Exploration of the role of diquarks in hadrons using lattice QCD

by

Patrick S. Varilly

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Bachelor of Science in Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2006

Author .....
Department of Physics
May 19, 2006

Certified by .....
John W. Negele
William A. Coolidge Professor of Physics
Thesis Supervisor

Accepted by .....
Professor David E. Pritchard
Senior Thesis Coordinator, Department of Physics

Exploration of the role of diquarks in hadrons using lattice QCD

by

Patrick S. Varilly

Submitted to the Department of Physics
on May 19, 2006, in partial fulfillment of the
requirements for the degree of
Bachelor of Science in Physics

Abstract

We perform a number of measurements relevant to nuclear and particle physics by using the tools of lattice QCD. We verify our lattice calculations by reproducing published meson masses. We then study the light quark distribution in a meson with one heavy quark. After improving our methods in the meson case, we conclude by looking at the correlation between the two light quarks in a baryon. We find evidence for these quarks binding into spatially extended diquarks.

Thesis Supervisor: John W. Negele
Title: William A. Coolidge Professor of Physics

Acknowledgments

First and foremost, I'd like to thank my family: for instilling in me a passion for ideas, rather than an obsession with lucre; for the sacrifices they made in giving me and my brother the best possible education; and for their unwavering support in my journey through life. I'd also like to thank my friends at the Institute, for making MIT an enjoyable place to live, and for believing in my sanity, despite mounting evidence to the contrary. Finally, I'd like to thank Prof. John Negele: for taking a gamble when taking me on; for letting me do things my way; and most importantly, for showing me how physicists do physics.

Contents

Chapter 1

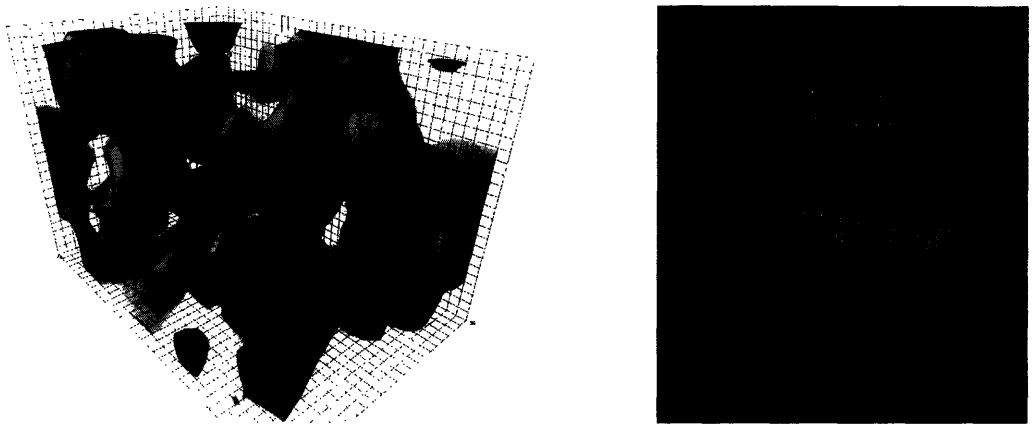
Introduction

Quantum chromodynamics (QCD) is the theory of the strong force. It postulates the existence of *quarks* and *gluons*, and describes their dynamics and interactions.

As usual, we can use perturbative expansions to calculate the predictions of QCD. Unfortunately, at the "low" energies that dominate life outside of particle accelerators, these expansions diverge. Therefore, in order to extract useful predictions, we must solve QCD with nonperturbative methods. Lattice QCD is the only known such method that solves QCD *exactly*. Lattice QCD makes spacetime discrete and finite, so the theory can now be solved numerically on powerful computers.

Lattice QCD can be used to calculate many of the fundamental properties of our world from first principles. Two such calculations are depicted in Figure 1: the complex structure of empty space, and the forces that bind two quarks into a meson. We can also use lattice QCD to calculate experimentally accessible observables, such as hadron masses [3] and lifetimes of unstable particles [17]. In recent years, precision lattice QCD has come to the forefront; for example, the world's most accurate determination of the strong coupling constant α_S was performed on a lattice [26].

In this thesis, we use lattice QCD to explore the structure of *diquarks* in baryons. Diquarks are pairs of quarks whose dynamics are strongly correlated. We study this correlation in the case of a baryon with a heavy quark and two light quarks forming a diquark. Our program is the following. First, we reproduce literature results on meson masses to gain experience with lattice QCD and to construct the computer



(a) (b)

Figure 1-1: Visualizations of QCD phenomena (from [16]). (a) Action density of the vacuum; (b) Reduction in action density (flux tube) caused by the presence of two interacting heavy quarks (arrows depict gradient of action density deficit)

codes that we use in more complex calculations. Then we study in detail the light quark distribution in a meson with one heavy quark. After improving our methods in the meson case, we conclude by looking at the correlation between the two light quarks in a baryon.

The structure of the following chapters is as follows. In Chapter 2, we discuss the theoretical framework of lattice QCD and how it's derived from continuum QCD. We also show how to calculate observables within this framework. In Chapter 3, we discuss the statistical errors that dominate lattice results, and ways of reducing them. Finally, in Chapter 4, we report the results of the physics program outlined above.

Chapter 2

Fundamentals of Lattice QCD

In this chapter, we give a brief summary of the general structure of lattice QCD, and give pointers to the literature where these ideas are fully developed. An expanded introduction to lattice QCD setup is given in [7]. The review article [21] covers many of the technical steps in detail, and places the subject in a general framework. Finally, these two works are well complemented by the Ph. D. thesis [25], which contains a highly pedagogical exposition of lattice QCD using the Wilson action. Our discussion closely follows this last reference.

Throughout, we use units where $\hbar = c = 1$.

2.1 Path integrals

The theory describing quarks and their interactions, QCD, is a quantum field theory. Consequently, any observable quantity can be expressed as a path integral of the general form

$$Z := \int [Dq\, D\bar{q}\, DA] e^{iS[\bar{q},q,A]} f[\bar{q}, q, A]. \tag{2.1}$$

The notation used is highly schematic. We begin by dissecting the various terms of (2.1) below.

The label q represents a *quark field*: at every spacetime point x, we define a complex vector with 12 entries, denoted by $q^a_\alpha(x)$. The upper index is a *color index*.

The number of quark colors is known from experiment to be 3, so a ranges from 1 to 3. The lower index is a *spin index*. Quarks are spin-$\frac{1}{2}$ particles, so their fields have 4 spin components. Additionally, a spin of $\frac{1}{2}$ forces quarks to be fermions, so quark fields anti-commute:

$$\{q_\alpha^a(x), q_\beta^b(y)\} = 0.$$

This means quark fields must be represented as complex Grassmann variables.

The label $\bar{q}$ analogously represents an anti-quark field. As usual, $\bar{q}(x) = q^\dagger(x)\gamma^0$.

The label A represents a *gluon field*: at every spacetime point x, we define a real vector with 32 entries, denoted by $A_\mu^c(x)$. For a fixed μ and x, the A field determines a unique element of the su(3), given by $\mathbb{A}_\mu^{ab}(x) := A_\mu^c(x)\lambda_c^{ab}$. Here, λ_c^{ab} are the eight 3×3 Gellman matrices. We call c the *color index* of A, and μ, it's direction, one of x, y, z, or t.

The *action* S is a functional of these three fields at every spacetime point. In the standard continuum formulation (see [24, ch. 15]), S is given by the formula

$$S[\bar{q}, q, A] := \int d^4x\, \mathcal{L}(q(x), \partial_\mu q(x), \bar{q}(x), \partial_\mu \bar{q}(x), A(x)),$$

where $\mathcal{L}$ denotes the QCD Lagrangian. We shorten the integrand to $\mathcal{L}(x)$. It's given by the equation

$$\mathcal{L}(x) := -\frac{1}{4} F_{\mu\nu}^a(x) F^{a\mu\nu}(x) + \bar{q}(x)[i\gamma^\mu D_\mu(x) - m]q(x) \tag{2.2}$$

where

$$F_{\mu\nu}^a(x) := \partial_\mu A_\nu^a(x) - \partial_\nu A_\mu^a(x) + g f^{abc} A_\mu^b(x) A_\nu^c(x), \tag{2.3}$$

and

$$D_\mu(x) = \partial_\mu(x) - igA_\mu^a(x)\lambda^a. \tag{2.4}$$

The integration measure $[Dq\, D\bar{q}\, DA]$ integrates over all possible quark, anti-quark and gluon configurations, with a normalization chosen such that $Z = 1$ when $f = 1$.

All the above elements are independent of the observable being calculated. In-

deed, it is the term $f[\bar{q}, q, A]$ which encodes the observable that Z measures; the fundamental relation is the following one:

$$\langle 0|\mathcal{T} f(\hat{\bar{q}}, \hat{q}, \hat{A})|0\rangle = \int [Dq\, D\bar{q}\, DA] e^{iS[\bar{q},q,A]} f[\bar{q}, q, A]. \tag{2.5}$$

Here, the hatted quantities are Heisenberg creation and annihilation operators, and $\mathcal{T}$ is the time-ordering operator. This correspondence is the thrust for using path integrals in quantum field theory calculations. Its precise construction in the case of scalar fields may be found in [4, ch. 1]; for an expanded discussion, the reader is referred to [22, ch. 2].

2.2 Continuation to imaginary time

To obtain information about ground-state elements, we make an important change to the standard path integrals: we continue to imaginary time. That is, we make the replacement $t \longrightarrow -it$. The change has at least three important consequences:

1. It renders path integrals mathematically well-defined. A host of subtle convergence issues are hidden by the $[Dq]$ notation above; these are absent in the imaginary time formulation.

2. The evolution operator $\exp(-iHt)$ becomes $\exp(-Ht)$. This change makes it practical to compute matrix elements of hadronic ground states, as explained below.

