in later sections, after we have built the necessary formalism to interpolate to the physical valence quark mass.

5.5 Extrapolation to the D

Our treatment of the D is almost identical to our treatment of the D_s , except that we cannot simulate the valence light quark mass due to computational slowdown as $m_q \rightarrow 0$ (see Section 4.8.3). As a result, it is necessary to perform a chiral extrapolation down to the physical light quark mass.

5.5.1 Chiral Extrapolations in PQQCD

Ordinarily a chiral extrapolation contains a single parameter: the mass of the light quark. In our case, however, it is evident that most of the computational cost comes from varying the sea quark mass, while variations in the valence quark mass are comparatively inexpensive. It is therefore advantageous to vary the sea- and valence-quark masses independently — a technique called partially quenched QCD (PQQCD). Full QCD, in which the sea- and valence-quark masses are equal, can be thought of as a special case of PQQCD.

The theory is slightly complicated by the partially-quenched generalization because there are now two *different* species of light quarks that can run around inside loops. For this reason, there are additional parameters in our extrapolations.

5.5.2 Mass Splittings

We do not have a complete chiral perturbation theory (χ PT) for the masses of the D mesons, and are therefore forced to make an educated guess about how to attempt the extrapolations. The linearity of the extrapolation in m_{val} is evident from Figure 5.15. Unfortunately, the extrapolation in m_{sea} is nonlinear, so we cannot comfortably extrapolate to the physical meson masses. Instead, we content ourselves with extrapolations of mass splittings, as most nonlinear

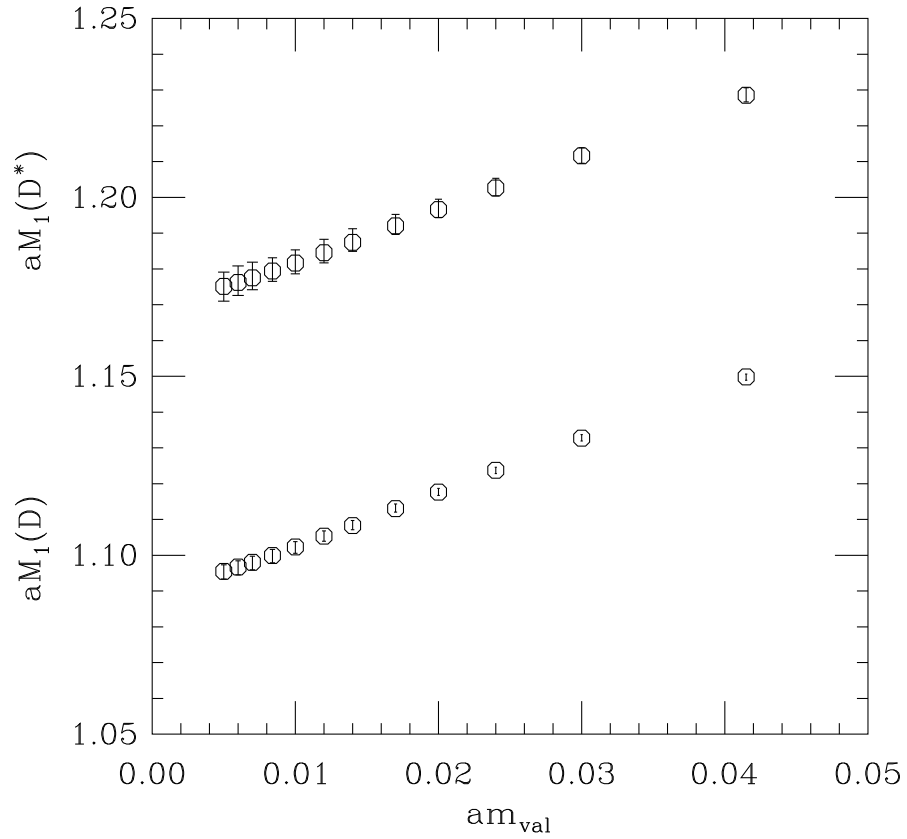


Figure 5.15: A m_{val} extrapolation in the D system, for the case of the 01/05 lattice. It is clear that, at accessible values of m_{val} , the extrapolation is linear. Furthermore, it is apparent that the same slope applies to both the D and the D^* .

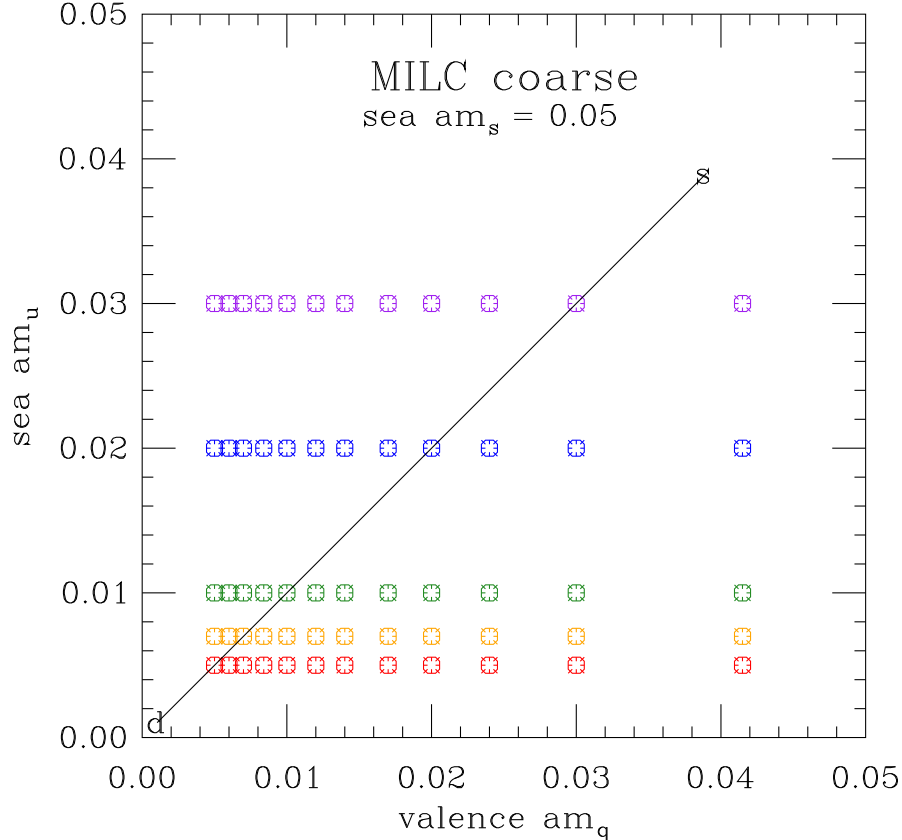


Figure 5.16: The selection of sea- and valence-quark masses we used for our PQS χ PT extrapolation. Points of the same color come from the same lattice, and are expected to be highly correlated. Points of different colors are expected to be uncorrelated. The line indicates “full” QCD, where $m_{\text{sea}} = m_{\text{valence}}$.

effects should cancel. We can therefore obtain a reasonable extrapolation by considering terms linear in both m_{val} and m_{sea} :

$$m_{D^*} - m_D \approx Am_{\text{val}} + Bm_{\text{sea}} + C \quad (5.33)$$

and so on for other splittings. It is clear from Figure 5.15 that, for the 01/05 lattice, $aM_1(D^* - D) = 0.079$, which corresponds to $M_1(D^* - D) \approx 127$ MeV.

We can obtain an extrapolation for all lattices by performing a single combined fit to a large parameter space. The first step is to perform bootstrap fits using several valence quark masses on each of several lattices that have different sea quark masses. Our selection of data is shown in Figure 5.16. Bootstrap data from each of these points in the input parameter

phase-space is used in a single combined fit.

Because we are fitting a 2-dimensional function to 60 data points, the resulting fits are difficult to understand, and not very informative. We simply summarize our results in Table 6.3 and note that each fit had $\chi^2 < 1$.

5.5.3 Staggered Chiral Perturbation Theory

Now let us continue on to investigate the full χ PT for the D decay constant. The situation is complicated further because we have staggered quarks. As described in Section 3.3.2, the staggered quark formalism leaves us with four tastes of quarks. We have to take this into account in any extrapolation we perform. In addition, we must account for the $\mathcal{O}(a^2)$ errors from taste-changing interactions.

The resulting theory is referred to as staggered χ PT, or S χ PT. Like the usual χ PT, we fit to the form

$$f_{Q_q} \sqrt{M_{Q_q}} = \alpha \left(1 + \frac{1}{16\pi^2 f^2} \Delta f_q + \dots \right). \quad (5.34)$$

where the Δf_q contains the chiral logs and the \dots are terms analytic in m_q and $\sum_f m_f$. In our case, however, the form of the chiral logs is complicated by the different tastes of fermions. Our fits use a functional form provided by Aubin and Bernard [72], which is summarized in Appendix E.

We are fitting a 2-dimensional function to 60 data points, making the resulting fit too complicated to understand from a single plot. Instead I show slices through the input parameter space. It is important to remember, however, that all of these slices came from the *same* fit. A slice through parameter space corresponding to full QCD is shown in Figure 5.18. This extrapolation gives our best determination of the D decay constant, $f_D = 237_{-4}^{+14}$ MeV.

The added information from this extrapolation also allows us to obtain a slight improvement on our previous results for the D_s decay constant. The results of Section 5.4.5 were restricted to a single value of the strange quark mass, $am_s = 0.0415$. We can compare to those results by taking a slice through the current fit where the valence quark mass is held constant and only the sea quark mass can vary. Of course, hindsight tells us that strange quark mass is not exact. We can use the results obtained from our PQS χ PT extrapolation to

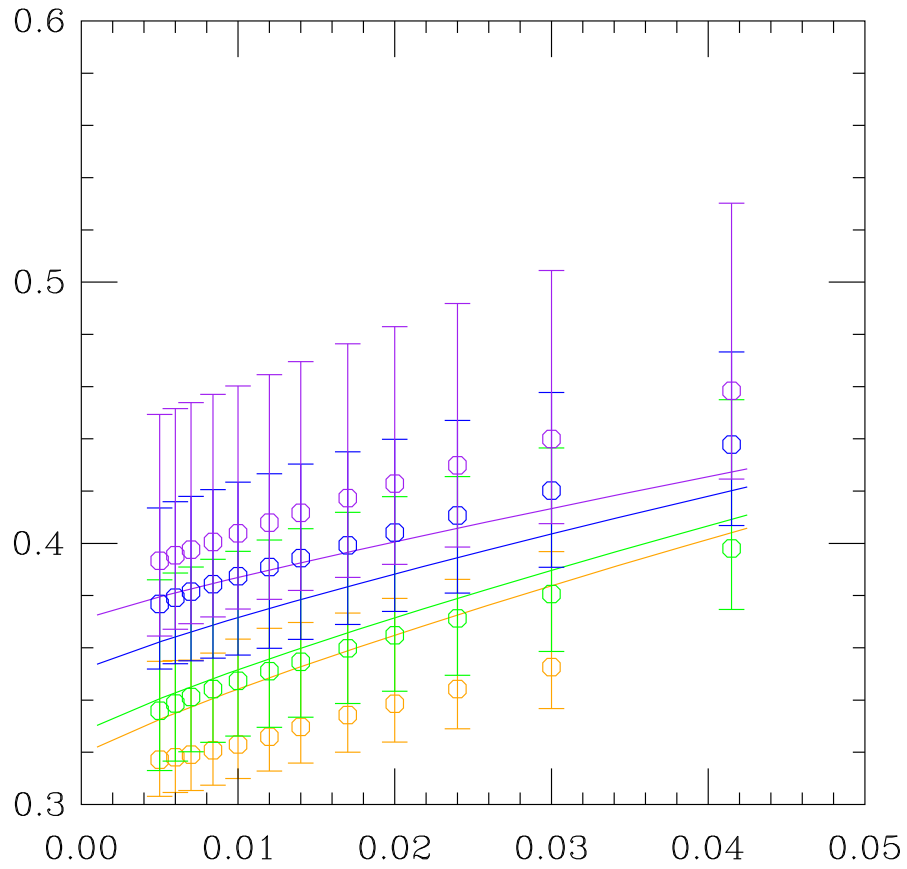


Figure 5.17: Slices through the $\text{PQS}\chi\text{PT}$ extrapolation. Correlated fits for each am_{sea} are shown.

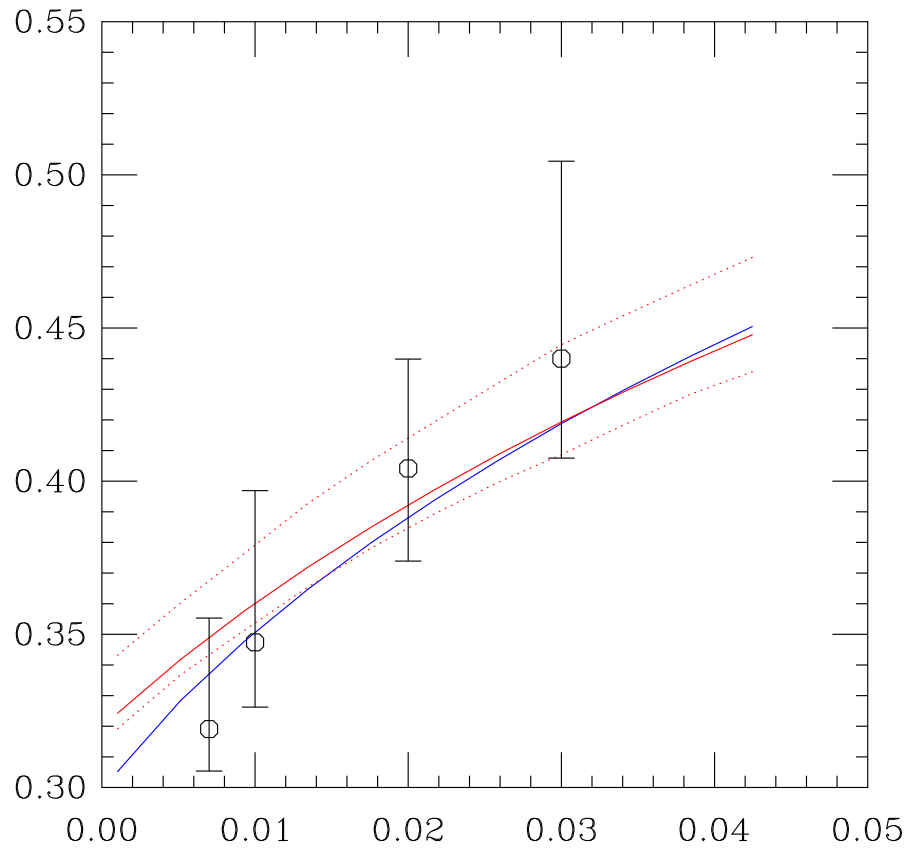


Figure 5.18: A slice through our PQS χ PT extrapolation corresponding to full QCD. Only the data corresponding to full QCD is shown, but the fit was to the full set of partially-quenched data. The blue line indicates the fit, including taste violations. The red line (and dotted errors) shows the prediction when taste violations have been turned off.

interpolate to the correct strange quark mass. The final result for the D_s is therefore adjusted slightly, to $f_{D_s} = 289_{-5}^{+15}$ MeV.