3. Spacetime becomes *Euclidean.* The metric for taking dot products of vectors changes from $g_{\mu\nu}$ to $\eta_{\mu\nu}$, so space and time are on a truly identical footing. As a practical consequence, raised and lowered spatial indices in tensorial quantities become indistinguishable. By convention, we lower all spatial indices. Another consequence is that the Dirac gamma matrices change, as described below.

The Euclidean counterparts of equations (2.1), (2.2), (2.4) and (2.3) are developed fully in [25, pp. 166–172], and are written down here:

$$Z_E := \int [Dq\, D\bar{q}\, DA] e^{-S[\bar{q},q,A]} f[\bar{q}, q, A]; \tag{2.6}$$

$$\mathcal{L}_E(x) := \frac{1}{4} F^a_{\mu\nu}(x) F^a_{\mu\nu}(x) + \bar{q}(x)[\gamma_\mu D_\mu(x) + m] q(x); \tag{2.7}$$

$$F^a_{\mu\nu}(x) := \partial_\mu A^a_\nu(x) - \partial_\nu A^a_\mu(x) - g f^{abc} A^b_\mu(x) A^c_\nu(x); \tag{2.8}$$

$$D_\mu(x) := \partial_\mu(x) + ig A^a_\mu(x) \lambda^a. \tag{2.9}$$

Note that we've: (a) redefined $A^a_0 \to i A^a_0$; (b) extracted an overall minus sign from $\mathcal{L}$ into the definition of Z_E; and (c) redefined the spatial γ matrices as $\gamma_i \to i\gamma_i$, leaving γ^0 unchanged. This last redefinition is consistent with a Euclidean metric, which requires that

$$\{\gamma_\mu, \gamma_\nu\} = \eta_{\mu\nu}.$$

2.3 Ground state observables

The Hamiltonian H will in general commute with a number of operators representing conserved quantities, or *quantum numbers*, like charge and momentum. Thus H has a block diagonal structure: the evolution operator doesn't mix states with different quantum numbers. Index these blocks by K, and index the eigenstates of H within each block by i. Denote the eigenstates by $|K, i\rangle$ and their energies by $E_{K,i}$. Order the blocks such that $E_{0,0} < E_{1,0} < \cdots$. With this notation, the unity operator can be written as

$$\hat{\mathbb{I}} = \sum_{K,i} |K, i\rangle\langle K, i|. \tag{2.10}$$

The most useful property of the continuation to imaginary time is the following: the evolution operator can be used to construct the ground state for a given set of quantum numbers. Suppose we have a state $|A\rangle$ satisfying $\langle M, i|A\rangle = 0$ for all $M < N$

and all i, but $\langle N,0|A\rangle \neq 0$. Then, by Equation (2.10),

$$\begin{aligned} e^{-H\tau}|A\rangle &= \sum_{K,i} e^{-H\tau}|K,i\rangle\langle K,i|A\rangle, \\ &= \sum_{K,i} |K,i\rangle e^{-E_{K,i}\tau}\langle K,i|A\rangle, \\ &= |N,0\rangle e^{-E_{N,0}\tau}\langle N,0|A\rangle + \mathcal{O}(e^{-(\Delta E)\tau}), \\ &\xrightarrow{\tau\to\infty} |N,0\rangle e^{-E_{N,0}\tau}\langle N,0|A\rangle. \end{aligned}$$

Here, ΔE is the energy difference between $|N,0\rangle$ and the next lowest energy eigenstate that overlaps $|A\rangle$.

We can exploit this property of Euclidean time to calculate the ground state matrix elements of any operator. Suppose we calculate the following matrix element:

$$X := \langle 0|\mathcal{T}e^{-HT}\hat{O}_f(x_f)\hat{O}_X(x)\hat{O}_i(x_i)|0\rangle.$$

Since the $\hat{O}$'s are Heisenberg operators, we can write them in terms of evolution operators and Schrödinger operators:

$$\hat{O}(\vec{x},t) := e^{Ht}\,\hat{O}(\vec{x})\,e^{-Ht}.$$

This change results in the following expansion for X:

$$X = \langle 0|e^{-H(T-t_f)}\,\hat{O}_f(\vec{x}_f)\,e^{-H(t_f-t)}\,\hat{O}_X(\vec{x})\,e^{-H(t-t_i)}\,\hat{O}_i(\vec{x}_i)\,e^{-Ht_i}|0\rangle.$$

The unique vacuum state $|\Omega\rangle$, with energy E_0, and the state $|0\rangle$ have the same quantum numbers. Assuming they overlap, $e^{-Ht}|0\rangle \to e^{-E_0 t}|\Omega\rangle$ as $t \to \infty$. Inserting unities around $|O\rangle$ operators, we thus get

$$\begin{aligned} X = e^{-E_0(T-t_f)}\langle\Omega|\hat{O}_f(\vec{x}_f)|A\rangle e^{-E_A(t_f-t)}\langle A|\hat{O}_X(\vec{x})|B\rangle e^{-E_B(t-t_i)}\langle B|\hat{O}_i(\vec{x}_i)|\Omega\rangle e^{-E_0 t_i} \\ + \text{terms of the form } \mathcal{O}(-\Delta E\tau). \end{aligned}$$

Here, the states $|A\rangle$ and $|B\rangle$, with energies E_A and E_B, are the lowest energy eigenstates overlapping $\hat{O}_f^\dagger(\vec{x}_f)|\Omega\rangle$ and $\hat{O}_i(\vec{x}_i)|\Omega\rangle$.

Suppose now that $|A\rangle = |B\rangle$. Then, rearranging the terms,

$$X = e^{-E_A(t_f-t_i)}\langle A|\hat{O}_X(\vec{x})|A\rangle e^{-E_0(T-t_f)}\langle\Omega|\hat{O}_f(\vec{x}_f)|A\rangle\langle A|\hat{O}_i(\vec{x}_i)|\Omega\rangle e^{-E_0 t_i} + \text{terms of the form } \mathcal{O}(-\Delta E\tau).$$

For any fixed operator $O_X(\vec{x})$, such as $\hat{\mathbb{I}}$, we see that $-dX/dt_f \to (E_A - E_0)$ as all time separations grow. This fact is used to compute the rest mass of state $|A\rangle$, for example a pion or a proton, in Section 4.1.

Moreover, consider the following matrix element:

$$Y := \langle 0|\mathcal{T}e^{-HT}\hat{O}_f(x_f)\hat{O}_i(x_i)|0\rangle.$$

Through similar manipulations, we obtain

$$Y = e^{-E_A(t_f-t_i)}e^{-E_0(T-t_f)}\langle\Omega|\hat{O}_f(\vec{x}_f)|A\rangle\langle A|\hat{O}_i(\vec{x}_i)|\Omega\rangle e^{-E_0 t_i} + \text{terms of the form } \mathcal{O}(-\Delta E\tau).$$

Thus, as time separations grow,

$$\frac{X}{Y} \to \langle A|\hat{O}_X(\vec{x})|A\rangle. \tag{2.11}$$

In this way, we can calculate ground state matrix elements of any operator.

2.4 Lattices and the Wilson action

To evaluate on a computer the infinite-dimensional functional integrals of Equation (2.6), we need to discretize spacetime, and all operators that depend on spacetime continuity.

We discretize spacetime first. Consider a finite cuboidal region spacetime of volume L^3 and time extent T. We impose a regular isotropic cubical grid with spacing

a, taking both L and T to be multiples of a (see Figure 2-1). We call this grid *the lattice*. Henceforth, we measure L and T in units of a, making them integers. For the present, we ignore the boundaries and focus on the bulk.

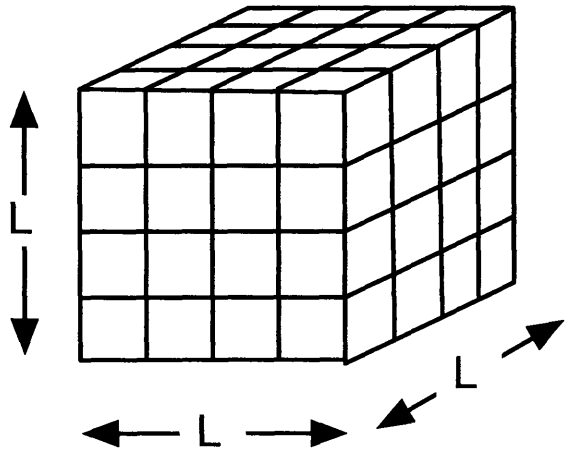


Figure 2-1: A 3-D timeslice of a 4-D spacetime lattice

For the quark fields, we can easily discretize the functional integral: associate with each lattice site x the 12 Grassman generators $q^a_\alpha(x)$ and their conjugates $\bar{q}^a_\alpha(x)$. The integration measure $[Dq\, D\bar{q}]$ is now a finite Grassman integral of dimension $24\times L^3\times T$, which we can evaluate.

We could do the same assignment for the gauge fields $A^a_\mu(x)$, but this leads to a path integral whose terms are not gauge invariant: gauge invariance would hold only in the *continuum limit* $a \to 0$.