Finally, it is helpful to realize that experimentalists might be better served by knowing a very precise prediction of the ratio

$$R_{q/s} \equiv \frac{f_D \sqrt{m_D}}{f_{D_s} \sqrt{m_{D_s}}}. \quad (5.35)$$

We can determine this ratio better than the individual decay constants because their statistical errors are largely correlated and therefore cancel out. Additionally, taking the ratio cancels some of the parameters in the PQS χ PT, therefore allowing for a somewhat more constrained fit. A plot showing a slice through the full QCD parameter space of the extrapolation is given in Figure 5.19. One can see from the plot that we find $R_{q/s} = 0.834_{-0.014}^{+0.011}$ (statistical errors only) in the chiral limit.

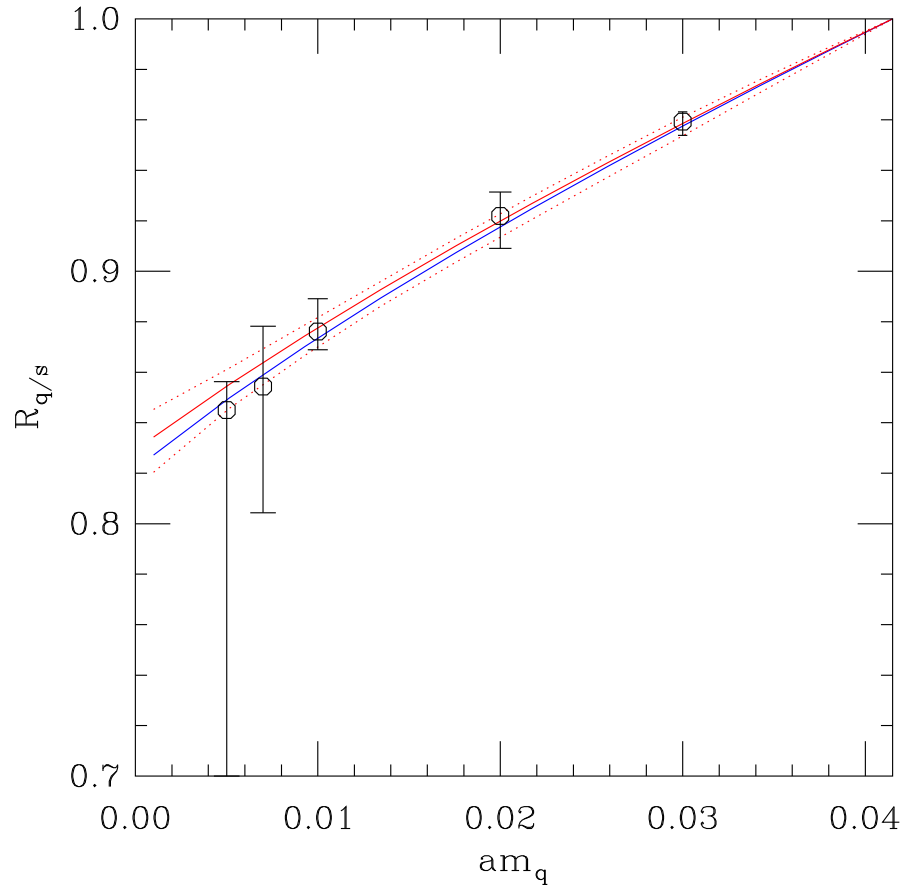


Figure 5.19: A slice corresponding to full QCD through the PQS χ PT extrapolation for the f_D/f_{D_s} ratio. Only the data corresponding to full QCD is shown, but the fit was to the full set of partially-quenched data. The blue line indicates the fit, including taste violations. The red line (and dotted errors) shows the prediction when taste violations have been turned off.

Chapter 6

Results

After performing the meson propagator fits, there are a few remaining tasks before we can compare to experiment. As we have already seen, it is necessary to use our value for a^{-1} to convert all lattice results into physical units, such as GeV. We must also perform any necessary chiral extrapolation, using the techniques of Sections 5.5.1 and 5.5.3. Also, to eliminate any artifacts from the use of a finite lattice spacing, it is important to repeat the calculations at different lattice spacings, and perform an extrapolation to the continuum limit. Finally, there may be some remaining uncertainties, such as finite volume effects, that can be addressed by performing simulations on lattices of different physical size and extrapolating to the infinite-volume limit. When quoting results, we must also consider any systematic errors from the tuning of lattice parameters.

6.1 The Charmonium System

After performing fits to the $c\bar{c}$ meson propagators, we have extracted mass splittings and kinetic masses of the $J = 0$ or $J = 1$ particles below threshold. By using the $1P - \overline{1S}$ splitting, we find the lattice spacings indicated in Figure 6.1. We see from the Figure that the lattice spacing is held roughly constant over the range of sea quark masses in the study.

I show the lattice prediction of the spectrum of mass splittings (to the $\overline{1S}$) in Figure 6.2. Numerical data for the splittings, including hyperfine splittings, are given in Table 6.1. As can be seen in Figures 6.3–6.6, the sea quark effects are mild.

To determine the offset for our energy splittings, I calculate the kinetic mass of the $\overline{1S}$

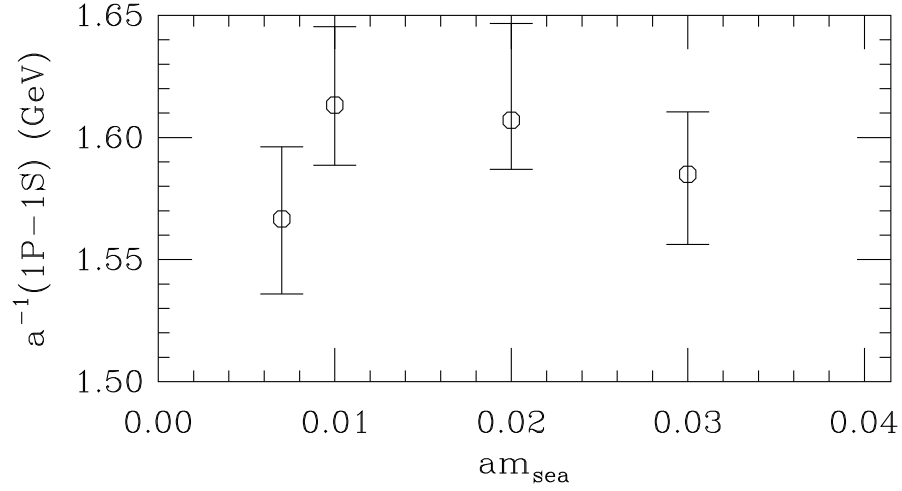


Figure 6.1: The inverse lattice spacings for the coarse lattices. The results for a^{-1} are 1.56(3), 1.61(3), 1.61(3), and 1.58(3) for am_{sea} of 0.007, 0.01, 0.02, and 0.03.

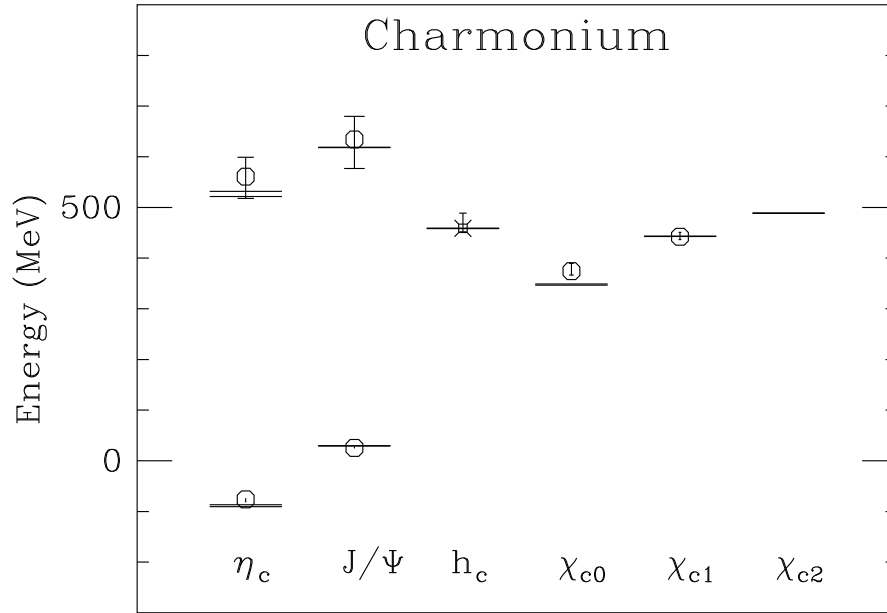


Figure 6.2: The charmonium spectrum shown in terms of mass splittings. The zero of the plot is the spin-averaged mass of the η_c and the J/ψ . Lines are the experimental observations, while circles indicate the lattice result.

Table 6.1: Mass splittings in the charmonium spectrum.

Quantity	Lattice prediction	Experiment	Plot
$h_c - \overline{1S}$	input	459 MeV	N/A
$\overline{2S} - \overline{1S}$	587^{+46}_{-50} MeV	610 MeV	Figure 6.3
$J/\psi - \eta_c$	100^{+4}_{-3} MeV	117 MeV	Figure 6.4
$\psi' - \eta'_c$	43^{+74}_{-66} MeV	32 MeV	Figure 6.5
$(\psi' - \eta'_c)/(J/\psi - \eta_c)$	0.4 ± 0.7	0.27	Figure 6.6
$h_c - \chi_{c0}$	81^{+29}_{-29} MeV	111 MeV	Figure 6.7
$h_c - \chi_{c1}$	10^{+23}_{-25} MeV	16 MeV	Figure 6.8

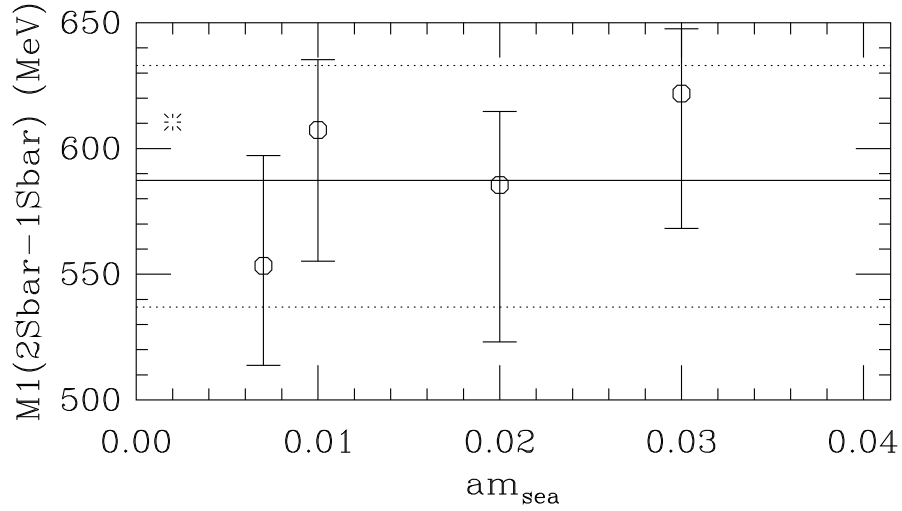


Figure 6.3: The $\overline{2S} - \overline{1S}$ splitting. Lattice values are shown at four values of m_{sea} , with a fit to a constant shown as a line with dotted errors. The experimentally-observed value is shown as a burst. The lattice prediction of $\overline{2S} - \overline{1S} = 587^{+46}_{-50}$ MeV is in agreement with the experimentally observed value of 610 MeV.

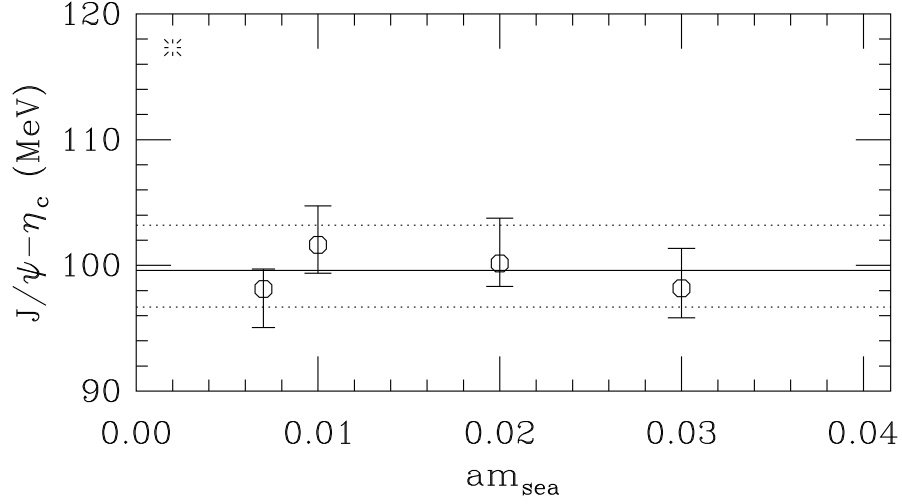


Figure 6.4: The $J/\psi - \eta_c$ hyperfine splitting. Lattice values are shown at four values of m_{sea} , with a fit to a constant shown as a line with dotted errors. The experimentally-observed value is shown as a burst. The lattice prediction of $J/\psi - \eta_c = 100_{-3}^{+4}$ MeV is significantly less than the experimentally observed value of 117 MeV. The discrepancy is most likely due to the $\mathcal{O}(\alpha_s a)$ discretization errors of the heavy quark action.

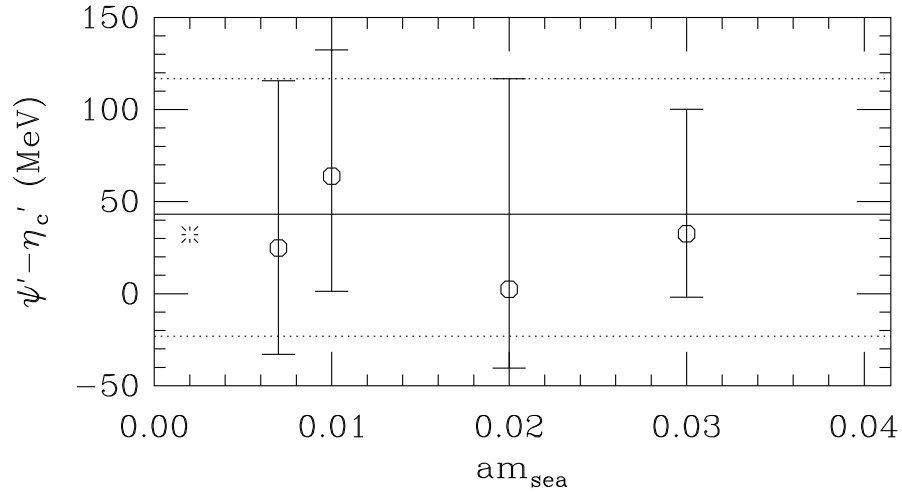


Figure 6.5: The $\psi' - \eta'_c$ hyperfine splitting. Lattice values are shown at four values of m_{sea} , with a fit to a constant shown as a line with dotted errors. The experimentally-observed value is shown as a burst. The lattice prediction of $\psi' - \eta'_c = 43_{-66}^{+74}$ MeV is consistent with the experimentally observed value of 32 MeV. The large errors on the lattice prediction are due to the lack of statistical significance for the excited state fits.

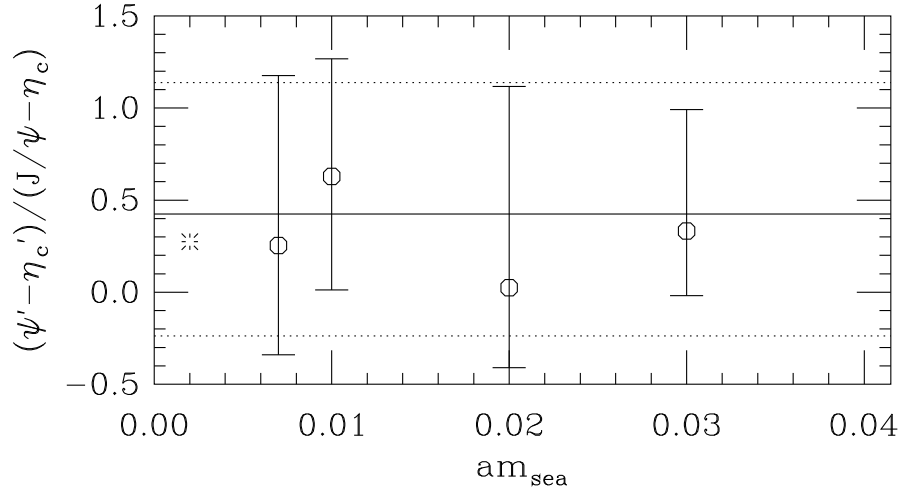


Figure 6.6: The ratio of hyperfine splittings: $(\psi' - \eta'_c)/(J/\psi - \eta_c)$. Lattice values are shown at four values of m_{sea} , with a fit to a constant shown as a line with dotted errors. The experimentally-observed value is shown as a burst. The lattice prediction of $(\psi' - \eta'_c)/(J/\psi - \eta_c) = 0.4 \pm 0.7$ is consistent with the experimentally observed value of 0.27. The large errors on the lattice prediction are due to the lack of statistical significance for the excited state fits, however some systematic errors cancel in the ratio, making this a somewhat more accurate prediction than the individual hyperfine splittings.

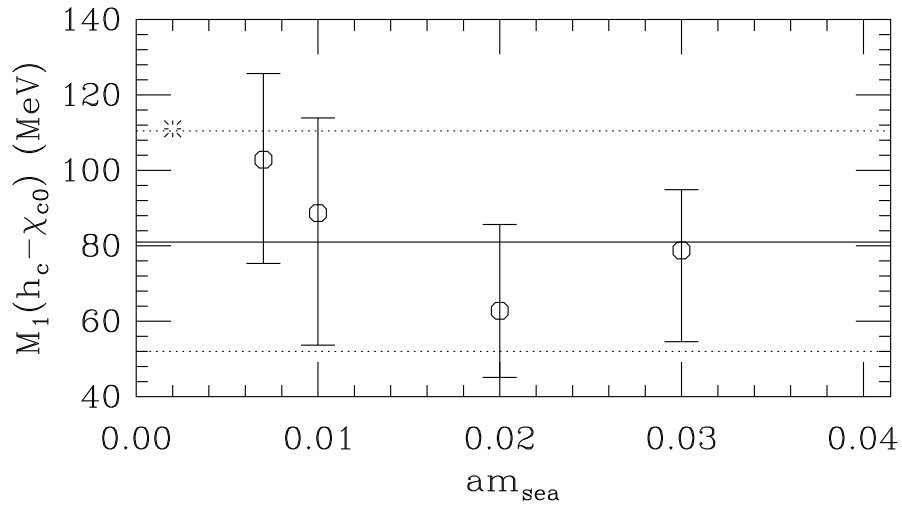


Figure 6.7: The $h_c - \chi_{c0}$ splitting, which tests the fine splitting resulting from the chromoelectric interaction. Lattice values are shown at four values of am_{sea} , with a fit to a constant shown as a line with dotted errors. The experimentally-observed value is shown as a burst. The lattice prediction of $h_c - \chi_{c0} = 81^{+29}_{-29}$ MeV is consistent with the experimentally observed value of 111 MeV. It is expected that reducing the $\mathcal{O}(\alpha_s a)$ discretization errors of the heavy quark action will reduce these errors.

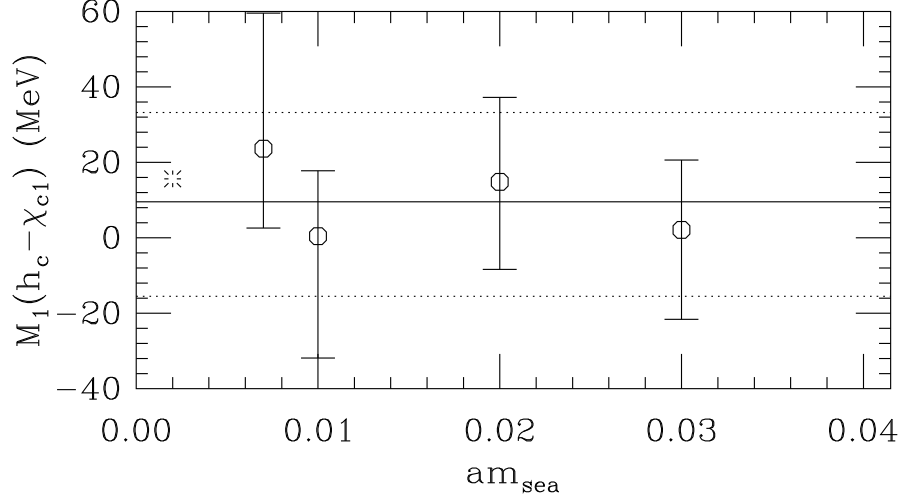


Figure 6.8: The $h_c - \chi_{c1}$ splitting, which tests the fine splitting resulting from the chromoelectric interaction. Lattice values are shown at four values of am_{sea} , with a fit to a constant shown as a line with dotted errors. The experimentally-observed value is shown as a burst. The lattice prediction of $h_c - \chi_{c1} = 10^{+23}_{-25}$ MeV is in agreement with the experimentally observed value of 16 MeV. It is expected that reducing the $\mathcal{O}(\alpha_s a)$ discretization errors of the heavy quark action will reduce these errors.

spin-averaged state. This is done by fitting to the dispersion relation shown in Eq. (3.44). A plot showing the result for each value of m_{sea} is given in Figure 6.9. We find $M_2(\overline{1S}) = 3.57^{+0.11}_{-0.08}$ GeV, which is about 16% higher than the experimental value of 3.07 GeV. This inconsistency is an expected result of the discretization errors of the heavy-quark action, as explained in Section 5.4.4. Future improvements in the heavy-quark action (to remove the $\mathcal{O}(a^2)$ errors) will be needed to eliminate this inconsistency. The spin-averaged splittings are not affected by the $\mathcal{O}(\alpha_s a)$ discretization errors and have uncertainties at the 5% – 10% level. Other splittings have larger errors, giving 25% or larger uncertainties. These results are part of a larger project which includes additional lattice spacings [73].

6.2 The D_s and D Systems

Splittings for the D_s system are given in Table 6.2, and for the D system in Table 6.3. In each case, the splittings are at the 5% – 10% level. We do not have an absolute scale for these quantities as the functional form of the chiral expansion which takes into account staggered quarks is not yet known. The splittings to the $D_{sJ}^*(2317)$ is of particular interest

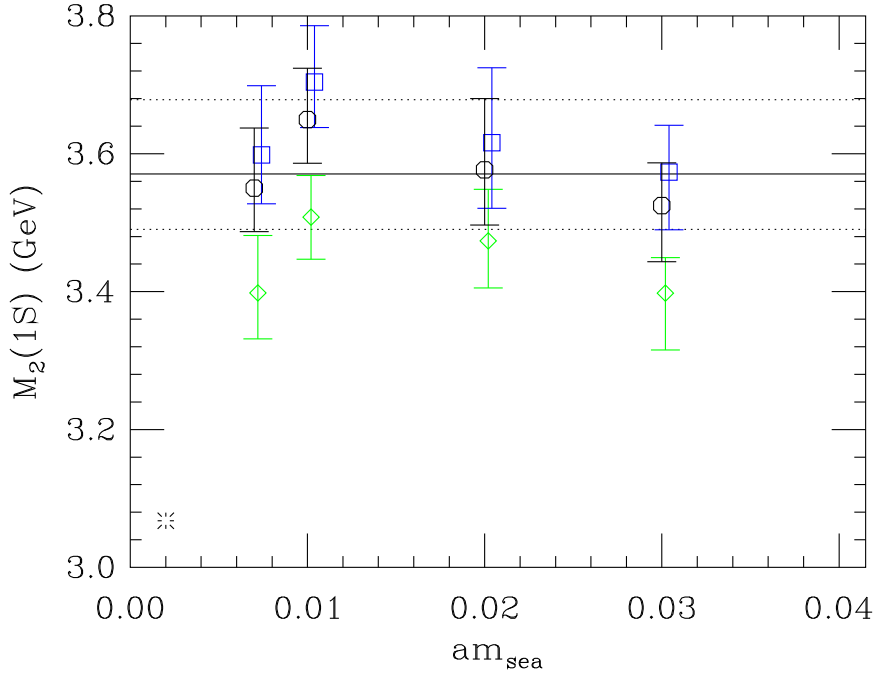


Figure 6.9: The kinetic mass of the charmonium $1S$. The lattice values for the η_c are shown in green diamonds and the values for the J/ψ in blue squares. The black circles indicate the $\overline{1S}$ (the spin-average of the η_c and the J/ψ). The deviation from experiment is not unexpected — see the text.

Table 6.2: Calculated splittings in the D_s system. The $D_s^* - D_s$ splitting is in agreement with experiment; the $D_{sJ}^* - D_s$ and $D_{sJ}^{**} - D_s^*$ splittings are larger than what has been observed.

Quantity	Lattice prediction	Experiment
$D_s^* - D_s$	129^{+9}_{-7} MeV	144 MeV
$D_{sJ}^* - D_s$	483^{+30}_{-25} MeV	349 MeV
$D_{sJ}^{**} - D_s^*$	469^{+30}_{-24} MeV	347 MeV

Table 6.3: Calculated splittings in the D system. The $D^* - D$ splitting is in agreement with experiment; the $D_1 - D$ and $D_2^* - D^*$ splittings are a bit low.

Quantity	Lattice prediction	Experiment
$D^* - D$	129^{+9}_{-7} MeV	141 MeV
$D_1 - D$	483^{+30}_{-25} MeV	558 MeV
$D_2^* - D^*$	469^{+30}_{-24} MeV	596 MeV

Table 6.4: Decay constants f_{D_s} and f_D compared to experiment.

Quantity	Lattice prediction	Experiment
f_{D_s}	289_{-5}^{+15} MeV	267 ± 33 MeV
f_D	237_{-4}^{+14} MeV	$223 \pm 16_{-9}^{+7}$ MeV
$\frac{f_D \sqrt{m_D}}{f_{D_s} \sqrt{m_{D_s}}}$	$0.834_{-0.014}^{+0.011}$	~ 0.814

because the $J^P = 0^+$ still needs confirmation, and because the mass of 2317 MeV had not been predicted. We find the mass of this state to be larger than the observed mass, but our systematic uncertainties are too large to draw any conclusions. It is interesting to note, however, that our results are in agreement with a previous lattice calculation [74].

The decay constants for the D_s and D are compared to experiment in Table 6.4. The decay constants are determined at the 10% level — the experimental determination for f_D is comparable, and we await an improved determination of f_{D_s} from the CLEO-c collaboration. These results are part of a larger project which includes additional lattice spacings [75, 76].

Chapter 7

Conclusions

Lattice QCD is a maturing field. The asqtad action treats light quarks with $\mathcal{O}(\alpha_s a^2, a^4)$ errors, the clover action in the Fermilab interpretation treats heavy quarks with $\mathcal{O}(\alpha_s a, a^2)$ errors, and the improved gauge action has $\mathcal{O}(\alpha_s a^2)$ errors. These lattice actions can already be used to calculate many quantities, including spectra of stable particles and decay constants. All errors are now controlled, and can be estimated or eliminated.

This work presented an analysis of charmonium and charmed mesons from the lattice perspective. The calculations have uncertainties at the 5% – 10% level. Because errors are controlled, we have been able to directly compare lattice results to experimental measurements. The agreement seen by this comparison gives confidence that future results will match or predict [77] experimental measurements. Having high confidence in quantitative predictions of QCD is an important step in searching for physics beyond the Standard Model.

It is important to recognize, however, that our work is not yet complete. Current lattice computations must be carefully tuned to optimize the delicate balance between box size (L), lattice spacing (a), the light-quark mass (m_q), and statistics. A short-term solution is to build tera-scale compute clusters or special purpose machines [56, 57]. However, because of the severe power-laws each of these quantities obeys, the future of the field depends on improving the lattice actions to higher order. Our largest discretization errors arise from the heavy-quark action, which has $\mathcal{O}(\alpha_s a, a^2)$ errors. The development of an $\mathcal{O}(a^2)$ -improved asymmetric heavy-quark action is already in progress (see Appendix F). Other groups are investigating still other ideas, including domain wall, overlap, and twisted mass fermions. With future actions it will be possible to reduce both systematic and statistical uncertainties,

and investigate more difficult systems. Current heavy-light analyses are limited by statistical uncertainties. Increased statistics will allow us to better determine the parameters of the chiral extrapolation, which is currently a large systematic uncertainty. Also, improved actions will reduce the heavy-quark systematic uncertainties. It is expected that heavy-light chiral extrapolations will soon approach the few-percent level.

Appendix A

Estimation of Operator Effects

The coefficients of our lattice actions are short-distance quantities. Therefore they can be determined perturbatively, first at tree level, then with tadpoles, and finally with a full one-loop calculation. However, for high levels of improvement the operators and coefficients are not (yet) known. Our current lattice action for heavy quarks is the clover action, which is improved to $\mathcal{O}(\alpha_s a, a^2)$. We would like to estimate the discretization error introduced by those effects.

A straightforward way to obtain these estimates is to use a potential model to determine the wavefunction of a bound state, and then evaluate the expectation values of various operators. We might expect a potential model to provide a reliable estimate because the charm quarks are non-relativistic, and the short-distance potential is approximately Coulombic. This was done for the case of charmonium and bottomonium using two likely candidates for the potential model:

- Cornell Model: This is a simple potential well with a Coulomb + linear form [78, 79]:

$$V(r) = -\frac{\kappa}{r} + \frac{r}{a^2}. \tag{A.1}$$

It is the simplest model that includes both confinement and asymptotic freedom.

- Richardson Potential: This potential minimizes the number of parameters required to achieve both asymptotic freedom and linear confinement [80]. The form of the potential

is given in terms of the integral

$$V(r) = \frac{8\pi}{33 - 2n_f} \Lambda \left(\Lambda r - \frac{f(\Lambda r)}{\Lambda r} \right) \quad (\text{A.2})$$

where

$$f(t) = \left[1 - 4 \int_1^\infty \frac{dq}{q} \frac{e^{-qt}}{[\ln(q^2 - 1)]^2 + \pi^2} \right]. \quad (\text{A.3})$$

The wavefunction can be found by evaluating the Schrödinger equation in the presence of the potential well, and adjusting the energy level to find a wavefunction with the proper number of nodes (0 nodes for the ground state) that falls off as $r \rightarrow \infty$. Radial excitations merely require additional nodes in the wavefunctions. We can also model the orbital angular momentum with another simple change to the wavefunction.¹

The free parameters in the models (κ, a and M, Λ) must be tuned to achieve correct physics. This is most easily done by tuning for the correct meson mass for the ground (1S) and excited (2S) states. The parameters resulting from such tuning are given in Table A.1.