Wilson [27] wrote down an alternate lattice discretization that conserves gauge invariance at finite a. His idea was to use not the $A^a_\mu(x)$'s as the fundamental degrees of freedom, but the *link variables* denoted by $U^{ab}_\mu(x)$, defined as

$$U^{ab}_\mu(x) = \mathcal{P}\exp\Big[ig\int_{P_\mu} A^c_\nu(r)\lambda^{ab}_c\, dr_\nu\Big].$$

Here, P_μ is the straight line path from $x + a\hat{\mu}$ to x, and $\mathcal{P}$ is the path-ordering operator. Notice that while $A^c_\mu(x)\lambda^{ab}_c \in \mathrm{su}(3)$, the link variable $U^{ab}_\mu(x) \in \mathrm{SU}(3)$. An

important property of link variables is the following:

$$(U^\dagger)^{ab}_\mu(x) = U^{ab}_{-\mu}(x + a\hat{\mu}). \tag{2.12}$$

Link variables have well-known gauge transformation properties. A gauge transform is specified by an SU(3) field $g^{aa'}(x) = \exp[i\alpha^c(x)\lambda_c]^{aa'}$. In the continuum, the following mapping leaves the action unchanged:

$$q^a_\alpha(x) \to g^{aa'}(x) q^{a'}_\alpha(x),$$
$$A^a_\mu(x) \to A^a_\mu(x) + \frac{i}{g}\partial_\mu \alpha^a(x).$$

(The gauge coupling g and the gauge change field $g^{aa'}(x)$ should not be confused). From these rules, it follows that,

$$U^{ab}_\mu(x) \to g^{aa'}(x) U^{a'b'}_\mu(x) (g^\dagger)^{b'b}(x + a\hat{\mu}).$$

The ∂_μ in the transformation of $A^a_\mu(x)$ is what makes the naive discretization of the gluon fields lack gauge invariance. If, instead, we write the lattice action in terms of only q's, $\bar{q}$'s and U's, we can make the lattice formulation gauge invariant.

Wilson discretized the gauge part of the Lagrangian,

$$\mathcal{L}_G(x) := \frac{1}{4} F^a_{\mu\nu}(x) F^a_{\mu\nu}(x) \quad \text{(continuum)}$$

as follows:

$$\mathcal{L}_G(x) := \sum_{\mu<\nu} \beta(1 - \frac{1}{N_c} \operatorname{Re} P_{\mu\nu}) \quad \text{(lattice)}. \tag{2.13}$$

Here, the number of colors (3) is denoted by N_c. The gauge-invariant *plaquette* $P_{\mu\nu}(x)$ is given by (suppressing matrix indices)

$$P_{\mu\nu}(x) := \operatorname{Tr}\left[U_\nu(x) U_\mu(x + \nu) U^\dagger_\nu(x + \mu) U^\dagger_\mu(x)\right], \tag{2.14}$$

and

$$\beta := \frac{2N_c}{g^2}.$$

As $a \to 0$, we can see that $\int d^4x\, \mathcal{L}_G^{\text{lat}}(x) \to \sum_x \mathcal{L}_G^{\text{con}}(x) + \mathcal{O}(a^2)$. The essentials of the proof can be seen when $N_c = 1$, where the algebra is a lot simpler because we're dealing only with phase factors, and not unitary matrices. Following [21, pp.12–15], we approximate

$$U_\mu^\dagger(x) \approx \exp[-igaA_\mu(x + a\hat{\mu}/2)]$$

Substituting this expansion into (2.14), we obtain

$$\begin{aligned} P_{\mu\nu}(x) &\approx \exp[iga[-A_\nu(x + a\hat{\nu}/2) - A_\nu(x + a\hat{\nu} + a\hat{\mu}/2) \\ &\qquad + A_\mu(x + a\hat{\mu} + a\hat{\nu}/2) + A_\mu(x + a\hat{\mu}/2)], \\ &\approx \exp[iga^2[\partial_\mu A_\nu(x) - \partial_\nu A_\mu(x)], \\ &= \exp[iga^2 F_{\mu\nu}]. \end{aligned}$$

When $N_c = 1$, we have that $\beta = \frac{1}{g^2}$: the factor of 2 arises from the relation $\text{Tr}[\lambda^a \lambda^b] = \frac{1}{2}\,\delta^{ab}$ for the fundamental representation of $\text{su}(N_c)$; the convention for U(1) can be changed to the standard convention by changing $A_\mu(x) \to A_\mu(x)/\sqrt{2}$, which we avoid here. Continuing,

$$\begin{aligned} \mathcal{L}_G^{\text{lat}} &\approx \beta \sum_{\mu<\nu} (1 - \cos(ga^2 F_{\mu\nu})), \\ &\approx a^4 \sum_{\mu<\nu} \frac{\beta}{2} g^2 F_{\mu\nu}^2, \\ &\approx a^4 \sum_{\mu,\nu} \frac{\beta}{4} g^2 F_{\mu\nu}^2, \\ &\approx a^4 \sum_{\mu,\nu} \frac{1}{4} F_{\mu\nu}^2. \end{aligned}$$

A more complete derivation that proves the $\mathcal{O}(a^2)$ error of Equation (2.13) for the general $\text{SU}(N)$ case is found in the reference above.

It remains to discretize the fermionic part of the action. To this end, note that

$D_\mu(x)$ is given by, and sometimes defined by, the equation

$$D_\mu^{ab}(x)q_\beta^b := \lim_{a\to 0} \frac{U_\mu^{ab}(x)q_\beta^b(x+a\hat{\mu}) - q_\beta^a(x)}{a}.$$

By stopping the limiting process at a finite but small a, we arrive at the lattice covariant derivative,

$$D_\mu^{ab}(x)q_\beta^b := \frac{1}{a}[U_\mu^{ab}(x)q_\beta^b(x+a\hat{\mu}) - q_\beta^a(x)] + \mathcal{O}(a) \quad \text{(forward difference)}.$$

As in standard finite difference schemes, we actually obtain better accuracy by using a center difference scheme:

$$D_\mu^{ab}(x)q_\beta^b := \frac{1}{2a}[U_\mu^{ab}(x)q_\beta^b(x+a\hat{\mu}) - U^{\dagger}{}_\mu^{ab}(x-a\hat{\mu})q_\beta^a(x-a\hat{\mu})] + \mathcal{O}(a^2) \quad \text{(center difference)}. \tag{2.15}$$

Now consider the standard fermionic Lagrangian part of (2.7),

$$\mathcal{L}_F(x) := \bar{q}(x)[\gamma_\mu D_\mu(x) + m]q(x).$$

Subtituting (2.15), we obtain

$$\begin{aligned}\mathcal{L}_F(x) = {} & m\bar{q}_\alpha^a(x)mq_\alpha^a(x) \\ & + \bar{q}_\alpha^a(x)\sum_{\mu=1}^{4}\frac{1}{2a}[\gamma_{\mu,\alpha\beta}U_\mu^{ab}(x)q_\beta^b(x+a\hat{\mu}) - \gamma_{\mu,\alpha\beta}U^{\dagger}{}_\mu^{ab}(x-a\hat{\mu})q_\beta^b(x-a\hat{\mu})].\end{aligned}$$

We can reduce this expression by using (2.12) adopting the following two conventions:

$$\gamma_{-\mu,\alpha\beta} = -\gamma_{\mu,\alpha\beta} \quad \text{and} \quad U^{\dagger}{}_\mu^{ab}(x) = U_{-\mu}^{ab}(x+a\hat{\mu}).$$

Then,

$$\mathcal{L}_F(x) = m\bar{q}_\alpha^a(x)q_\alpha^a(x) - \sum_{\mu=\pm 1}^{\pm 4}\frac{1}{2a}[\bar{q}_\alpha^a(x)\gamma_{\mu,\alpha\beta}U^{\dagger}{}_\mu^{ab}(x-a\hat{\mu})q_\beta^b(x-a\hat{\mu})]. \tag{2.16}$$

At this point we'd be done, except the Lagrangian above has a fatal problem commonly called *fermion doubling*. The lattice can only represent field configurations of momenta at least $\sim \pi/a$. It can be shown that in each dimension, to each low momentum mode there corresponds a high-momentum mode (near $\sim \pi/a$) with equal energy. Because our methods for extracting matrix elements of operators crucially rely on the properties of the energy spectrum (see Section 2.3), the problem is particularly serious.

Fermion doubling is an artifact of the discretization, and appears in other simpler contexts. For example, suppose we were solving the initial-value problem for the Dirac equation (in Minkowski spacetime) on a 3D L^3 grid:

$$i\gamma^0\partial_0\psi(x) = (m - i\gamma^i\partial_i)\psi(x).$$

If we discretize the ∂_μ with a center difference operator as we did with D_μ, then the right hand side is the same for $\psi(\vec{x}, t = t_0) = \exp(i\pi x/(La))f(y,z)$ as for $\psi(\vec{x}, t = t_0) = \exp(i\pi x(L-1)/(La))f(y,z)$, leading to two solutions with wildly different momenta that have the same energy. One solution here is to add a term involving the second derivative of ψ, which would raise the energy of the higher momentum solution; if the term is proportional to a, it drops out in the continuum limit:

$$(i\gamma^\mu\partial_\mu + a\partial_\mu\partial^\mu - m)\psi(x) = 0.$$

Adding an analogous second derivative term to Equation (2.16) leads to the *Wilson action*. Concretely, we'll add a term of the form $-(a/2)\bar{q}^a_\alpha(x)D^{aa'}_\mu(x)D^{a'b}_\mu(x)q^b_\beta(x)$ and discretize D^2 as a central-difference covariant Laplacian:

$$(D^2)^{ab}(x)q^b_\beta(x) := \frac{1}{a^2}\sum_{\mu=1}^{4}[U^{\dagger ab}_\mu(x - a\hat{\mu})q^b_\beta(x - a\hat{\mu}) - 2q^a_\beta(x) + U^{ab}_\mu(x)q^b_\beta(x + a\hat{\mu})].$$

Finally, performing the modification to Equation (2.16) yields

$$\mathcal{L}_F(x) = (m - \tfrac{4}{a})\bar{q}^a_\alpha(x)q^a_\alpha(x) - \sum_{\mu=\pm1}^{\pm4} \frac{1}{a}[\bar{q}^a_\alpha(x)\tfrac{1}{2}(1+\gamma_\mu)_{\alpha\beta}U^{\dagger ab}_\mu(x-a\hat{\mu})q^b_\beta(x-a\hat{\mu})]. \quad (2.17)$$

The Wilson action has a unique property: the spin matrices linking quark fields at adjacent sites are *spin projectors*. This fact is enormously useful. Most importantly, it's essential to building a *transfer matrix* interpretation of path integrals; that is, to show that Equation (2.5) holds exactly for the Wilson action at finite lattice spacing [18]. It also allows the construction of good preconditioners (the so-called *even-odd* preconditioners) used when numerically solving for propagators as described below. It also finds use in the hopping parameter expansion described below as well.