Once appropriate parameters are determined, we use the potential well (shown in Fig-

¹ In one dimension, the Schrödinger equation is

$$H\psi(x) = \left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right) \psi(x) = E\psi(x). \quad (\text{A.4})$$

We can write this more simply as

$$\psi_n'' = \frac{2m}{\hbar^2} [V(x) - E_n] \psi_n(x). \quad (\text{A.5})$$

In three dimensions with radial symmetry, the Schrödinger equation becomes

$$H\psi(\mathbf{r}) = \left(-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right) \psi(\mathbf{r}) = E\psi(\mathbf{r}). \quad (\text{A.6})$$

After separating out the angular dependence we get the radial Schrödinger equation:

$$\left(\frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} \right) R_{nl}(r) - \frac{2m}{\hbar^2} \left[V(r) + \frac{l(l+1)\hbar^2}{2mr^2} \right] R_{nl}(r) + \frac{2mE}{\hbar^2} R_{nl}(r) = 0. \quad (\text{A.7})$$

Substituting $u_{nl}(r) \equiv rR_{nl}(r)$ gives

$$u_{nl}''(r) = \frac{2m}{\hbar^2} \left[V(x) + \frac{l(l+1)\hbar^2}{2mr^2} - E_{nl} \right] u_{nl}(r). \quad (\text{A.8})$$

Note that this is the same as the one-dimensional case but with

$$V(x) \rightarrow V(x) + \frac{l(l+1)\hbar^2}{2mr^2} \quad (\text{A.9})$$

$$u_{nl}(0) = 0 \quad (\text{ie, } V(r < 0) = \infty). \quad (\text{A.10})$$

Table A.1: Table of potential model values.

Potential	Parameters
Cornell	$\kappa = 0.52$, $a = 2.34 \text{ GeV}^{-1}$, $m_c = 1.840 \text{ GeV}$, $m_b = 5.174 \text{ GeV}$
Richardson	$\Lambda = 0.3975$, $m_c = 1.4935 \text{ GeV}$, $m_b = 4.892 \text{ GeV}$

Table A.2: Potential model predictions for various charmonium states. The value indicates the result from the Richardson potential, while the error indicates the deviation if the Cornell model had been used instead.

Quantity \ State	1S	1P	2S	2P	3S	3P
m (GeV)	3.097	3.519(6)	3.688	3.955(13)	4.099(12)	4.312(27)
$\langle v^2/c^2 \rangle$	0.234(35)	0.248(48)	0.285(45)	0.314(51)	0.349(51)	0.378(58)
$\langle p^2 \rangle$ (GeV ²)	0.522(152)	0.553(125)	0.636(179)	0.701(188)	0.778(229)	0.843(242)
$\langle r^2 \rangle^{1/2}$ (fm)	0.423(47)	0.672(50)	0.856(96)	1.060(118)	1.214(143)	1.390(163)

ure A.1) to determine the appropriate wavefunctions. Plots of the resulting wavefunctions for charmonium and bottomonium are shown in Figure A.2. The wavefunctions encode all information about the mesons (to the extent that the potential model approximation is correct).

Using these wavefunctions one can obtain lots of information about the mesons, as shown in Tables A.2 and A.3. The agreement between the Richardson model predictions and experiment is impressive. For example, the Richardson model predicts the $\Upsilon(2S) - \Upsilon(1S)$ and $\Upsilon(3S) - \Upsilon(1S)$ splittings to high accuracy.

As an aside, we note that this technique may also be used to determine which of the several $\mathcal{O}(a^2)$ operators are most critical to obtain improvement. One might, for example, choose to leave out a particularly expensive (in terms of computational requirements) operator if it could be determined that its effects would be minor.

Table A.3: Potential model predictions for various bottomonium states. The value indicates the result from the Richardson potential, while the error indicates the deviation if the Cornell model had been used instead.

Quantity \ State	1S	1P	2S	2P	3S	3P
m (GeV)	9.460	9.903(55)	10.02(3)	10.26(5)	10.35(5)	10.53(6)
$\langle v^2/c^2 \rangle$	0.082(14)	0.068(3)	0.075(1)	0.077(2)	0.084(1)	0.087(1)
$\langle p^2 \rangle$ (GeV ²)	1.961(617)	1.638(95)	1.798(237)	1.839(181)	1.998(279)	2.094(262)
$\langle r^2 \rangle^{1/2}$ (fm)	0.224(23)	0.394(11)	0.505(28)	0.655(28)	0.750(44)	0.880(47)

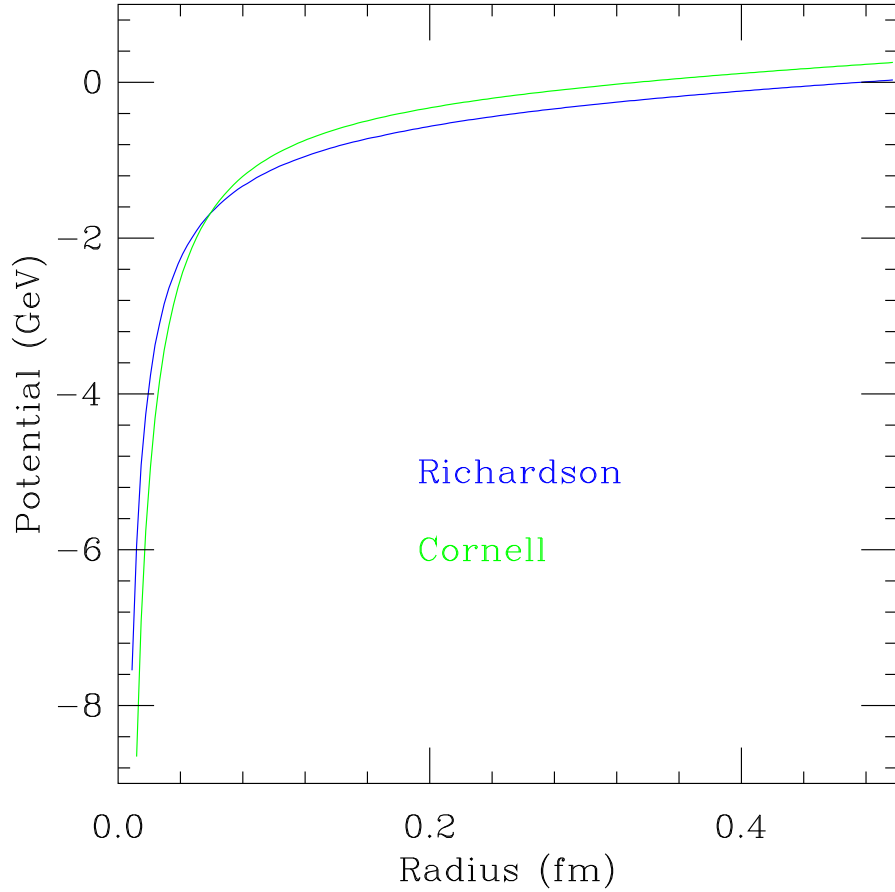


Figure A.1: Plots of the Cornell and Richardson potential wells. The Cornell model is shown in green, the Richardson model in blue.

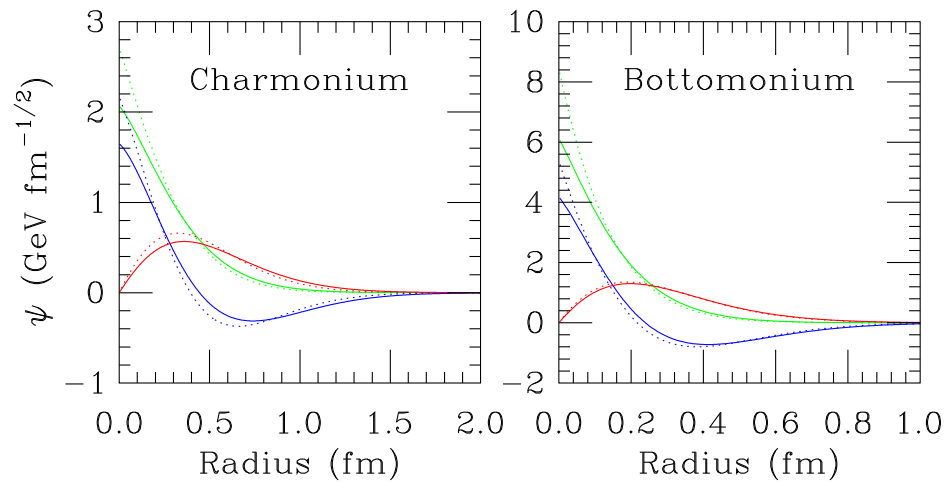


Figure A.2: Plots of the Schrödinger wavefunctions for charmonium and bottomonium. The solid lines show the $1S$, $2S$, and $1P$ states, calculated using the Richardson potential. The dotted lines show the same for the Cornell model.

Appendix B

Grassmann Variables

We represent fermions through the use of Grassmann variables η : numbers which anticommute such that $\{\eta_i, \eta_j\} = 0$. For each fermion, we will use a pair of these Grassmann variables: (η^+, η) . Because they anticommute, the square of a Grassmann variable is zero, and the most general function involving a single pair takes the form

$$F(\eta^+, \eta) = F^{(00)} + F^{(01)}\eta + F^{(10)}\eta^+ + F^{(11)}\eta^+\eta. \quad (\text{B.1})$$

The derivative is defined by

$$\partial_{\eta^+} F(\eta^+, \eta) \equiv F^{(10)} + F^{(11)}\eta \quad (\text{B.2})$$

$$\partial_{\eta} F(\eta^+, \eta) \equiv F^{(01)} - F^{(11)}\eta^+ \quad (\text{B.3})$$

which means

$$\partial_{\eta} \partial_{\eta^+} F(\eta^+, \eta) = F^{(11)} = -\partial_{\eta^+} \partial_{\eta} F(\eta^+, \eta). \quad (\text{B.4})$$

Furthermore, we define the integral by

$$\int d\eta d\eta^+ F(\eta^+, \eta) \equiv F^{(11)} \equiv - \int d\eta^+ d\eta F(\eta^+, \eta). \quad (\text{B.5})$$

Therefore we can integrate

$$\int d\eta^+ d\eta e^{-\lambda\eta^+\eta} = \int d\eta^+ d\eta (1 - \lambda\eta^+\eta) = \lambda. \quad (\text{B.6})$$

This can be extended to the case of a generic Gaussian integral:

$$\int d\eta_1^+ d\eta_1 \cdots d\eta_N^+ d\eta_N e^{-\sum_{i,j} \eta_j^+ A_{ji} \eta_i} = \det A. \quad (\text{B.7})$$

Another useful integral is

$$\int d\eta_1^+ d\eta_1 \cdots d\eta_N^+ d\eta_N e^{-\sum_{i,j} \eta_j^+ A_{ji} \eta_i} \eta_{j_1}^+ \eta_{i_1}^+ \cdots \eta_{j_n}^+ \eta_{i_n}^+ = \det A \sum_{k_1 \cdots k_n} \epsilon_{j_1 j_2 \cdots j_n}^{k_1 k_2 \cdots k_n} A_{k_1 i_1}^{-1} \cdots A_{k_n i_n}^{-1} \quad (\text{B.8})$$

where

$$\epsilon_{j_1 j_2 \cdots j_n}^{k_1 k_2 \cdots k_n} \equiv \begin{cases} 1 \\ -1 \\ 0 \end{cases} \text{ if } (j_1 j_2 \cdots j_n) \text{ is } \begin{cases} \text{even} \\ \text{odd} \\ \text{no} \end{cases} \text{ permutation of } (i_1 i_2 \cdots i_n). \quad (\text{B.9})$$

Appendix C

Numerical Techniques

C.1 Singular Matrices

The matrix inverse can be accomplished via LU decomposition, which requires $\mathcal{O}(N^3)$ operations. Using the LU decomposition, we can compute the determinant as¹

$$\det \mathbf{LU} = \det \mathbf{L} \det \mathbf{U} = \prod_i U_{ii}. \quad (\text{C.1})$$

Therefore computing a determinant is no harder than calculating a matrix inverse.

If the eigenvalues of a matrix differ by several orders of magnitude, it may impede the calculation of its inverse. One may define the norm of a matrix in terms of a vector norm as

$$\|\mathbf{A}\| \equiv \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A} \cdot \mathbf{x}\|}{\|\mathbf{x}\|}, \quad (\text{C.2})$$

where, in the Euclidean norm,

$$\|\mathbf{x}\|_2 = \left(\sum_i |x_i|^2 \right)^{1/2}. \quad (\text{C.3})$$

Using this definition, the norm measures the maximum stretching the matrix does to any vector. Defining the condition number of the matrix as

$$\text{cond}(\mathbf{A}) \equiv \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\| = \left(\max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A} \cdot \mathbf{x}\|}{\|\mathbf{x}\|} \right) \cdot \left(\min_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A} \cdot \mathbf{x}\|}{\|\mathbf{x}\|} \right)^{-1} \quad (\text{C.4})$$

we see the condition number gives the ratio of the maximum stretching to the maximum

¹Here we use the fact that the \mathbf{L} matrix has 1's along the main diagonal.

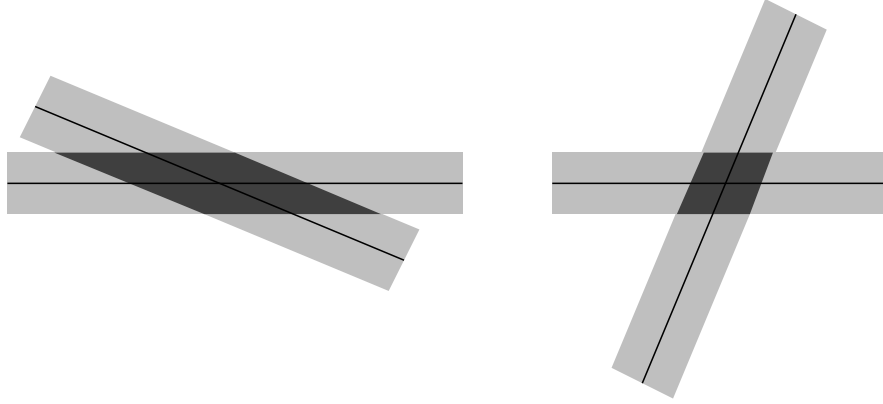


Figure C.1: A pictorial representation of the importance of the condition number of a linear system. The region of intersection of two lines with errors is best determined if the lines are orthogonal to each other.

shrinking.

The condition number may be used to understand the accuracy we may expect from solving a matrix equation. As a pictorial example, see Figure C.1.

C.2 Solving a System of Linear Equations

For some applications, we do not explicitly need the matrix inverse, but only to solve a system of linear equations. Rather than inverting the matrix, we can simply guess a solution and then use an iterative procedure to cause it to converge to the correct value. These iterative methods can potentially converge faster than the $\mathcal{O}(N^3)$ operations required by direct matrix inversion. In particular, iterative methods work well for sparse matrices.

The minimal residual algorithm to solve the matrix equation $\mathbf{M}\boldsymbol{\xi} = \boldsymbol{\eta}$ proceeds as follows:

- Take a guess of the correct answer, and denote it $\boldsymbol{\xi}_0$. Then calculate the residual (the difference from the correct answer) as $\mathbf{r}_0 = \boldsymbol{\eta} - \mathbf{M}\boldsymbol{\xi}_0$.
- Repeat the following until the desired level of convergence has been achieved, incrementing k at each step:

1. $\mathbf{s} = \mathbf{M}\mathbf{r}_k$

2. $\alpha = \frac{\mathbf{s}^\dagger \mathbf{r}_k}{\mathbf{s}^\dagger \mathbf{s}}$

3. $\boldsymbol{\psi}_{k+1} = \boldsymbol{\psi}_k + \omega\alpha\mathbf{r}_k$
4. $\mathbf{r}_{k+1} = \mathbf{r}_k - \omega\alpha\mathbf{s}$

Note that ω , the relaxation parameter, can be tuned to accelerate convergence.

Other, more complicated, algorithms exist that may provide faster convergence.² A common alternative to minimal residual is to use a conjugate gradient method. Here one tries to minimize the gradient of the residual, rather than the residual itself.

In addition, preconditioning the matrix may speed convergence in many cases. This means multiplying by another matrix in an attempt to reduce the condition number of the matrix.

C.3 Multidimensional Curvefitting

The simplex method is an intuitive algorithm for multidimensional curvefitting [81, §10.4]. In this method, we define a simplex³ in the parameter space. Each iteration begins by calculating the χ^2 at each vertex of the simplex and identifying the high point (worst χ^2) and best point (lowest χ^2). Initially, we reflect the high point through the opposite face of the simplex. If its value decreases so much that it becomes the new best point, we can try extrapolating it even further through the opposite face. If, on the other hand, it is still the high point, then we can try to make it better by finding an intermediate point along the perpendicular to the opposite face which is better. Finally, if no such intermediate point can be found, then we simply contract the entire simplex about the best point. Possible steps are shown in Figure C.2. This algorithm has the advantage that it is very simple to understand (think of an amoeba working its way through a complex space to a more desirable location). Also, it only requires function evaluations — the derivative is not required.

Other methods are more mathematical, and can provide faster convergence. Near a minimum $\mathbf{x} = \mathbf{m}$, we can Taylor expand the χ^2 function to find

$$\chi^2(\mathbf{m} + \mathbf{a}) = \chi^2(\mathbf{m}) + \partial_{\mathbf{x}}\chi^2(x)|_{\mathbf{m}+\mathbf{a}} \cdot \mathbf{a} + \frac{1}{2!}\mathbf{a} \cdot \partial_{\mathbf{x}}^2\chi^2(\mathbf{x})|_{\mathbf{m}+\mathbf{a}} \cdot \mathbf{a} + \dots \quad (\text{C.5})$$

²An introduction to several iterative methods for matrix inversion is given in Ref. [27, §11.5].

³A simplex is a geometrical shape which forms a minimal basis for a space. In a D -dimensional space, it must therefore have $D + 1$ vertices.

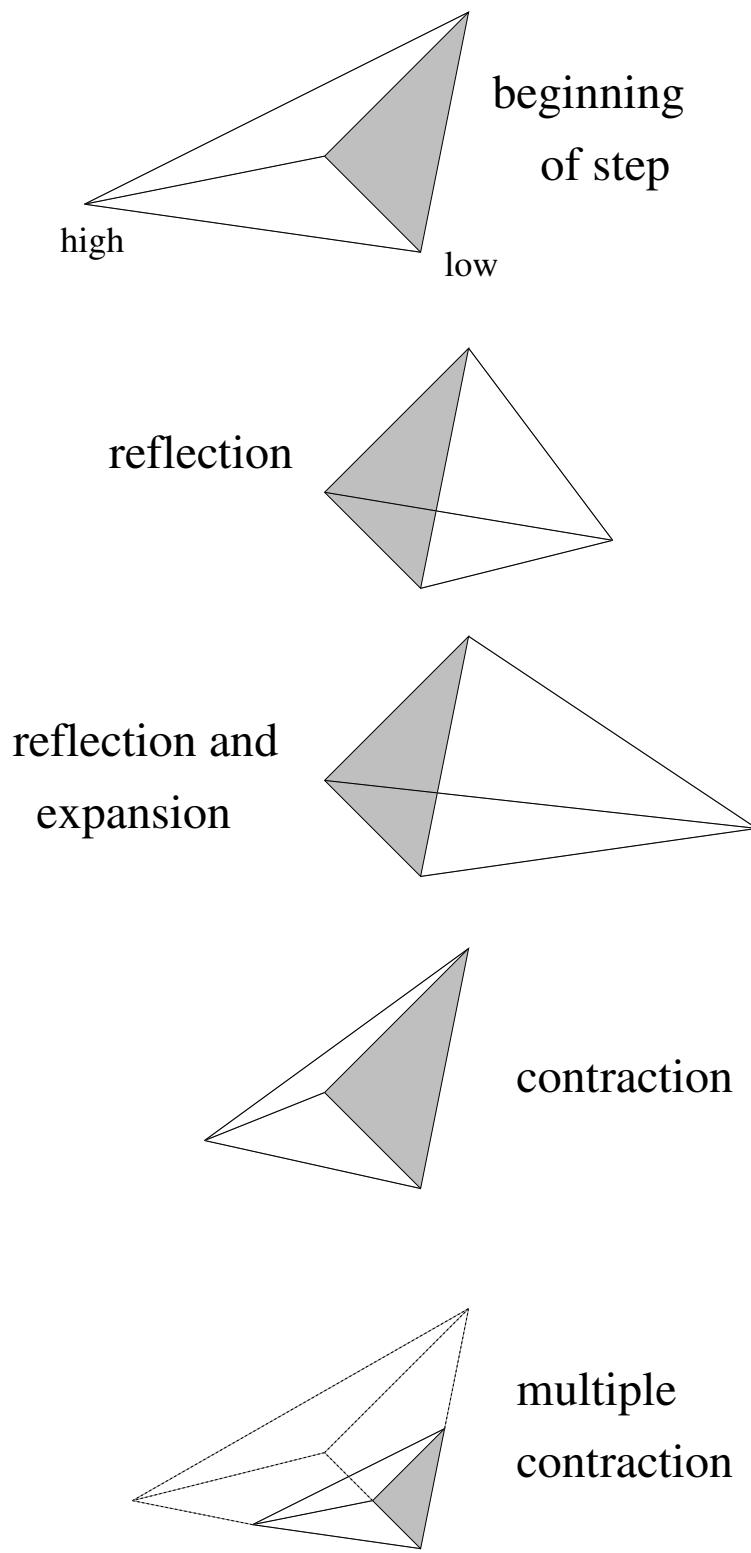


Figure C.2: Possible steps for the amoeba algorithm.

$$\approx \gamma - \mathbf{d} \cdot \mathbf{a} + \frac{1}{2} \mathbf{a} \cdot \mathbf{D} \cdot \mathbf{a} \quad (\text{C.6})$$

where $\gamma = \chi^2(\mathbf{m})$, $d_i = -\partial_{x_i} \chi^2(x)|_{\mathbf{m}+\mathbf{a}}$, and $D_{i,j} = \partial_{x_i, x_j} \chi^2(x)|_{\mathbf{m}+\mathbf{a}}$. Here the vector \mathbf{a} contains the deviation from the ideal fit parameters. Where this is a good approximation, we can calculate the best-fit parameters by differentiating with respect to \mathbf{a} and setting the result to 0. We find $\mathbf{d} = \mathbf{D} \cdot \mathbf{a}$, or $\mathbf{a} = \mathbf{D}^{-1} \cdot \mathbf{d}$. This is commonly referred to as the inverse-Hessian method. Where the approximation is *not* good, all we can do is follow a method of steepest descent, and step down the gradient.

Multi-dimensional fitting can be accomplished efficiently using the Levenberg-Marquardt method [81, §15.5]. This method combines the linear convergence of steepest descent (far from the minimum) with the quadratic convergence of the inverse-Hessian method (near the minimum). It does this by combining the two methods while introducing a scale parameter, λ , that controls their relative weight, with large λ favoring the steepest descent method, and small λ favoring the inverse-Hessian method. In practice, we simply take $\mathbf{D} \rightarrow \mathbf{D} + \lambda \mathbf{1}$. By adjusting λ during the fitting process, the method can smoothly cross over between the two extremes. When the fitter takes a good step, it decreases λ , which increases the relative weight of the inverse-Hessian method. If it takes a bad step, it increases λ , moving it back to a steepest descent method. The method therefore automatically finds a balance, shifting to get quadratic convergence as the minimum is approached.

As with any multi-dimensional fitting routine, we must restart the algorithm after reaching the minimum. We do this by setting $\lambda = 1$ and repeating the fit, starting from its previous fit result. The algorithm will therefore be encouraged to take a large first step, which will take it away from the minimum it had found. This is necessary to avoid the case where we find a local minimum that differs from the global minimum. While there is no known method to guarantee one finds the global minimum, restarting the fitting algorithm one or more times can give a greater confidence in the result. If the minimum it had previously found was the true global minimum, the algorithm will find it again quickly (since it's already starting nearby), so little additional computer time is required. The benefit, however, is that if it was *not* the true global minimum it has a chance to move away from the local minimum and find a better

one.

C.4 Singular Value Decomposition

It is clearly not possible to find the inverse of a singular matrix. What is less obvious is what happens in the case of a *nearly*-singular matrix. In a computer, the finite precision means that we could get a poor inverse due to roundoff errors if a matrix is nearly singular. A technique to obtain the most accurate possible inverse is to invert the matrix for those eigenvectors which dominate the solution. This technique is called Singular Value Decomposition (SVD).

One may obtain an eigenvalue-like decomposition for any $m \times n$ matrix \mathbf{A} as

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{C.7}$$

where \mathbf{U} is an $m \times m$ orthogonal matrix, \mathbf{V} is an $n \times n$ orthogonal matrix, and $\mathbf{\Sigma}$ is an $m \times n$ diagonal matrix with the singular values on the diagonal. For a rank-deficient matrix, one or more of the singular values may be zero (to within machine precision). We can now compute the pseudoinverse as

$$\mathbf{A}^+ = \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T \tag{C.8}$$

where the pseudoinverse of a scalar is defined as

$$\sigma^+ = \begin{cases} 1/\sigma, & \sigma \neq 0 \\ 0, & \sigma = 0 \end{cases} . \tag{C.9}$$

Note that the ratio of the maximum singular value to the minimum singular value gives an estimate of the condition number of the matrix. We can therefore reduce the condition number (to allow for a more accurate matrix inverse, for example) by discarding those singular values which do not contribute to within some specified precision. For example, after identifying the maximum singular value, σ_{\max} , we could discard any singular values smaller than $\sigma_{\max}/10^{15}$.

Note that, by definition, it should⁴ be possible to decompose any matrix into its singular

⁴In practical implementations, this is not always the case. During the course of this research it was discovered that the SVD routine called by Numerical Python would fail for some matrices. It turned out that their

values. This is therefore a robust technique whenever a matrix inverse is required, and is highly recommended in cases where the matrix might have a large condition number.

implementation (in `dlapack_lite.c`) used an iterative method to converge on the decomposition. In rare cases, that iterative routine would hit its maximum iteration limit and fail. The fix, which I suggested in bug 601052, was to increase the maximum iteration count [82].

Appendix D

Model Function Derivatives

The calculation of the model-function derivative is a somewhat tedious task. For speed, however, it is helpful to hand-code the derivatives.

Consider the model function described in Section 5.3.1:

$$G_{\text{src,snk}}(t) = 2 \sum_n Z_{\text{src}}^{(n)} Z_{\text{snk}}^{(n)} \exp(-E_n t) \cosh(-E_n t) \quad (\text{D.1})$$

where

$$\begin{aligned} E_1 &= E_0 + e^{\epsilon_1} \\ E_2 &= E_1 + e^{\epsilon_2} \\ E_3 &= E_2 + e^{\epsilon_3} \\ &\vdots \end{aligned} \quad (\text{D.2})$$

The Marquardt-Levenberg algorithm, in order to determine which way is downhill, needs the partial derivative with respect to each parameter. It is simplest to consider the cases of the E_0 and ϵ_i , and the $Z_{\text{smearing}}^{(n)}$ separately.

We can find the energy parameter derivatives most easily by considering the generic problem (where $y = y(x)$):

$$\frac{\partial}{\partial x} e^y \cosh y = \frac{\partial y}{\partial x} e^y (\cosh y + \sinh y). \quad (\text{D.3})$$

In the case of interest, we have $y = -(E_0 + \epsilon_1 + \epsilon_2 + \dots + \epsilon_n)t$. Therefore, $\partial_{E_0} y = -t$ for all

n . We find the partial derivative of our model function is just

$$\frac{\partial G_{\text{src,snk}}(t)}{\partial E_0} = -2t \sum_n Z_{\text{src}}^{(n)} Z_{\text{snk}}^{(n)} \exp(-E_n t) (\cosh(-E_n t) + \sinh(-E_n t)). \quad (\text{D.4})$$

This is not hard to generalize to the derivatives with respect to the ϵ_i . For this case, $\partial_{\epsilon_i} y = -t$ for $n \geq i$, and zero otherwise. In equation form,

$$\frac{\partial G_{\text{src,snk}}(t)}{\partial \epsilon_i} = -2t \sum_{n \geq i} Z_{\text{src}}^{(n)} Z_{\text{snk}}^{(n)} \exp(-E_n t) (\cosh(-E_n t) + \sinh(-E_n t)). \quad (\text{D.5})$$

The derivatives with respect to the coefficients are simpler, but there is one minor catch: the two smearings might be identical, which would lead to a Z^2 term. We therefore find

$$\frac{\partial G_{\text{src,snk}}(t)}{\partial Z_{\text{src}}^{(n)}} = \begin{cases} 2Z_{\text{snk}}^{(n)} \exp(-E_n t) \cosh(-E_n t) & \text{for src} \neq \text{snk} \\ 4Z_{\text{snk}}^{(n)} \exp(-E_n t) \cosh(-E_n t) & \text{for src} = \text{snk} \end{cases} \quad (\text{D.6})$$

Because we don't decide how many energy levels to include in a fit until runtime, any program we write must be completely general in that respect. The equations of the above two sections are realized through the Python code shown in Figure D.1. At the end of this code, the array `dfdp` will contain the partial derivatives with respect to each parameter.