One more step remains: to find the analog to $[DA^a_\mu(x)]$. Since we've made the link variables our primary objects, we would like to integrate over all their possible values in some "uniform" manner. The desired integration measure is called the *group-invariant measure*. We describe it here only in passing and refer the reader to [7, ch. 8] for further details.

Essentially, we want to give meaning to the integral

$$\int dU f(U), \quad U \in \mathrm{SU}(N).$$

When $N = 1$, an obvious candidate emerges:

$$\int dU f(U) \xrightarrow{N=1} \frac{1}{2\pi}\int_{-\pi}^{\pi} d\theta\, f(e^{i\theta}).$$

The integration is "uniform" in the sense that $\frac{1}{2\pi}\int_{-\pi}^{\pi} d\theta\,(1) = 1$ and

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} d\theta\, f(e^{i\theta}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} d\theta\, f(e^{i(\theta+\alpha)}) \quad \text{for any } \alpha \in \mathbb{R}.$$

That is, no θ-direction is privileged by the integral. The two properties can be

generalized to any compact group G as follows

$$\int dg\, 1 = 1 \quad \text{and} \quad \int dg f(g) = \int dg f(g'g) \quad \text{for all } g' \in G.$$

It can be shown these two properties define a unique integration measure. Concretely, let G be a continuous group parametrized by N real parameters denoted by α_i. We can define a *metric tensor* $M_{ij}(\vec{\alpha})$ as follows:

$$M_{ij}(\vec{\alpha}) = \mathrm{Tr}[g^{-1}(\partial_i g)g^{-1}(\partial_j g)],$$

where $g = g(\vec{\alpha})$ and $\partial_i = \partial/\partial\alpha_i$. The group-invariant integral is then given by

$$\int dg\, f(g) = K \int d\vec{\alpha}\, |\det M(\vec{\alpha})|^{1/2} f(g(\vec{\alpha})),$$

with K a normalization factor needed to obtain $\int dg\, 1 = 1$.

In this manner, we can integrate over the value of the all link variables $U_\mu^{ab}(x)$. A useful consequence is that the gauge integration is now over a compact domain (the parameters of the group elements), instead of over all of $\mathbb{R}^N$ for some N large as is needed in the continuum. In particular, the set of gauge configurations gauge-equivalent to a given $U_\mu^{ab}(x)$ field is also compact, so integrating over it doesn't produce infinite answers. Indeed, there is no need to gauge fix in the lattice.

To conclude this section, let's summarize the essential results. We discretize spacetime into a finite, evenly spaced grid (spacing a) with $L^3 \times T$ sites. To each site x, we associate:

- The 12 Grassmann generators $q_\alpha^a(x)$;
- Their conjugate generators $\bar{q}_\alpha^a(x)$;
- An SU(3) 4-vector $U_\mu^{ab}(x)$.

We then rewrite the integral (2.6) as follows:

$$Z \to \int \prod_x \left[\prod_\mu dU_\mu(x) \prod_{a,\alpha} (dq^a_\alpha(x) d\bar{q}^a_\alpha(x)) \right] e^{-S[\bar{q},q,U]} f(\bar{q},q,U) \quad \text{(no sum over } a,\alpha\text{)}. \tag{2.18}$$

The action we use is the *Wilson action*:

$$S[\bar{q},q,U] := a^4 \sum_x \left[\sum_{\mu<\nu} \beta(1 - \frac{1}{N_c} \operatorname{Re} P_{\mu\nu}) \right.$$
$$\left. + (m - \tfrac{4}{a}) \bar{q}^a_\alpha(x) q^a_\alpha(x) - \sum_{\mu=\pm 1}^{\pm 4} \frac{1}{a} [\bar{q}^a_\alpha(x) \tfrac{1}{2}(1+\gamma_\mu)_{\alpha\beta} U^{\dagger}{}^{ab}_{\mu}(x - a\hat{\mu}) q^b_\beta(x - a\hat{\mu})] \right]$$

The quark and gluon fields transform as follows under a gauge transform $g^{aa'}(x)$:

$$\begin{aligned} q^a_\alpha(x) &\to g^{aa'}(x) q^{a'}(x); \\ \bar{q}^a_\alpha(x) &\to \bar{q}^{a'}(x) (g^\dagger)^{a'a}(x); \\ U^{ab}_\mu(x) &\to g^{aa'}(x) U^{a'b'}_\mu(x) (g^\dagger)^{b'b}(x + a\hat{\mu}); \\ U^{\dagger}{}^{ab}_{\mu}(x) &\to (g^\dagger)^{aa'}(x + a\hat{\mu}) U^{a'b'}_\mu(x) g^{b'b}(x). \end{aligned}$$

The grid spacing a functions as a regulator for the path integrals. As $a \to \infty$, we approach the physical *continuum limit*, where the results of our calculations should match experiments.

At the beginning of the section, we glossed over the issue of boundary conditions. We can now deal with it appropriately. In most cases, it is appropriate to use periodic boundary conditions in space. This is much like solid-state physics, where you simulate one unit cell of a crystal. In that case, it is desirable for a wavefunction and its periodic images to interact, since that's the physics that you want to capture. In QCD, periodic images are an artifact that must be corrected for by either (a) computing in a volume L^3 large enough that image effects are negligible; or (b) assuming some form of the effect for images on matrix elements and compensating for it (see Section 4.2). We can enforce periodic boundary conditions along the ν axis by setting $U_\mu(x) = U_\mu(x + L\hat{\nu})$. Antiperiodic boundary conditions, on the other hand, result

from setting $U_\mu(x) = -U_\mu(x + L\hat{\nu})$. Finally, hard-wall boundary conditions in time faithfully correspond to taking the expectation value of the path integral integrand with respect to $|0\rangle$. They are implemented in practice by setting $U_t^\dagger(\vec{x}, (T-1)) = 0$, which has the equivalent effect of deleting all the terms in the action that link quark and gluon fields at $t = 0$ to those at $t = (T-1)$.

2.5 Evaluating the path integral

2.5.1 Breaking down the path integral

To calculate matrix elements, we have to evaluate the immense, but finite, integral in Equation (2.18). The strategy can be summarized as follows. The gauge integral can be written as a regular integral over a finite domain, but the Grassmann integral over the quark fields cannot. However, for a given gauge field, this quark field integral is a Gaussian integral, so we can evaluate it analytically using Wick's theorem. We then calculate the gauge integral with Monte Carlo techniques.

We now elaborate on this prescription. Equation (2.18) can be written as follows:

$$Z = \int DU\, e^{-S_G[U]} \int d\bar{q}\, dq\, e^{-S_F[\bar{q},q,U]} f(\bar{q}, q, U).$$

We've introduced condensed notation for the integration measures and split the action into: (a) terms independent of q and $\bar{q}$, collected in S_G, which stands for *gauge action*; and (b) all other terms, collected in S_F, which stands for *fermion action.*

Let's introduce some additional conventions and notations to simplify our expressions further. We begin by using units of length in which $a = 1$, so we'll stop including a's in our formulas. Next, for a given U field, we define the following so-called *D-slash* operator:

$$(\not{D}q)^a_\alpha(x) := \sum_{\mu=\pm1}^{\pm4} (1 + \gamma_\mu)_{\alpha\beta} U^{\dagger}{}^{ab}_{\mu}(x - a\hat{\mu}) q^b_\beta(x - a\hat{\mu}).$$

The $\not{D}$ operator can also be viewed as a matrix over color, spin *and* space, as follows:

$$\not{D}^{ab}_{\alpha\beta}(x,y) := \sum_{\mu=\pm 1}^{\pm 4} (1+\gamma_\mu)_{\alpha\beta} U^{\dagger ab}_{\mu}(x-a\hat{\mu})\delta_{x-a\hat{\mu},y}, \tag{2.19}$$

where δ is the Kronecher delta. Thus $(\not{D}q)^a_\alpha(x) = \sum_y \not{D}^{ab}_{\alpha\beta}(x,y) q^b_\beta(y)$.

The fermion action is given by the formula

$$S_F[\bar{q},q,U] = \sum_x (m+4)\bar{q}^a_\alpha(x) q^a_\alpha(x) - \tfrac{1}{2}\bar{q}^a_\alpha(x)(\not{D}q)^a_\alpha(x).$$

We now introduce the parameter $\kappa := 1/(2m+8)$. For reasons discussed below, we call κ the *hopping parameter.* Next, change variables $q^a_\alpha(x) \to \sqrt{2\kappa}\, q^a_\alpha(x)$ and $\bar{q}^a_\alpha(x) \to \sqrt{2\kappa}\, \bar{q}^a_\alpha(x)$. This introduces an overall factor of $\kappa^{2\times 3\times 4\times L^3\times T}$ into the measure, which we can redefine away, and changes the fermion action to

$$S_F[\bar{q},q,U] \to \sum_x \bar{q}^a_\alpha(x) q^a_\alpha(x) - \kappa \bar{q}^a_\alpha(x)(\not{D}q)^a_\alpha(x).$$

Finally, we introduce the *Wilson-Dirac matrix*, denoted by M:

$$M^{ab}_{\alpha\beta}(x,y) := \delta^{ab}_{\alpha\beta}\delta_{x,y} - \kappa \not{D}^{ab}_{\alpha\beta}(x,y).$$

With this definition,

$$S_F[\bar{q},q,U] = \sum_{x,y} \bar{q}^a_\alpha(x) M^{ab}_{\alpha\beta}(x,y) q^a_\alpha(y).$$

Note that M depends implicitly on the U field through $\not{D}$. In more condensed matrix notation, where q and $\bar{q}$ are a column vector and a row vector, both indexed by color, spin *and lattice site*, the above formulas read

$$M = \hat{\mathbb{I}} - \kappa \not{D}, \tag{2.20}$$

and

$$S_F = \bar{q} M q.$$

There are three main reasons for rewriting the fermion equations as we did above:

1. The fermion integral is now manifestly a Gaussian integral, which can be calculated by Wick's theorem.

2. The $\not{D}$ operator is used extensively in all lattice calculations: its evaluation consumes the bulk of the computational time of lattice QCD codes. By isolating it and writing all lattice expression in terms of $\not{D}$, a concerted optimization effort can be made *once* on a single $\not{D}$ implementation, and the benefits follow immediately to all dependent codes [19].