It should be noted that this is one of the simplest model functions we consider. This entire procedure must be repeated for each model function we wish to optimize. To reduce the likelihood of introducing bugs, we rely heavily on the automatic differentiation of Section 5.1.4.

```

dfdp = [] # Start with an empty dfdp array

sum = 0 # Take care of the energy parameters
for elevel in range(self.n_lvl): # E0 affects all energy levels
    sum += (src_spect[elevel]
            *snk_spect[elevel]
            *(-self.Tmid*local_cosh(espect[elevel]*Tmid_t)
              +Tmid_t*local_sinh(espect[elevel]*Tmid_t)
              )
            *2*local_exp(-spect[elevel]*self.Tmid)
            )
dfdp.append(sum)

for i in range(1,self.n_lvl): # eps_i affects all energy levels >= i
    sum = 0
    for elevel in range(i,self.n_lvl):
        sum += (src_spect[elevel]
                *snk_spect[elevel]
                *2*self.par[i]
                *(-self.Tmid*local_cosh(espect[elevel]*Tmid_t)
                  +Tmid_t*local_sinh(espect[elevel]*Tmid_t)
                  )
                *2*local_exp(-spect[elevel]*self.Tmid)
                )
    dfdp.append(sum)

for elevel in range(self.n_lvl): # Take care of the source/sink parameters
    for i in range(self.n_src_snk):
        if (i==source or i==sink):
            if (source==sink):
                dfdp.append(2*src_spect[elevel]
                            *local_exp(-spect[elevel]*self.Tmid)
                            *2*local_cosh(espect[elevel]*Tmid_t)
                            )
            else:
                dfdp.append(src_spect[elevel]
                            *snk_spect[elevel]
                            *local_exp(-spect[elevel]*self.Tmid)
                            *2*local_cosh(espect[elevel]*Tmid_t)
                            /self.src_snk(elevel,i)
                            )
        else:
            dfdp.append(zeros(len(t), 'd'))

```

Figure D.1: Sample Python code that implements the derivative of Eq. (D.1): Eq. (D.6).

Appendix E

Partially Quenched Staggered Chiral Perturbation Theory

Partially quenched staggered chiral perturbation theory (PQS χ PT) treats the independent extrapolation of light sea and valence quarks to their continuum values. The chiral expansion has the form

$$f_{Q_q} \sqrt{M_{Q_q}} = \alpha \left(1 + \frac{1}{16\pi^2 f^2} \Delta f_q + \dots \right) \quad (\text{E.1})$$

where the Δf_q contains the chiral logs and the \dots are terms analytic in m_q and $\sum_f m_f$. Following Aubin and Bernard [83, 84, 72], we write the 16 tastes as the multiplets¹ P,A,T,V,I with degeneracies 1,4,6,4,1, we can write the contributions to the chiral log as

$$\Delta f_q = -\frac{1 + 3g_\pi^2}{2} (h^{av} + h^I + h^A + h^V). \quad (\text{E.2})$$

The first contribution is simply an average over the 16 meson tastes:

$$h^{av} = \frac{1}{16} \sum_{\xi=1}^{16} \sum_{f=1}^N M_{qf,\xi}^2 \log(M_{qf,\xi}^2). \quad (\text{E.3})$$

The other contributions are, for the case $N_f = 2 + 1$,

$$h^I = \frac{1}{N} \frac{d}{dM_{qq,I}^2} \left(\sum_{b=1}^{N-1} R_b^{[N-1, N-1]} (\{M^{qI}\}; \{\mu^I\}) M_{b,I}^2 \log M_{b,I}^2 \right) \quad (\text{E.4})$$

¹We use the notation P,A,T,V,I to refer to the pseudoscalar, axial-vector, tensor, vector, and isosinglet states, respectively.

$$h^A = a^2 \delta'_A \frac{d}{dM_{qq,A}^2} \left(\sum_{b=1}^N R_b^{[N,N-1]} (\{M^{qA}\}; \{\mu^A\}) M_{b,A}^2 \log M_{b,A}^2 \right) \quad (\text{E.5})$$

and similarly for h^V with the substitution $A \rightarrow V$. Here

$$\{M^{qI}\} = \{M_{qq,I}, M_{\eta,I}\} \text{ with } M_{\eta,I}^2 = \frac{M_{U,I}^2}{3} + \frac{2M_{S,I}^2}{3} \quad (\text{E.6})$$

$$\{\mu^I\} = \{M_{U,I}, M_{S,I}\} \quad (\text{E.7})$$

$$\{M^{qA}\} = \{M_{qq,A}, M_{\eta,A}, M_{\eta',A}\} \quad (\text{E.8})$$

$$\{\mu^A\} = \{M_{U,A}, M_{S,A}\} \quad (\text{E.9})$$

and

$$R_j^{[n,k]}(\{m\}; \{\mu\}) \equiv \frac{\prod_{a=1}^k (\mu_a^2 - m_j^2)}{\prod_{i \neq j} (m_i^2 - m_j^2)}. \quad (\text{E.10})$$

Finally, we learn from fits to f_π that [85, Eq. 47]

$$r_1^2 a^2 \delta'_A = -0.28(3)(5) \quad (\text{E.11})$$

$$r_1^2 a^2 \delta'_V = -0.11(8)^{(+21)}_{(-4)}. \quad (\text{E.12})$$

These values will be taken as priors in our f_D fits. We also set g_π to its experimentally measured value [86].

The h^A and h^V are suppressed by a factor of a^2 . These terms embody the $\mathcal{O}(a^2)$ discretization errors resulting from taste-changing interactions. While we must include these terms in fits to lattice data, we will remove them when taking the continuum limit $a \rightarrow 0$.

Taking the derivatives for h^I , h^V , and h^A is straightforward, if tedious. One finds:

$$\begin{aligned} h^I &= \frac{1}{3} \left[\frac{((f-q)(s-q) + ((f-q)(s-q) - q(s-q) - q(f-q)) \log(q/\mu))}{\eta - q} \right. \\ &\quad \left. + \frac{(f-q)(s-q)q \log(q/\mu) - (f-\eta)(s-\eta)\eta \log(\eta/\mu)}{(\eta-q)^2} \right] \quad (\text{E.13}) \\ h^A &= a^2 \delta'_A \left[\frac{(f-q)(s-q) + ((f-q)(s-q) - q(s-q) - q(f-q)) \log(q/\mu)}{(\eta-q)(\eta'-q)} \right. \\ &\quad \left. + \frac{((\eta-q) + (\eta'-q))(f-q)(s-q)q \log(q/\mu)}{(\eta-q)^2(\eta'-q)^2} \right] \end{aligned}$$

Table E.1: Taste splittings, for a meson of taste Γ . Values obtained from light mesons [87].

Γ	degeneracy	$\delta(\Gamma)$
P (Goldstone)	1	0.0000000
A	4	0.0295374
T	6	0.0471298
V	4	0.0631816
I (singlet)	1	0.0772650

$$- \left[\frac{(f - \eta)(s - \eta)\eta \log(\eta/\mu)}{(q - \eta)^2(\eta' - \eta)} - \frac{(f - \eta')(s - \eta')\eta' \log(\eta'/\mu)}{(q - \eta')^2(\eta - \eta')} \right] \quad (\text{E.14})$$

Here I have used the abbreviated notation $M_{q,I}^2 \rightarrow q$, etc., so \sqrt{q} has units of mass. In the above, f refers to the sea quark, q refers to the valence quark, and μ is a renormalization scale².

Finally, the masses must be adjusted for taste splittings. We use the formula

$$a^2 m_\pi^2(\Gamma) = \delta(\Gamma) + (a\mu)[am_X + am_Y] \quad (\text{E.15})$$

where $a\mu = 2.3647$ on the coarse lattice, and the $\delta(\Gamma)$ are defined in Table E.1.

²The renormalization scale comes from the mass of the singlet contribution to the η' .

Appendix F

Asymmetric Actions

In this work, we utilize a symmetric lattice, where the lattice spacing a is the same in the temporal and three spatial directions. While this is a natural way to approach a relativistic system, some improvement is possible for heavy quarks. The heavy quarks move non-relativistically, so they have an asymmetry in their 4-momenta. It is possible to correct for this asymmetry (and essentially boost the quarks to a relativistic frame). This can be done by treating the spatial and temporal components of the operators differently.

This approach, called the Fermilab approach [5], is best applied to the charm and bottom quarks. It is fairly straightforward to determine the necessary operators for improvement to $\mathcal{O}(a)$. The only dimension 3 operator is still $\bar{\psi}\psi$, but the dimension 4 operator now splits into two operators:

$$\bar{\psi}\not{D}\psi \rightarrow \begin{cases} \bar{\psi}\gamma_0 D_0\psi \\ \bar{\psi}\boldsymbol{\gamma}\cdot\mathbf{D}\psi \end{cases} . \quad (\text{F.1})$$

At dimension 5, the splitting is even more complex:

$$\bar{\psi}\not{D}^2\psi \rightarrow \begin{cases} \bar{\psi}D_0^2\psi \\ \bar{\psi}(\boldsymbol{\gamma}\cdot\mathbf{D})^2\psi \end{cases} \quad (\text{F.2})$$

$$i\bar{\psi}\sigma_{\mu\nu}F_{\mu\nu}\psi \rightarrow \begin{cases} i\bar{\psi}\boldsymbol{\Sigma}\cdot\mathbf{B}\psi \\ \bar{\psi}\boldsymbol{\alpha}\cdot\mathbf{E}\psi \\ \bar{\psi}[\gamma_0 D_0, \boldsymbol{\gamma}\cdot\mathbf{D}]\psi \end{cases} . \quad (\text{F.3})$$

Redundant operators must be removed and the coefficients determined — a task that is

complicated by the fact that there are many more operators to take into account. Without axis-interchange symmetry, the transformation in Eq. (3.72) must be generalized to

$$\begin{aligned}\psi &\rightarrow \exp[\epsilon a(\not{D} + m) + \delta a \boldsymbol{\gamma} \cdot \mathbf{D}] \psi \\ \bar{\psi} &\rightarrow \bar{\psi} \exp[\bar{\epsilon} a(\not{D} + m) + \bar{\delta} a \boldsymbol{\gamma} \cdot \mathbf{D}].\end{aligned}\tag{F.4}$$

Applying this transformation leads to the dimension-five interactions listed in Eqs. (F.2) and (F.3). From the four independent parameters only three combinations survive: $\epsilon + \bar{\epsilon}$, $\delta + \bar{\delta}$, and $\delta - \bar{\delta}$. The coefficients of the three terms $\bar{\psi} D_0^2 \psi$, $\bar{\psi} (\boldsymbol{\gamma} \cdot \mathbf{D})^2 \psi$, and $\bar{\psi} [\gamma_0 D_0, \boldsymbol{\gamma} \cdot \mathbf{D}] \psi$ can be chosen arbitrarily. The first two terms can be used to solve the doubling problem; the third coefficient is chosen to be zero for simplicity (and performance).

The asymmetry modifies the fermion propagator slightly. As a result, the normalization of the fermion field becomes

$$e^{M_1 a/2} \psi(x) = \sqrt{1 - 6r_s \kappa_s} \psi_n,\tag{F.5}$$

where r_s and κ_s refer to the Wilson parameter and hopping parameter in a spatial dimension, and $M_1 \equiv E(0)$ is the rest mass. Results from calculations utilizing an $\mathcal{O}(a)$ asymmetric lattice action are promising [88].

It turns out that, to $\mathcal{O}(a)$, the physics depends only on the kinetic mass (the M_2 of Eq. (3.44)) regardless of the tuning of the asymmetry parameter [88, Figure 5]. It is therefore not necessary to tune the asymmetry parameter ζ . In this work, we have simply set $\zeta = 1$.

The next step is to construct the $\mathcal{O}(a^2)$ improved asymmetric action. At dimension 6, there are two classes of operators: bilinear operators $\bar{\psi}(x) O^b \psi(x)$ where O^b is of dimension 3, and quadratic operators $(\bar{\psi}(x) O^q \psi(x))^2$ where O^q is of dimension 0. The quadratic operators describe four-fermion interactions. A straightforward, but tedious, calculation shows the bilinear terms can be written as

$$\begin{aligned}O_4(x) &= \bar{\psi}(x) \gamma_\mu D_\mu^3 \psi(x) \\ O_5(x) &= \bar{\psi}(x) D^2 \not{D} \psi(x)\end{aligned}$$

$$O_6(x) = \bar{\psi}(x)\not{D}D^2\psi(x) \quad (\text{F.6})$$

$$O_7(x) = \bar{\psi}(x)\gamma_\mu[D_\nu, F_{\mu\nu}]\psi(x)$$

$$O_8(x) = \bar{\psi}(x)\not{D}^3\psi(x).$$

We have already seen O_4 (the Naik term) in Section 3.4.4. It is discretized as

$$\begin{aligned} O_4^L &\equiv \bar{\psi}(x)\gamma_\mu(D_\mu^3)^L\psi(x) \\ &= \bar{\psi}(x)\sum_\mu\frac{\gamma_\mu}{(2a)^3}\left[U_\mu(x+2\hat{\mu})U_\mu(x+\hat{\mu})U_\mu(x)\psi(x+3\hat{\mu})-3U_\mu(x)\psi(x+\hat{\mu})\right. \\ &\quad \left.+3U_\mu^\dagger(x-\hat{\mu})\psi(x-\hat{\mu})-U_\mu^\dagger(x-3\hat{\mu})U_\mu^\dagger(x-2\hat{\mu})U_\mu^\dagger(x-\hat{\mu})\psi(x-3\hat{\mu})\right]. \end{aligned} \quad (\text{F.7})$$

This operator reaches sites 3 hops away.

For the quartic terms, everything is squared so charge, parity, and discrete lorentz symmetries don't restrict our operators. In addition, we must allow for a gauge symmetry by writing the operators as $(\bar{\psi}\Gamma t^a\psi)^2$. Finally, we must conserve flavor symmetry (allowing for a four-quark interaction of two different flavors) so we must include the flavor generators β^A . These considerations add to our list the operators:

$$\begin{aligned} O_9(x) &= (\bar{\psi}(x)t^a\psi(x))^2 & O_{10}(x) &= (\bar{\psi}(x)t^a\beta^A\psi(x))^2 \\ O_{11}(x) &= (\bar{\psi}(x)\gamma_5 t^a\psi(x))^2 & O_{12}(x) &= (\bar{\psi}(x)\gamma_5 t^a\beta^A\psi(x))^2 \\ O_{13}(x) &= (\bar{\psi}(x)\gamma_\mu t^a\psi(x))^2 & O_{14}(x) &= (\bar{\psi}(x)\gamma_\mu t^a\beta^A\psi(x))^2 \\ O_{15}(x) &= (\bar{\psi}(x)\gamma_5\gamma_\mu t^a\psi(x))^2 & O_{16}(x) &= (\bar{\psi}(x)\gamma_5\gamma_\mu t^a\beta^A\psi(x))^2 \\ O_{17}(x) &= (\bar{\psi}(x)\sigma_{\mu\nu} t^a\psi(x))^2 & O_{18}(x) &= (\bar{\psi}(x)\sigma_{\mu\nu} t^a\beta^A\psi(x))^2. \end{aligned} \quad (\text{F.8})$$

It is expected that the 5+10 symmetric dimension 6 operators will expand out to 16+18 operators in the asymmetric case. To determine which operators are redundant, one must begin with the field redefinitions

$$\begin{aligned} A_\mu &\rightarrow A_\mu + a^2\left(\frac{1}{2}\epsilon\sum_\nu[D_\nu, F_{\mu\nu}] + \epsilon'\bar{\psi}\gamma_\mu\psi\right) \\ \psi &\rightarrow \psi + a\epsilon_1\not{D}\psi + a^2\epsilon_2 D^2\psi + a^2\epsilon_3\not{D}^2\psi \\ \bar{\psi} &\rightarrow \bar{\psi} + a\epsilon_1\bar{\psi}\not{D} + a^2\epsilon_2\bar{\psi}D^2 + a^2\epsilon_3\bar{\psi}\not{D}^2. \end{aligned} \quad (\text{F.9})$$

Choosing the form of these operators and eliminating the redundant ones is a current area of research [89, 90]. It will then be necessary to determine the coefficients of these operators. As each coefficient is calculated, we can immediately implement the corresponding term into our lattice actions. Although we won't be guaranteed complete $\mathcal{O}(a^2)$ improvement until all coefficients have been calculated, we should be able to reap some benefits of $\mathcal{O}(a^2)$ improvement as additional terms are added.

References

- [1] Kenneth G. Wilson. Confinement of quarks. *Phys. Rev.*, D10(8):2445–2459, Oct 1974.
- [2] Michael Dine. The strong CP problem. In *Flavor Physics for the Millennium*, pages 349–369, 2000.
- [3] B. A. Thacker and G. Peter Lepage. Heavy quark bound states in lattice QCD. *Phys. Rev.*, D43:196–208, 1991.
- [4] G. Peter Lepage, Lorenzo Magnea, Charles Nakhleh, Ulrika Magnea, and Kent Hornbostel. Improved nonrelativistic QCD for heavy quark physics. *Phys. Rev.*, D46:4052–4067, 1992.
- [5] Aida X. El-Khadra, Andreas S. Kronfeld, and Paul B. Mackenzie. Massive fermions in lattice gauge theory. *Phys. Rev. D*, 55(7):3933–3957, 1997.
- [6] T. A. Armstrong et al. Observation of the 1P_1 state of charmonium. *Phys. Rev. Lett.*, 69:2337–2340, 1992.
- [7] C. Baglin et al. Search for the 1P_1 charmonium state in $\bar{p}p$ annihilations at the CERN intersecting storage rings. *Phys. Lett.*, B171:135, 1986.
- [8] S. Eidelman et al. Review of particle physics. *Phys. Lett.*, B592:1, 2004.
- [9] B. Aubert et al. Observation of a narrow meson decaying to $D_s^+\pi^0$ at a mass of 2.32 GeV/ c^2 . *Phys. Rev. Lett.*, 90:242001, 2003.
- [10] D. Besson et al. Observation of a narrow resonance of mass 2.46 GeV/ c^2 in the $D_s^{*+}\pi^0$ final state, and confirmation of the $D_{sJ}^*(2317)$. *AIP Conf. Proc.*, 698:497–502, 2004.

- [11] N. Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531–532, 1963.
- [12] M. Kobayashi and T. Maskawa. CP violation in the renormalizable theory of weak interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.
- [13] Francis Halzen and Alan D. Martin. *Quarks & Leptons: An Introductory Course in Modern Particle Physics*. Wiley, New York, 1984.
- [14] Lincoln Wolfenstein. Parametrization of the Kobayashi-Maskawa matrix. *Phys. Rev. Lett.*, 51:1945, 1983.
- [15] M. Ciuchini et al. Unitarity triangle analysis in the standard model and sensitivity to new physics. *ECONF*, C0304052:WG306, 2003.
- [16] G. Bonvicini et al. Improved measurement of $\mathcal{B}(D^+ \rightarrow \mu^+\nu)$ and the pseudoscalar decay constant f_{D^+} . <http://www.lepp.cornell.edu/~dgc/cleo-fd.pdf>, 2005.
- [17] G. Bonvicini et al. Measuring $B(D^+ \rightarrow \mu^+\nu)$ and the pseudoscalar decay constant $f(D^+)$. *Phys. Rev.*, D70:112004, 2004.
- [18] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Addison-Wesley, New York, 1995.
- [19] Kenneth G. Wilson. Quarks and strings on a lattice. In A. Zichichi, editor, *New Phenomena in Subnuclear Physics*. Plenum Press, 1977. International School of Subnuclear Physics, Erice, Sicily, July 11 – August 1, 1975.
- [20] H. B. Nielsen and M. Ninomiya. Absence of neutrinos on a lattice. 1. Proof by homotopy theory. *Nucl. Phys.*, B185:20, 1981.
- [21] H. B. Nielsen and M. Ninomiya. Absence of neutrinos on a lattice. 2. Intuitive topological proof. *Nucl. Phys.*, B193:173, 1981.
- [22] István Montvay and Gernot Münster. *Quantum Fields on a Lattice*. Cambridge University Press, Cambridge, 1994.

- [23] Ivan Horvath. Ginsparg-Wilson-Luescher symmetry and ultralocality. *Phys. Rev.*, D60:034510, 1999.
- [24] James Ward Brown and Ruel V. Churchill. *Complex Variables and Applications*. McGraw-Hill, sixth edition, 1996.
- [25] Leonard Susskind. Lattice fermions. *Phys. Rev.*, D16:3031–3039, 1977.
- [26] Karl Jansen. Actions for dynamical fermion simulations: Are we ready to go? <http://www-zeuthen.desy.de/~kjansen/tsukuba.pdf>, July 2003.
- [27] Michael T. Heath. *Scientific Computing: An Introductory Survey*. WCB/McGraw-Hill, St. Louis, 1997.
- [28] Steven A. Gottlieb, W. Liu, D. Toussaint, R. L. Renken, and R. L. Sugar. Hybrid molecular dynamics algorithms for the numerical simulation of quantum chromodynamics. *Phys. Rev.*, D35:2531–2542, 1987.
- [29] S. Aoki et al. An exact algorithm for any-flavor lattice QCD with Kogut-Susskind fermion. *Comput. Phys. Commun.*, 155:183–208, 2003.
- [30] Karl Jansen. Actions for dynamical fermion simulations: Are we ready to go? *Nucl. Phys. Proc. Suppl.*, 129:3–16, 2004.
- [31] Claude W. Bernard and Maarten F. L. Golterman. Partially quenched gauge theories and an application to staggered fermions. *Phys. Rev.*, D49:486–494, 1994.
- [32] C. T. H. Davies et al. High-precision lattice QCD confronts experiment. *Phys. Rev. Lett.*, 92:022001, 2004.
- [33] K. Symanzik. Continuum limit and improved action in lattice theories. 1. principles and ϕ^4 theory. *Nucl. Phys.*, B226:187–204, 1983.
- [34] M. Lüscher and P. Weisz. On-shell improved lattice gauge theories. *Commun. Math. Phys.*, 97:59–77, 1985.

- [35] M. Luscher and P. Weisz. Computation of the action for on-shell improved lattice gauge theories at weak coupling. *Phys. Lett.*, B158:250, 1985.
- [36] B. Sheikholeslami and R. Wohlert. Improved continuum limit lattice action for QCD with Wilson fermions. *Nucl. Phys.*, B259:572–596, 1985.
- [37] G. Peter Lepage and Paul B. Mackenzie. Renormalized lattice perturbation theory. *Nucl. Phys. Proc. Suppl.*, 20:173–176, 1991. Lattice 90.
- [38] G. Peter Lepage and Paul B. Mackenzie. On the viability of lattice perturbation theory. *Phys. Rev.*, D48:2250–2264, 1993.
- [39] Satchidananda Naik. On-shell improved lattice action for QCD with Susskind fermions and asymptotic freedom scale. *Nucl. Phys.*, B316:238, 1989.
- [40] J.-F. Lagaë and D. K. Sinclair. Improving the staggered quark action to reduce flavour symmetry violations. *Nucl. Phys. Proc. Suppl.*, 63:892–894, 1998.
- [41] J.-F. Lagaë and D. K. Sinclair. Improved staggered quark actions with reduced flavour symmetry violations for lattice QCD. *Phys. Rev.*, D59:014511, 1999.
- [42] G. Peter Lepage. Flavor-symmetry restoration and symanzik improvement for staggered quarks. *Phys. Rev.*, D59:074502, 1999.
- [43] Tom Blum et al. Improving flavor symmetry in the Kogut-Susskind hadron spectrum. *Phys. Rev.*, D55:1133–1137, 1997.
- [44] Kostas Orginos, Doug Toussaint, and R. L. Sugar. Variants of fattening and flavor symmetry restoration. *Phys. Rev.*, D60:054503, 1999.
- [45] Mark G. Alford, W. Dimm, G. P. Lepage, G. Hockney, and P. B. Mackenzie. Lattice QCD on small computers. *Phys. Lett.*, B361:87–94, 1995.
- [46] Steven A. Gottlieb, W. Liu, D. Toussaint, R. L. Renken, and R. L. Sugar. Empirical study of the hybrid molecular dynamics approach to the simulation of QCD. *Phys. Rev.*, D36:3797, 1987.

- [47] Gordon E. Moore. Cramming more components onto integrated circuits. *Electronics*, 38(8), April 1965.
- [48] Aida X. El-Khadra. Lattice QCD. In S. Keller and H. D. Wahl, editors, *Physics in Collision*, pages 209–232, France, 1995. Editions Frontieres.
- [49] G. Peter Lepage and John H. Sloan. Precision Υ spectroscopy and fundamental parameters from NRQCD. *Nucl. Phys. Proc. Suppl.*, 34:417–424, 1994.
- [50] Aida X. El-Khadra, Andreas S. Kronfeld, Paul B. Mackenzie, Sinead M. Ryan, and James N. Simone. B and D meson decay constants in lattice QCD. *Phys. Rev.*, D58:014506, 1998.
- [51] Matthew Wingate, Junko Shigemitsu, Christine T. H. Davies, G. Peter Lepage, and Howard D. Trottier. Heavy-light mesons with staggered light quarks. *Phys. Rev.*, D67:054505, 2003.
- [52] Steven A. Gottlieb. Benchmarking and tuning the MILC code on clusters and supercomputers. *Nucl. Phys. Proc. Suppl.*, 106:1031–1033, 2002.
- [53] Steven A. Gottlieb. Cost of dynamical quark simulations with improved staggered quarks. *Nucl. Phys. Proc. Suppl.*, 106:189–190, 2002.
- [54] M. Fischler. *The ACPMAPS System: A Detailed Overview*. FERMILAB-TM-1780.
- [55] Neal Margulis. *i860 Microprocessor Architecture*. McGraw-Hill, 1990.
- [56] D. Chen et al. QCDOC: A 10-teraflops scale computer for lattice QCD. *Nucl. Phys. Proc. Suppl.*, 94:825–832, 2001.
- [57] P. Boyle et al. The QCDOC project. *Nucl. Phys. Proc. Suppl.*, 140:169–175, 2005.
- [58] Massimo Di Pierro et al. Charmonium with three flavors of dynamical quarks. *Nucl. Phys. Proc. Suppl.*, 119:586–591, 2003.
- [59] G. P. Lepage et al. Constrained curve fitting. *Nucl. Phys. Proc. Suppl.*, 106:12–20, 2002.

- [60] Colin Morningstar. Bayesian curve fitting for lattice gauge theorists. *Nucl. Phys. Proc. Suppl.*, 109A:185–191, 2002.
- [61] A. Ukawa. Computational cost of full QCD simulations experienced by CP-PACS and JLQCD collaborations. *Nucl. Phys. Proc. Suppl.*, 106:195–196, 2002.
- [62] R. E. Wengert. A simple automatic derivative evaluation program. *Commun. ACM*, 7(8):463–464, 1964.
- [63] UKQCD Collaboration. Private communication, 2005.
- [64] C. Aubin et al. Light hadrons with improved staggered quarks: Approaching the continuum limit. *Phys. Rev.*, D70:094505, 2004.
- [65] S. Collins, R. G. Edwards, Urs M. Heller, and J. H. Sloan. Υ and J/Ψ spectroscopy using clover fermions in the presence of dynamical quarks. *Nucl. Phys. Proc. Suppl.*, 47:455–458, 1996.
- [66] Andreas S. Kronfeld. Binding energies in nonrelativistic field theories. *Nucl. Phys. Proc. Suppl.*, 53:401–404, 1997.
- [67] Aida X. El-Khadra. Phenomenology with lattice QCD: A consumer’s guide to lattice QCD results. In *Frascati Physics Series*, volume XI, pages 277–298, 1998. International Conference on Physics in Collision, Frascati, Italy, 17–19 Jun 1998.
- [68] Junpei Harada, Shoji Hashimoto, Andreas S. Kronfeld, and Tetsuya Onogi. Perturbative calculation of $\mathcal{O}(a)$ improvement coefficients. *Phys. Rev.*, D67:014503, 2003.
- [69] Masataka Okamoto et al. Semileptonic decays of D mesons in unquenched lattice QCD. *Nucl. Phys. Proc. Suppl.*, 129:334–336, 2004.
- [70] Junpei Harada et al. Application of heavy-quark effective theory to lattice QCD. II: Radiative corrections to heavy-light currents. *Phys. Rev.*, D65:094513, 2002.
- [71] M. Nobes et al. Private communication.

- [72] C. Bernard. Notes on heavy-light decay constants in staggered chiral perturbation theory. Private communication, March 2004.
- [73] S. Gottlieb et al. Onium masses with three flavors of dynamical quarks. To be published in Nucl. Phys. Proc. Suppl., 2006.
- [74] Gunnar S. Bali. The $D_{sJ}^+(2317)$: What can the lattice say? *Phys. Rev.*, D68:071501, 2003.
- [75] C. Aubin et al. Charmed meson decay constants in three-flavor lattice QCD. *Phys. Rev. Lett.*, 95:122002, 2005.
- [76] J. Simone et al. The decay constants f_{D_s} and f_{D^+} in three-flavor lattice QCD. To be published in Nucl. Phys. Proc. Suppl., 2006.
- [77] A. Kronfeld et al. Predictions from lattice QCD. To be published in Nucl. Phys. Proc. Suppl., 2006.
- [78] E. Eichten, K. Gottfried, T. Kinoshita, K. D. Lane, and T.-M. Yan. Charmonium: The model. *Phys. Rev.*, D17(11):3090–3117, 1978.
- [79] E. Eichten, K. Gottfried, T. Kinoshita, K. D. Lane, and T.-M. Yan. Charmonium: Comparison with experiment. *Phys. Rev.*, D21(1):203–233, 1980.
- [80] John L. Richardson. The heavy quark potential and the Υ , J/ψ systems. *Physics Letters*, 82B(2):272–274, 1979.
- [81] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in FORTRAN: The Art of Scientific Computing*. Cambridge University Press, 2nd edition, 1992.
- [82] Damian Menscher. SVD did not converge. Bug Report, August 2002. http://sourceforge.net/tracker?func=detail&atid=101369&aid=601052&group_id=1369.
- [83] C. Aubin and C. Bernard. Pion and kaon masses in staggered chiral perturbation theory. *Phys. Rev.*, D68:034014, 2003.