3. The compactness of formulas using $\not{D}$ eases their manipulation. In particular, the hopping parameter expansion described below can be derived immediately from Equation (2.20).

In the present context, Wick's theorem reads:

$$\begin{aligned}
&\int d\bar{q}\, dq\, e^{-\bar{q}Mq} q^{a_1}_{\alpha_1}(x_1) \bar{q}^{b_1}_{\beta_1}(y_1) \cdots q^{a_N}_{\alpha_N}(x_N) \bar{q}^{b_N}_{\beta_N}(y_N) \\
&\quad = (\det M)(2\kappa)^N \sum_P (-1)^P (M^{-1})^{a_1 b_{P1}}_{\alpha_1 \beta_{P1}}(x_1, y_{P1}) \cdots (M^{-1})^{a_N b_{PN}}_{\alpha_N \beta_{PN}}(x_N, y_{PN}). \quad (2.21)
\end{aligned}$$

Here, P is the set of all permutations of the integers $1, \ldots, N$, and $(-1)^P$ denotes the P's sign.

In the *quenched approximation*, the determinant $(\det M)$ is ignored, which *significantly* reduces computation time. It can be shown [21] that this replacement is equivalent to suppressing the effect of sea quarks, that is, virtual quark-antiquark pairs that pop in and out of the vacuum. Empirically, observables calculated within this approximation are 10–15% different from experiment [3]. Throughout this thesis, we'll always work in the quenched approximation.

Once Wick's theorem is applied, we must still do a gauge integral of the form

$$I := \int DU e^{-S_G[U]} g(U).$$

This we do by Monte Carlo integration: we generate a finite set of gauge fields U, picked at random with a probability distribution $\exp(-S_G[U])$, then take the average value of g with respect to them. The mechanics of generating such gauge fields will not be described here; the interested reader may consult [7, 6] for further details. An important aspect of Monte Carlo algorithms is that they are statistical in nature, so the calculated integrals will have quantifiable errors associated with them. We estimate these errors with the *jackknife method*, summarised in Section 2.7.

2.5.2 Propagators

As can be gleaned from Equation (2.21), the fundamental building blocks of any lattice calculation are the entries of the matrix M^{-1}. If we computed all elements of M^{-1}, then the calculations would be unmanageable by present standards. The matrix is square with $3 \times 4 \times L^3 \times T$ rows and columns of complex numbers. Moreover, although M is sparse, its inverse is dense. Simply storing the result for a typical lattice ($16^3 \times 32$) would take up 4.6 TB of space; calculating it would be prohibitively expensive. So instead of computing all of M^{-1}, we compute a few linear combinations of its columns, which we refer to as *propagators*. In perturbative language, these correspond to dressed quark propagators.

Essentially, we want to calculate the quantity $P^{ab}_{\alpha\beta}(x)$ in the following expression:

$$P^{ab}_{\alpha\beta}(x) := \sum_y (M^{-1})^{ac}_{\alpha\gamma}(x,y) S^{cb}_{\gamma\beta}(y),$$

by solving the linear system

$$\sum_x M^{ca}_{\gamma\alpha}(y,x) P^{ab}_{\alpha\beta}(x) = S^{cb}_{\gamma\beta}(y). \tag{2.22}$$

The vector $S^{cb}_{\gamma\beta}(y)$ is called the *source*, and specifies a set of linear combination of M^{-1}'s columns. For example, to get M^{-1} for a fixed $y = Y$, we'd solve (2.22) with $S^{cb}_{\gamma\beta}(y) = \delta^{cb}_{\gamma\beta}\delta_{y,Y}$. Section 3.1 discusses the use of nontrivial sources.

In practice, Equation (2.22) is solved piecewise. We run through all 12 possible values of (b, β) and fix them in turn in the equation. This results in 12 standard matrix-vector linear system, which can be solved iteratively with a numerical method such as conjugate gradient or minimum residue.

Before proceeding, let's examine M^{-1} further. Starting from (2.20), we can expand M^{-1} in powers of κ as follows:

$$M^{-1} = 1 + \kappa \not{D} - \kappa^2 \not{D}^2 + \kappa^3 \not{D}^3 - \cdots . \tag{2.23}$$

This expansion is called the *hopping parameter expansion*, and has a beautiful graphical interpretation.

Let $P_\mu := \frac{1}{2}(1 + \gamma_\mu)$ be the spin projection matrices along the $\hat{\mu}$ direction (not to be confused with the P vector in Equation (2.22)). Then, suppressing spin, color but not space indices, equation (2.19) reads

$$\not{D}(x, y) := \sum_{\mu=\pm 1}^{\pm 4} P_\mu U^\dagger{}_\mu(x - \hat{\mu})\delta_{x-\hat{\mu},y}.$$

Square this operator to get

$$\not{D}^2(x, y) = \sum_z \not{D}(x, z)\not{D}(z, y) = \sum_{\mu=\pm 1}^{\pm 4} \sum_{\nu=\pm 1}^{\pm 4} P_\mu U^\dagger{}_\mu(x - \hat{\mu}) P_\nu U^\dagger{}_\nu(x - \hat{\mu} - \hat{\nu})\delta_{x-\hat{\mu}-\hat{\nu},y}.$$

In general, let $\hat{e}_1, \ldots, \hat{e}_L$ be a sequence of L directions, with $\hat{e}_i \in \{\pm\hat{x}, \pm\hat{y}, \pm\hat{z}, \pm\hat{t}\}$. Then,

$$\not{D}^L(x, y) = \sum_{\substack{\text{paths } \hat{e}_i \\ \text{of length } L \\ \text{to } x \text{ from } y}} \prod_{i=1}^{L} [P_{\hat{e}_i} U^\dagger{}_{\hat{e}_i}(x - \textstyle\sum_{j \le i} \hat{e}_j)],$$

Substituting this expression into (2.23) we finally obtain the following:

$$M^{-1}(x,y) = \sum_{L=0}^{\infty}\left[(-\kappa)^L \sum_{\substack{\text{paths } \hat{e}_i \\ \text{of length } L}} \prod_{i=1}^{L}[P_{\hat{e}_i}U^{\dagger}_{\hat{e}_i}(x - \textstyle\sum_{j\le i}\hat{e}_j)] \cdot \delta_{x-\sum_i \hat{e}_i, y}\right]. \qquad (2.24)$$

We note the following fact in passing. The P's and the U's commute with each other, as they are spin and color matrices respectively. So we may rearrange each term in (2.24) as a product of: κ^L, a product of spin projectors and a product of link variables. By virtue of P_μ being a projector, we have $P_\mu P_{-\mu} = 0$. Thus, we can ignore all paths that go back on themselves.

We can "draw" (2.24) by drawing all the paths to x from y, and associating to it the respective propagator matrix. This construction is shown in Figure 2-2. Recall that M^{-1} results from Wick's theorem as follows:

$$(M^{-1})^{ab}_{\alpha\beta}(x,y) = \frac{1}{2\kappa}\int d\bar{q}\, dq\, e^{-\bar{q}Mq} q^a_\alpha(x)\bar{q}^b_\beta(y).$$

This corresponds to creating a quark at y and annihilating it at x. The propagator is literally the "sum over paths" of the accumulated color and spin rotations that the quark undergoes as it travels from y to x, with heavier quarks (higher m) suppressing longer paths (because they have a lower κ). We can represent full matrix elements diagramatically by showing representative paths in their propagators.

An important limit we use in this thesis is that of infinitely massive quarks, which we call *heavy quarks*. In particular, note that as $m \to \infty$, we get $\kappa \to 0$, so the shortest paths from x to y dominate. We denote heavy quarks and their generators by the subscript H. Physically, we know an infinitely heavy quark cannot move in space, so $(M_H^{-1})(\vec{x}, t_f; \vec{y}, t_i) = 0$ if $\vec{x} \neq \vec{y}$. Assume $t_f > t_i$. Then it follows from the hopping parameter expansion that:

$$\kappa^{-(t_f - t_i)}(M_H^{-1})(\vec{x}, t_f; \vec{x}, t_i) = \tfrac{1}{2}(1+\gamma_0)U_t^{\dagger}(\vec{x}, t_i - 1)\cdots U_t^{\dagger}(\vec{x}, t_f) + \mathcal{O}(\kappa). \qquad (2.25)$$

The factors of $\kappa^{-(t_f - t_i)}$ typically cancel in the ratios of path integrals we calculate

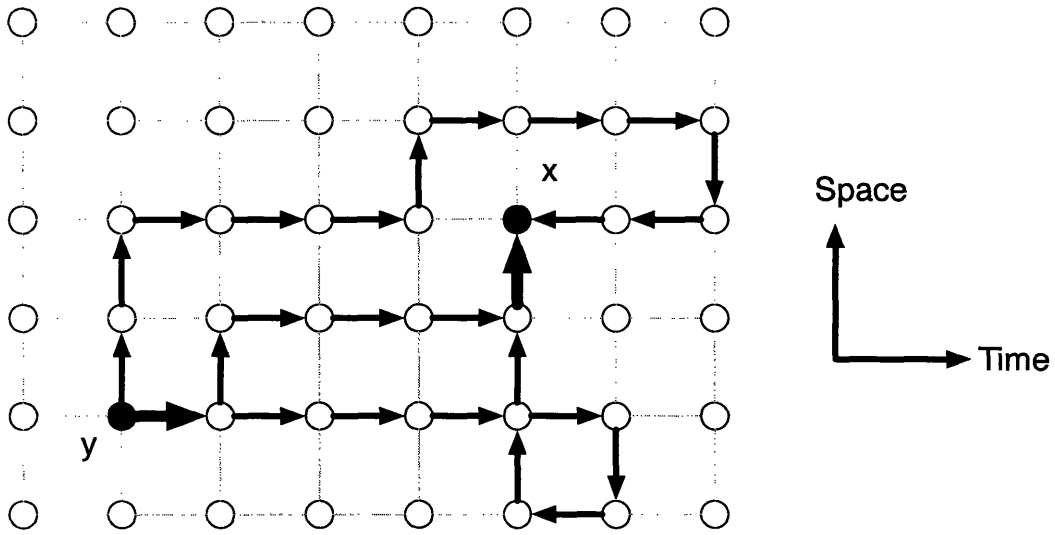


Figure 2-2: Three representative paths in the hopping parameter expansion of $M^{-1}(x,y)$. Notice how paths can go backwards in time (quark-antiquark pairs) and how paths can self-cross. The thick arrows denote link variables shared by two distinct paths.

(see Equation (2.11)).

There is one further crucial property of propagators, summarized by the formula

$$\gamma^5 (M^{-1})^\dagger (x,y) \gamma^5 = (M^{-1})(y,x), \tag{2.26}$$

where $\gamma_5 := \gamma_1\gamma_2\gamma_3\gamma_4$. That is, knowing $M^{-1}(x,y)$ (sometimes called a *forward propagator*), we can compute $M^{-1}(y,x)$ (correspondingly called a *backward propagator*). The importance of this trick is computational. The matrix elements we can compute have either the source location (y) or the sink location (x) fixed. If we could only use the linear system in (2.22) to calculate propagators, only the source location could be fixed.

Equation (2.26) can be shown by directly computing $\gamma^5 M^\dagger(x,y)\gamma^5 = M(y,x)$ and taking the inverse of both sides (since $\gamma_5\gamma_5 = 1$, we have $\gamma_5^{-1} = \gamma_5$).

2.6 Matrix elements evaluated in this thesis

In this thesis, we study the properties of heavy-light and heavy-light-light systems. As model particles for both, we take a pseudoscalar kaon and a sigma. We need operators J that act on the vacuum to annihilate states with the quantum numbers of these particles; the operator $\bar{J} := J^\dagger \gamma_0$ can then be used to create such particles from the vacuum.

Below, we used the symbols u and d to mean light quark fields (no contractions between u and d quarks) and s to mean a heavy quark. These are the sources we've used:

- For the meson case, we use a pseudoscalar kaon source $J(x) = \bar{s}^a_\alpha(x)\gamma_{5,\alpha\beta}\delta^{ab}d^b_\beta(x)$. In more condensed notation, this reads $J = \bar{s}\gamma_5 d$. The conjugate source is $\bar{J} = -\bar{d}\gamma^5 s$.

- For the baryon case, we use a Σ^0 source $J_\alpha(x) = \epsilon^{abc}(u^b_\beta(x)\Gamma_{\beta\gamma}d^c_\gamma(x))s^a_\alpha(x)$. We used $\Gamma = C\gamma^5 = \gamma^1\gamma^3$, corresponding to the "good" diquark sources of [1]. In more condensed notation, this reads $J = (u^T C\gamma^5 d)s$. The conjugate source is $\bar{J} = -\bar{s}(\bar{d}^T C\gamma^5 \bar{u})$. When evaluating baryon matrix elements, we project onto the positive parity states only ($J_\alpha \mapsto \frac{1}{2}(1+\gamma^0)_{\alpha\alpha'}J_{\alpha'}$), and average over both spins.

In the continuum theory, the operator

$$j^\mu_q(x) := \bar{q}(x)\gamma^\mu q(x) \tag{2.27}$$

is a Nöther conserved charge of the QCD Lagrangian. That is, $\partial_\mu j^\mu_q = 0$, which is the relativistic continuity equation encoding local charge conservation of quark type q. In particular, $j^0_q(x)$ is the *density* of quarks of type q at x. It is the closest object in field theory to a probability density in ordinary quantum mechanics. We denote it with a symbol of its own:

$$\rho_q(x) := \bar{q}(x)\gamma^0 q(x).$$

In the lattice, the current defined in (2.27) is not exactly conserved. Instead, by applying Nöther's theorem to the Wilson action, we find the following *point-split current* is conserved:

$$j_q^{\mu\mathrm{PS}}(x) := \bar{q}(x+\hat{\mu})\tfrac{1}{2}(1+\gamma^\mu)U_\mu^\dagger(x)q(x) - \bar{q}(x)\tfrac{1}{2}(1-\gamma^\mu)U_\mu(x)q(x+\hat{\mu}).$$

In the continuum limit, the U factors tend to 1 while $x+\hat{\mu} \to x$, so we recover (2.27). Analogously, the continuum current should be "almost" conserved on the lattice, with deviations decreasing as $a \to 0$.

We experimented with both types of current during the course of this work, but eventually settled on the continuum current: the computational effort involved is substantially smaller and the continuum current is more useful when averaging a density over many timeslices as discussed in section 3.4.

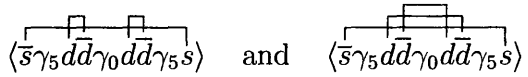
As an example of how these matrix elements were calculated, we derive the density expectation for a meson in terms of propagators. The expression to calculate is:

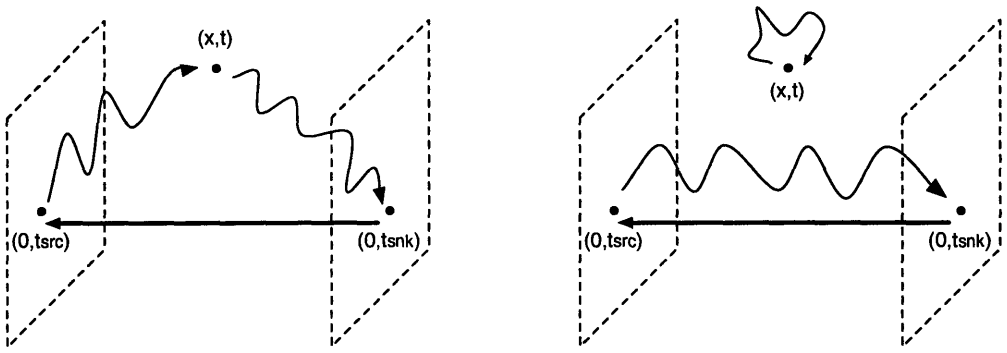
$$\langle\rho(\vec{x})\rangle := \frac{\langle J(\vec{0},t_{\mathrm{snk}})\rho(\vec{x},t)\bar{J}(\vec{0},t_{\mathrm{src}})\rangle}{\langle J(\vec{0},t_{\mathrm{snk}})\bar{J}(\vec{0},t_{\mathrm{src}})\rangle}. \tag{2.28}$$

Here, t_{src} and t_{snk} are two arbitrary times both far enough from t that the ground state of source J has been filtered out. Expanding J and ρ in terms of quark fields we get the following expression for the numerator,

$$\begin{aligned}&\langle J(\vec{0},t_{\mathrm{snk}})\rho(\vec{x},t)\bar{J}(\vec{0},t_{\mathrm{src}})\rangle \\ &\qquad = -\langle \bar{s}^a_\alpha(\vec{0},t_{\mathrm{snk}})\gamma_{5,\alpha\beta}d^a_\beta(\vec{0},t_{\mathrm{snk}})\bar{d}^c_\gamma(\vec{x},t)\gamma^0_{\gamma\delta}d^c_\delta(\vec{x},t)\bar{d}^e_\epsilon(\vec{0},t_{\mathrm{src}})\gamma_{5,\epsilon\eta}s^e_\eta(\vec{0},t_{\mathrm{src}})\rangle.\end{aligned}$$

Schematically, we can take the following two Wick contractions:





(a) (b)

Figure 2-3: Visual representations of the two possible contractions of a heavy-light meson density insertion (see (2.29)).

When all is said and done, we obtain the following expression for the numerator:

$$\begin{aligned}\langle J(\vec{0},t_{\text{snk}})\rho(\vec{x},t)\bar{J}(\vec{0},t_{\text{src}})\rangle = \qquad (2.29)\\ \gamma_{5,\epsilon\eta}(M_H^{-1})^{ea}_{\eta\alpha}(\vec{0},t_{\text{src}};\vec{0},t_{\text{snk}})\gamma_{5,\alpha\beta}(M^{-1})^{ac}_{\beta\gamma}(\vec{0},t_{\text{snk}};\vec{x},t)\gamma^0_{\gamma\delta}(M^{-1})^{ce}_{\delta\epsilon}(\vec{x},t;\vec{0},t_{\text{src}})\\ -\gamma_{5,\epsilon\eta}(M_H^{-1})^{ea}_{\eta\alpha}(\vec{0},t_{\text{src}};\vec{0},t_{\text{snk}})\gamma_{5,\alpha\beta}(M^{-1})^{ae}_{\beta\epsilon}(\vec{0},t_{\text{snk}};\vec{0},t_{\text{src}})\gamma^0_{\gamma\delta}(M^{-1})^{cc}_{\delta\gamma}(\vec{x},t;\vec{x},t).\end{aligned}$$

This equation is shown diagrammatically in Figure 2.6.

The first term of (2.29) is now straightforward (if laborious) to compute using the methods of the previous sections. The second term, usually called a *loop diagram* (see Fig. 2.6(b)), is more problematic, because to compute it, we'd have to calculate propagators with sources at every point in timeslice t of the lattice: such a computation is prohibitively expensive. As a result, we simply ignore the contribution of this last term to the probability amplitude. Physically, this corresponds to using a more complex source whose loop diagrams all cancel; in this case, the source $J = \bar{s}\gamma_5 d - \bar{s}\gamma_5 u$ would achieve that goal.