- [84] C. Aubin and C. Bernard. Pseudoscalar decay constants in staggered chiral perturbation theory. *Phys. Rev.*, D68:074011, 2003.
- [85] C. Aubin et al. Light pseudoscalar decay constants, quark masses, and low energy constants from three-flavor lattice QCD. *Phys. Rev.*, D70:114501, 2004.
- [86] A. Anastassov et al. First measurement of $\Gamma(D^{*+})$ and precision measurement of $m(D^{*+}) - m(D^0)$. *Phys. Rev.*, D65:032003, 2002.
- [87] C. Bernard. Taste splittings. <http://www.physics.wustl.edu/~cb/splittings.ps>, March 2004.
- [88] Z. Sroczynski et al. First results from the asymmetric $\mathcal{O}(a)$ improved Fermilab action. *Nucl. Phys. Proc. Suppl.*, 83-84:971–973, 2000.
- [89] M. B. Oktay, A. X. El-Khadra, A. S. Kronfeld, P. B. Mackenzie, and J. N. Simone. A relativistic $\mathcal{O}(a^2)$ improved action for heavy quarks. *Nucl. Phys. Proc. Suppl.*, 119:464–466, 2003.
- [90] M. B. Oktay, A. X. El-Khadra, A. S. Kronfeld, and P. B. Mackenzie. A more improved lattice action for heavy quarks. *Nucl. Phys. Proc. Suppl.*, 129:349–351, 2004.
- [91] John B. Kogut and Leonard Susskind. Hamiltonian formulation of Wilson’s lattice gauge theories. *Phys. Rev.*, D11:395, 1975.
- [92] Kenneth G. Wilson. Quantum chromodynamics on a lattice. In Maurice Levy and Pronob Mitter, editors, *New Developments in Quantum Field Theory and Statistical Mechanics*. Plenum Press, 1977. Cargese Summer Inst., Cargese, France, Jul 12–31, 1976.
- [93] Kenneth G. Wilson. Monte Carlo calculations for the lattice gauge theory. *NATO Adv. Study Inst. Ser. B Phys.*, 59:363–402, 1979.
- [94] Claude W. Bernard, Aida El-Khadra, and Amarjit Soni. Lattice calculation of meson form-factors for semileptonic decays. *Nucl. Phys. Proc. Suppl.*, 9:186–192, 1989.

- [95] D. Toussaint. Error analysis of simulation results: A sample problem. In *From Actions to Answers*, 1990. From the TASI-89 Summer School, Boulder, Colo., Jun 4–30, 1989.
- [96] G. Peter Lepage. Simulating heavy quarks. *Nucl. Phys. Proc. Suppl.*, 26:45–56, 1992. Lattice 91.
- [97] Claude W. Bernard and Maarten F. L. Golterman. Chiral perturbation theory for the quenched approximation of QCD. *Phys. Rev.*, D46:853–857, 1992.
- [98] M. Fukugita, N. Ishizuka, H. Mino, M. Okawa, and A. Ukawa. Full QCD hadron spectroscopy with two flavors of dynamical Kogut-Susskind quarks on the lattice. *Phys. Rev.*, D47:4739–4769, 1993.
- [99] Andreas S. Kronfeld. Lattice QCD. In *Perspectives in the Standard Model*, 1991. TASI Summer School, Boulder, CO, Jun 2–28, 1991.
- [100] Andreas S. Kronfeld and Paul B. Mackenzie. Progress in QCD using lattice gauge theory. *Ann. Rev. Nucl. Part. Sci.*, 43:793–828, 1993.
- [101] Estia J. Eichten, Christopher T. Hill, and Chris Quigg. Properties of orbitally excited heavy-light mesons. *Phys. Rev. Lett.*, 71:4116–4119, 1993.
- [102] Andreas S. Kronfeld. The charm quark on the lattice. *Nucl. Phys. Proc. Suppl.*, 30:445–448, 1993.
- [103] P. B. Mackenzie. Lattice QCD. In *QCD and Beyond*, 1995. Theoretical Advanced Study Institute in Elementary Particle Physics, Boulder, Colorado, 4–30 Jun 1995.
- [104] Thomas DeGrand. Nonperturbative quantum field theory on the lattice. In *Fields, Strings and Duality*, pages 885–945, 1996. Theoretical Advanced Study Institute in Elementary Particle Physics, Boulder, CO, 2–28 Jun 1996.
- [105] Mark G. Alford, T. R. Klassen, and G. P. Lepage. Towards highly improved quark actions. *Nucl. Phys. Proc. Suppl.*, 53:861–866, 1997.
- [106] Mark G. Alford, T. R. Klassen, and G. P. Lepage. Improving lattice quark actions. *Nucl. Phys.*, B496:377–407, 1997.

- [107] G. Peter Lepage. Redesigning lattice QCD. In *Perturbative and Nonperturbative Aspects of Quantum Field Theory*, pages 1–48, 1996.
- [108] M. Luscher and P. Weisz. $\mathcal{O}(a)$ improvement of the axial current in lattice QCD to one-loop order of perturbation theory. *Nucl. Phys.*, B479:429–458, 1996.
- [109] Martin Luscher, Stefan Sint, Rainer Sommer, Peter Weisz, and Ulli Wolff. Non-perturbative $\mathcal{O}(a)$ improvement of lattice QCD. *Nucl. Phys.*, B491:323–343, 1997.
- [110] Hartmut Wittig. Verification of $\mathcal{O}(a)$ improvement. *Nucl. Phys. Proc. Suppl.*, 63:47–52, 1998.
- [111] G. P. Lepage. How to renormalize the Schrödinger equation. In *Campos do Jordao 1997, Nuclear Physics*, pages 135–180, 1997. Jorge Andre Swieca Summer School on Nuclear Physics, Sao Paulo, Brazil, 26 Jan – 7 Feb 1997.
- [112] T. Yoshie. Light hadron spectroscopy. *Nucl. Phys. Proc. Suppl.*, 63:3–15, 1998.
- [113] Peter Lepage. Perturbative improvement for lattice QCD: An update. *Nucl. Phys. Proc. Suppl.*, 60A:267–278, 1998.
- [114] Stephan Gusken. Dynamical quark effects in QCD. *Nucl. Phys. Proc. Suppl.*, 63:16–21, 1998.
- [115] Martin Luscher. Advanced lattice QCD. In *Probing the Standard Model of Particle Interactions*, pages 229–280, 1998. Summer School in Theoretical Physics, Les Houches, France, 28 Jul – 5 Sep 1997.
- [116] G. P. Lepage. Lattice QCD for novices. In *Strong Interactions at Low and Intermediate Energies*, pages 49–90, 1998. HUGS AT CEBAF, Newport News, Virginia, 26 May – 2 Jun 1998.
- [117] J. L. Goity and W. Roberts. A relativistic chiral quark model for pseudoscalar emission from heavy mesons. *Phys. Rev.*, D60:034001, 1999.
- [118] Jonathan L. Rosner. Lattice QCD and heavy quarks. *Nucl. Phys. Proc. Suppl.*, 73:29–42, 1999.

- [119] C. T. H. Davies et al. Scaling of the Υ spectrum in lattice NRQCD. *Phys. Rev.*, D58:054505, 1998.
- [120] R. Kenway. Improvement and the continuum limit for quenched light hadrons. *Nucl. Phys. Proc. Suppl.*, 60A:26–33, 1998.
- [121] Kostas Orginos and Doug Toussaint. Testing improved actions for dynamical Kogut-Susskind quarks. *Phys. Rev.*, D59:014501, 1999.
- [122] Sinead M. Ryan, A. X. El-Khadra, A. S. Kronfeld, P. B. Mackenzie, and J. N. Simone. B and D semileptonic decays to light mesons. *Nucl. Phys. Proc. Suppl.*, 83:328–330, 2000.
- [123] Shoji Hashimoto. B decays on the lattice. *Nucl. Phys. Proc. Suppl.*, 83:3–13, 2000.
- [124] Stephen R. Sharpe and N. Shores. Partially quenched QCD with non-degenerate dynamical quarks. *Nucl. Phys. Proc. Suppl.*, 83:968–970, 2000.
- [125] Claude W. Bernard et al. Scaling tests of the improved Kogut-Susskind quark action. *Phys. Rev.*, D61:111502, 2000.
- [126] D. E. Groom et al. Review of particle physics. *The European Physical Journal*, C15:1, 2000.
- [127] Andreas S. Kronfeld. Application of heavy-quark effective theory to lattice QCD. I: Power corrections. *Phys. Rev.*, D62:014505, 2000.
- [128] A. X. El-Khadra, Steven A. Gottlieb, A. S. Kronfeld, P. B. Mackenzie, and J. N. Simone. A study of sea-quark effects on the charmonium spectrum. *Nucl. Phys. Proc. Suppl.*, 83:283–285, 2000.
- [129] Andreas S. Kronfeld. B and D mesons in lattice QCD. In *High Energy Physics*, pages 1322–1327, 2000.
- [130] Claude W. Bernard et al. f_B for various actions: Approaching the continuum limit with dynamical fermions. *Nucl. Phys. Proc. Suppl.*, 94:346–349, 2001.

- [131] Claude W. Bernard et al. The QCD spectrum with three quark flavors. *Phys. Rev.*, D64:054506, 2001.
- [132] J. N. Simone, S. Hashimoto, A. S. Kronfeld, P. B. Mackenzie, and Sinead M. Ryan. The $\bar{B} \rightarrow D^* \ell \bar{\nu}$ form factor at zero recoil and the determination of $|V_{cb}|$. *Nucl. Phys. Proc. Suppl.*, 106:394–396, 2002.
- [133] Junpei Harada, Shoji Hashimoto, Andreas S. Kronfeld, and Tetsuya Onogi. Application of heavy-quark effective theory to lattice QCD. III: Radiative corrections to heavy-heavy currents. *Phys. Rev.*, D65:094514, 2002.
- [134] Howard D. Trotter, G. Peter Lepage, Paul B. Mackenzie, Quentin Mason, and Matthew A. Nobes. Highly improved naive and staggered fermions. *Nucl. Phys. Proc. Suppl.*, 106:856–858, 2002.
- [135] Aida X. El-Khadra, Andreas S. Kronfeld, Paul B. Mackenzie, Sinead M. Ryan, and James N. Simone. The semileptonic decays $B \rightarrow \pi \ell \nu$ and $D \rightarrow \pi \ell \nu$ from lattice QCD. *Phys. Rev.*, D64:014502, 2001.
- [136] Massimo Di Pierro and Estia Eichten. Excited heavy-light systems and hadronic transitions. *Phys. Rev.*, D64:114004, 2001.
- [137] C. Bernard et al. Heavy-light meson decay constants with $N_f = 3$. *Nucl. Phys. Proc. Suppl.*, 119:613–615, 2003.
- [138] Andreas S. Kronfeld. Uses of effective field theory in lattice QCD. In M. Shifman, editor, *Handbook of QCD*, volume 4, chapter 39, pages 2411–2477. World Scientific, 2002.
- [139] Damir Bećirević, Saša Prelovšek, and Jure Zupan. $B \rightarrow \pi$ and $B \rightarrow K$ transitions in partially quenched chiral perturbation theory. *Phys. Rev.*, D68:074003, 2003.
- [140] Steven A. Gottlieb. Results on improved KS dynamical configurations: Decay constants, spectrum and topology. *Nucl. Phys. Proc. Suppl.*, 128:72–81, 2004.
- [141] A. Dougall, R. D. Kenway, C. M. Maynard, and C. McNeile. The spectrum of D_s mesons from lattice QCD. *Phys. Lett.*, B569:41–44, 2003.

- [142] Massimo di Pierro et al. D_s spectrum and leptonic decays with Fermilab heavy quarks and improved staggered light quarks. *Nucl. Phys. Proc. Suppl.*, 129:328–330, 2004.
- [143] Andreas S. Kronfeld. Heavy quarks and lattice QCD. *Nucl. Phys. Proc. Suppl.*, 129:46–59, 2004.
- [144] D. G. Cassel. Contributions of charm physics to CKM parameters. *ECONF*, C0304052:WG501, 2003.
- [145] Thomas DeGrand. Lattice QCD at the end of 2003. *Int. J. Mod. Phys.*, A19:1337–1394, 2004.
- [146] C. Aubin et al. Pion and kaon physics with improved staggered quarks. *Nucl. Phys. Proc. Suppl.*, 129:227–229, 2004.
- [147] M. Okamoto et al. Semileptonic $D \rightarrow \pi/K$ and $B \rightarrow \pi/D$ decays in $2 + 1$ flavor lattice QCD. *Nucl. Phys. Proc. Suppl.*, 140:461–463, 2005.
- [148] P. Mackenzie et al. Semileptonic decays and unitarity constraints. To be published in *Nucl. Phys. Proc. Suppl.*, 2006.

Curriculum Vitae

Education

- Ph.D. Physics, University of Illinois at Urbana-Champaign, October 2005
Thesis title: Charmonium and Charmed Mesons with Improved Lattice QCD
Computational Science and Engineering Certificate
- M.S. Physics, University of Illinois at Urbana-Champaign, August 1998
- B.S. Honors Physics, Brigham Young University, August 1997
Thesis title: Modeling the Quantum Computer on the Classical Computer

Selected Publications

- “Predictions from Lattice QCD”, To be published in Nucl. Phys. Proc. Suppl. (2006)
- “The Decay Constants f_{D_s} and f_{D^+} in Three-Flavor Lattice QCD”, To be published in Nucl. Phys. Proc. Suppl. (2006)
- “Onium Masses with Three Flavors of Dynamical Quarks”, To be published in Nucl. Phys. Proc. Suppl. (2006)
- “Charmed Meson Decay Constants in Three-Flavor Lattice QCD”, Phys. Rev. Lett. **95**, 122002 (2005); [hep-lat/0506030](#)
- “Leptonic Decay Constants f_{D_s} and f_D in Three Flavor Lattice QCD”, Lattice 2004; Nucl. Phys. Proc. Suppl. **140**, 443 (2005); [hep-lat/0410030](#)

- “High-Precision Lattice QCD Confronts Experiment”, Phys. Rev. Lett. **92**, 022001 (2004); [hep-lat/0304004](#)
- “ D_s Spectrum and Leptonic Decays with Fermilab Heavy Quarks and Improved Staggered Light Quarks”, Lattice 2003; Nucl. Phys. Proc. Suppl. **129**, 328 (2004); [hep-lat/0310045](#)
- “Properties of Charmonium in Lattice QCD with 2+1 Flavors of Improved Staggered Sea Quarks”, Lattice 2003; Nucl. Phys. Proc. Suppl. **129**, 340 (2004); [hep-lat/0310042](#)

Talks and Posters

- “Reclaiming Email: Spam and Virus Filtering”, talk presented for the ITG Forum, September 7, 2004
- “Forensics on Unix Systems”, Talk presented for SigMil, April 16, 2004
- “Blowfish and MD5”, Talk presented for SigMil, November 14, 2003
- “Charmonium with Three Flavors of Dynamical Quarks”, poster presented at Lattice 2002; Nucl. Phys. Proc. Suppl. **119**, 586 (2003); [hep-lat/0210051](#)
- “Practical Password Storage and Retrieval”, talk presented for the Information Protection Seminar, October 10, 2001
- “Improving Lattice QCD to $\mathcal{O}(a^2)$ ”, talk presented to UIUC Physics Department, December 20, 2000

Conferences

- Reflections | Projections, October 22 – 24, 2004
- DefCon 12, July 30 – August 1, 2004
- Lattice 2004, June 21 – 26, 2004

- Lattice Gauge Theory SciDAC Collaboration Meeting, February 21 – 22, 2003
- Lattice 2002, June 24 – 29, 2002
- Theoretical Advanced Study Institute in Elementary Particle Physics, June 4 – 30, 2000

Professional Experience

- UNIX Systems Administrator, June 1998 – present
- Computer Security Consultant, May 2000 – present
- Passed the “IRIX Advanced System Administration” and “IRIX Network Administration I” certification exams
- Graded General Field Theory, UIUC, August – December, 1999
- Teaching Assistant, BYU, September – December, 1996

Selected Honors and Professional Affiliations

- Eagle Scout
- Benson Scholar — Presidential Scholarship at Brigham Young University
- $\Sigma\Pi\Sigma$ — the national physics honor society
- American Institute of Physics