The denominator of (2.28) poses no problem. Following the same logic as for the numerator, we obtain the result

$$\langle J(\vec{0},t_{\text{snk}})\bar{J}(\vec{0},t_{\text{src}})\rangle = \gamma_{5,\epsilon\eta}(M_H^{-1})^{ea}_{\eta\alpha}(\vec{0},t_{\text{src}};\vec{0},t_{\text{snk}})\gamma_{5,\alpha\beta}(M^{-1})^{ae}_{\beta\epsilon}(\vec{0},t_{\text{snk}};\vec{0},t_{\text{src}}).$$

2.7 Jackknife error estimation

Any lattice QCD calculation that produces a numerical output must produce a corresponding measure of the error inherent to the lattice approximation. There are two main sources of error in lattice calculations: statistical error, due to the Monte Carlo integral over gauge field, and systematic error, due to the discretization of spacetime and finite volume effects. Discretization errors are hard to quantify, but can usually be dealt with by calculating the lattice result various a's and then extrapolating to the continuum. Many finite volume effects, like images, can also be corrected for; for details, see Section 4.2. As for statistical errors, we quantify these using the *jackknife* method [9].

The main idea is the following. Suppose we have some distribution X with p.d.f. $p(x)$, from which we've sampled n values $x_1, \ldots, x_n$. Furthermore, we want to estimate some function of the distribution by using the sampled values. For example, suppose we wanted to calculate the expectation of some function f. The expectation is defined as follows,

$$\langle f \rangle := \int dx\, p(x) f(x),$$

and can be estimated from the sampled values by the formula

$$\langle f \rangle \approx \frac{1}{n} \sum_{i=1}^{n} f(x_n).$$

As $n \to \infty$, the approximation becomes exact. For n finite, we want to estimate the difference between our estimated expectation and the exact expectation.

In *bootstrap method*, we take the sampled values x_i as being the best available approximation of $p(x)$. That is, we assume

$$p(x) := \sum_{i=1}^{n} \frac{1}{n} \delta(x - x_i). \tag{2.30}$$

Then we ask, "if we drew many samples of n values from this new *delta distribution* and estimated $\langle f \rangle$ with them, how much would this estimate for f vary?" More

specifically, let

$$\overline{f} := \frac{1}{n}\sum_{i=1}^{n} f(x_i).$$

Denote by y^k a set of n samples $\{y_1^k, \ldots, y_n^k\}$ drawn from (2.30), with $1 \leq k \leq M$. Essentially, y^k is a selection *with replacement* of n elements from x_i. Let

$$\overline{f}_k := \frac{1}{n}\sum_{i=1}^{n} f(y_i^k).$$

Then the error of our estimate to $\overline{f}$ is given by the standard deviation of the $\overline{f}_k$'s, as follows:

$$\sigma_{\overline{f}} := \lim_{M\to\infty} \sqrt{\sum_{k=1}^{M} \frac{\left(\overline{f}_k - \overline{f}\right)^2}{M}} \qquad \text{(bootstrap estimate)}.$$

The *jackknife method* is an approximation to the bootstrap result which doesn't require taking the limit $M \to \infty$. Instead, we calculate n estimates for $\langle f \rangle$ by omitting each of the x_i's in turn, then use the standard deviation of these estimates to quantify the error on $\overline{f}$. Specifically, let

$$\overline{f}_{(k)} := \frac{1}{n-1}\sum_{\substack{i=1 \\ i\neq k}}^{n} f(x_i),$$

and

$$\overline{f}_{(\cdot)} := \frac{1}{n}\sum_{k=1}^{n} \overline{f}_{(k)}.$$

(Note that $\overline{f}_{(\cdot)} = \overline{f}$ here, but this coincidence doesn't generalize below). Then

$$\sigma_{\overline{f}} := \sqrt{\sum_{k=1}^{n} \frac{n-1}{n}\left(\overline{f}_{(k)} - \overline{f}_{(\cdot)}\right)^2} \qquad \text{(jackknife estimate)}.$$

The method can be generalized to arbitrary functionals of $p(x)$ in the natural way. Let $\theta(x_1, \ldots, x_n)$ be an estimator for some functional of $p(x)$ based on n samples drawn

from p. Generalizing the above definitions, we let

$$\overline{\theta}_{(k)} := \theta(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \quad \text{and} \quad \overline{\theta}_{(\cdot)} := \frac{1}{n}\sum_{k=1}^{n} \overline{\theta}_{(k)}.$$

Then $\overline{\theta}_{(\cdot)}$ is our best estimate of the value of the functional, and the estimate of its error is

$$\sigma_{\overline{\theta}} := \sqrt{\sum_{k=1}^{n} \frac{n-1}{n} \left(\overline{\theta}_{(k)} - \overline{\theta}_{(\cdot)}\right)^2}. \tag{2.31}$$

All error bars in this thesis depict jackknifed statistical errors.

In Section 4.1, we need to fit a curve to a set of data points; the data points have errors, so the fit parameters will have them too. The jackknife method provides a clean, if expensive, way of estimating those errors. Suppose our fit function is characterized by a single parameter θ. Then, to find $\overline{\theta}$, we do a χ^2 fit [20] to the entire data set. Then, we can sequentially fit the function to all the data points not in configuration c, with c ranging through the number of gauge configurations used: this calculates values for $\overline{\theta}_{(c)}$. Finally, Equation (2.31) allows us to calculate an error bar for θ. The procedure generalizes straight forwardly to multiple fit parameters.

2.8 Lattices used in this thesis

All results in this thesis were calculated on a lattice with the following properties:

1. Volume of 16^3, with 32 timeslices
2. Periodic boundary conditions in space, hard wall boundary conditions in time.
3. Gauge coupling β of 6.0. By measuring the string tension [2], it can be determined that the lattice spacing a is then around 0.101 ± 0.002 fm.
4. Total of 90 gauge configurations `OSU_Q60a`, taken from the NERSC [23].

Most propagators were calculated at $\kappa = 0.153$, which corresponds to a pion mass m_π of 821 ± 17 MeV, as shown in Section 4.1.

We wrote our codes using the QDP++ library [8].

Chapter 3

Noise reduction techniques

In an ideal scenario, we could straightforwardly apply the techniques from Chapter 2 to calculate any ground state matrix element. Unfortunately, the gauge integration is done with Monte Carlo techniques, so we can only *estimate* the matrix element to within a certain accuracy. Let N denote the number of gauge fields used in the Monte Carlo integration. As N increases, the error on the estimate decreases as $1/\sqrt{N}$. The computational cost per gauge field is nontrivial, so it's essential to find ways of minimizing the error through methods other than increasing N.

The key idea behind many such *noise-reduction* methods is that we can use many different sources J to measure the same ground state matrix element, given long enough time separations. By picking intelligent J's, we can increase the overlap with the ground state, as in Section 3.1. This technique allows us to measure operators at closer time-separations than otherwise, where stochastic errors are less substantial. In some cases, the form of the matrix element makes it easy to calculate large numbers of related observables, which we can then average over to effectively increase N; these methods are described in Sections 3.2, 3.4 and 3.5.

In general, the matrix elements will be complicated functions of the gauge fields $U_\mu(x)$. When we construct better sources, we end up making J a function of the gauge fields, which introduces stochastic noise. By "smearing" the gauge fields used to construct J, we can reduce the amount of noise introduced. This smearing is described in Section 3.3. Sometimes, however, we need to reduce the noise of the gauge links in

the time direction; for example, when dealing with heavy quarks, the main source of stochastic error comes from fluctuations in the gauge link line linking the source and sink at two timeslices. In this case, we can modify the gauge links locally, averaging gauge links not more than one lattice site away, such that fluctuations are reduced by an order of magnitude, but the calculated matrix element still corresponds to the desired continuum matrix element. This different smearing scheme is described in Section 3.6.

3.1 Extended sources and Wuppertal smearing

The first direct improvement to the sources we can make is to give them some spatial extent. A source such as $J(x) = \overline{u}(x)\gamma^5 d(x)$ is infinitely localised at x, whereas experimentally, a pseudoscalar pion (the ground state of J) has a charge radius of around 0.67 fm [10].

The most general way to give a quark spatial extent whilst maintaining its gauge transformation properties is to make the following replacement:

$$q^a_\alpha(x) \mapsto Q^a_\alpha(x) := \sum_{x'} S^{aa'}_{\alpha\alpha'}(x,x') q^{a'}_{\alpha'}(x').$$

Correspondingly,

$$\bar{q}^b_\beta(y) \mapsto \overline{Q}^b_\beta(y) := \sum_{y'} \bar{q}^{b'}_{\beta'}(y')(\gamma^0 S^\dagger \gamma^0)^{b'b}_{\beta'\beta}(y,y').$$

The matrix S is called the *smearing matrix.* To be able to interpret an extended quark source in the language of states evolving over time, the smearing matrix must be 0 for sites at different time-slices. Furthermore, for $Q^a_\alpha(x)$ and $q^a_\alpha(x)$ to have the same gauge transformation law, the matrix S must have the following gauge transformation law:

$$S^{aa'}_{\alpha\alpha'}(x,x') \xrightarrow{\text{gauge change } g(x)} g^{ab}(x) S^{bb'}_{\alpha\alpha'}(x,x')(g^\dagger)^{b'a'}(x').$$

This is only possible when $S(x,x')$ is a product of spatial link variables forming a chain

from x' to x (perhaps multiplied by some external color matrix that's independent of the U field), where the spin structure is unrestricted.

Consider the following path integral:

$$(F_{QQ})^{ab}_{\alpha\beta}(x,y) := \int [DU\, D\bar{q}\, Dq] e^{-S}\, Q^a_\alpha(x) \overline{Q}^b_\beta(y).$$

Expanding the definitions of Q and $\overline{Q}$, we obtain the following equation:

$$(F_{QQ})^{ab}_{\alpha\beta}(x,y) = \int [DU\, D\bar{q}\, Dq] e^{-S} \sum_{x',y'} S^{aa'}_{\alpha\alpha'}(x,x') q^{a'}_{\alpha'}(x') \bar{q}^{b'}_{\beta'}(y') (\gamma^0 S^\dagger \gamma^0)^{b'b}_{\beta'\beta}(y,y').$$

Splitting the gauge and fermion integrals and actions, and then factoring the S matrices out, we finally learn that

$$(F_{QQ})^{ab}_{\alpha\beta}(x,y) = \int [DU]\, e^{-S_G} \sum_{x',y'} S^{aa'}_{\alpha\alpha'}(x,x') \left(\int [D\bar{q}\, Dq] e^{-S_F}\, q^{a'}_{\alpha'}(x') \bar{q}^{b'}_{\beta'}(y') \right) (\gamma^0 S^\dagger \gamma^0)^{b'b}_{\beta'\beta}(y,y').$$

Recall from Equation (2.21) that, in the quenched approximation, the quantity in parentheses is equal to $(M^{-1})^{a'b'}_{\alpha'\beta'}(x',y')$. Let $F_{qq} := M^{-1}$ to make notation consistent. Further, define $S'(y',y) := \gamma^0 S^\dagger(y,y')\gamma^0$. Suppressing all indices and regarding S and F_{qq} as spacetime matrices, we see that

$$F_{QQ} = S F_{qq} S'.$$

The propagator F_{qq} is called a *point-to-point propagator*, whereas F_{QQ} is called a *smeared-to-smeared propagator.* A similar derivation shows $F_{qQ} := F_{qq}S' = \langle qQ \rangle$, and so it's called a *smeared-to-point propagator.*

Following the derivation above, it's possible to show [25, App. L] a version of Wick's theorem for extended quark sources Q: the only change is that point-to-point propagators F_{qq} are replaced by smeared to smeared propagators F_{QQ}.

In this thesis, we use *Wuppertal* smearing [12]. This scheme depends on two

parameters, α and N. The smearing matrix $S(x,x')$ has no spin structure, and consists of a sum of all spatial gauge link paths of length $l \leq N$ from x' to x, weighted by α^l. There exists a convenient iterative scheme to generate such an S for a fixed x'. Let

$$s^{aa'}_{(0)}(y) = \delta^{aa'}\delta(y-x').$$

Then perform the following iteration N times:

$$s^{aa'}_{(i+1)}(y) = \frac{1}{1+6\alpha}\left[s^{aa'}_{(i)}(y) + \alpha \sum_{\mu=\pm 1}^{\pm 3} (U^{\dagger}_{\mu})^{ab}(y-\hat{\mu}) s^{ba'}_{(i)}(y-\hat{\mu})\right]. \tag{3.1}$$

Then set $S^{aa'}_{\alpha\alpha'}(x,x') = s^{aa'}_{(N)}(x)\,\delta_{\alpha\alpha'}$.

The Wuppertal smearing matrix has an important property, which is obvious from the sum over paths description:

$$S(x,x') = S^{\dagger}(x',x). \tag{3.2}$$

Because it also lacks spin structure, implying that S and γ^0 commute, we conclude that $S' = S$. Thus, the formulas for a smeared propagators simplify to:

$$F_{QQ} = SF_{qq}S \quad \text{and} \quad F_{qQ} = F_{qq}S.$$

In summary, to calculate a smeared-to-smeared propagator at a fixed source point x', we'd calculate the smearing matrix $S(x,x')$ as above, use it as a source vector in (2.22) to obtain $F_{qQ}(x,x')$. Pre-multiplying this propagator by S corresponds to applying the smearing procedure (3.1) on the x coordinate of $F_{qQ}(x,x')$.

Using the hopping parameter expansion, we can visualize a smeared-to-smeared propagator as a chain of gauge links: see Figure 3-1.

The construction outlined above produces a smeared-to-smeared propagator from a fixed point source to any sink point in the lattice. Using (2.26), we can produce an

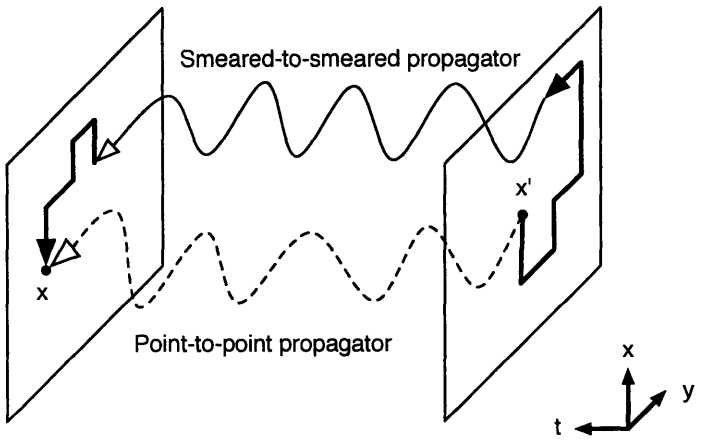


Figure 3-1: Visualizing propagators (z dimension suppressed). The bottom line represents one path in the hopping parameter expansion (2.24). The upper path represents one possible path in a smeared-to-smeared propagator: the heavy lines are pure gauge link chains from S weighted by a power of α, whereas the light line is a hopping path with spin structure, weighted by a power of κ.

analogous fixed-sink propagator:

$$\begin{aligned} F_{QQ}(x',x) &= \sum_{y,z} S(x',y) F_{qq}(y,z) S(z,x), \\ &= \sum_{y,z} S(x',y)\gamma^5 F^\dagger_{qq}(z,y)\gamma^5 S(z,x). \end{aligned}$$

Noting that γ^5 and S commute and using (3.2), we see that

$$\begin{aligned} F_{QQ}(x',x) &= \sum_{y,z} \gamma^5 S(x',y) F^\dagger_{qq}(z,y) S(z,x)\gamma^5, \\ &= \sum_{y,z} \gamma^5 (S^\dagger(z,x) F_{qq}(z,y) S^\dagger(x',y))^\dagger \gamma^5, \\ &= \sum_{y,z} \gamma^5 (S(x,z) F_{qq}(z,y) S(y,x'))^\dagger \gamma^5. \end{aligned}$$

Therefore,

$$F_{QQ}(x',x) = \gamma^5 F^\dagger_{QQ}(x,x')\gamma^5.$$

In other words, Equation (2.26) allows us to reverse the direction of both smeared and unsmeared propagators.

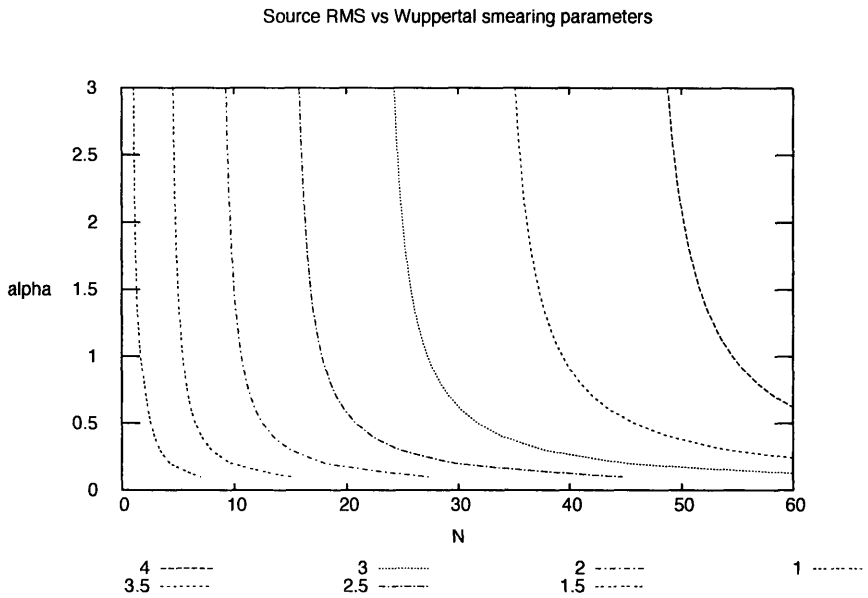


Figure 3-2: RMS (in lattice units) of Wuppertal smeared sources as a function of Wuppertal smearing parameters α and N (see Equation (3.3)).

To conclude, we calculate the spatial extent of smeared source as α and N are varied. The calculations were done on the lattices described in Section 2.8. Regarding the smearing matrix from x' to r as a kind color wavefunction of an extended quark at x', it's reasonable to make the following definition:

$$|\psi(\vec{r})|^2 := \mathrm{Tr}[S^\dagger(r,x')S(r,x')],$$

where the time components of t and t' are equal but otherwise irrelevant. We then define the RMS of an extended source as follows:

$$R(\alpha,N) := \left[\frac{\sum_{\vec{r}} r^2|\psi(\vec{r})|^2}{\sum_{\vec{r}}|\psi(\vec{r})|^2}\right]^{1/2}. \tag{3.3}$$

The RMS has the virtue of having a physical interpretation, whereas smearing parameters like α and N do not.

We calculate R for each gauge configuration to obtain a mean $\overline{R}$ with associated error. When $U = 1$, smearing corresponds to exploring all walks of N steps or less

from a given lattice point, so we'd expect $|\psi(\vec{r})|^2$ to be roughly Gaussian, and R to grow as $\sqrt{N}$. The gauge links act like complex rotations whose effects can partially cancel; thus we expect the same qualitative behavior, but with lower absolute RMS values.

Figure 3-2 shows a contour plot of source RMS as α and N are varied. This establishes that R is a strong function of N, but not of α for $\alpha > 1$. We decided to use $\alpha = 3.0$ throughout the rest of this work. Figure 3-3 plots source RMS as a function of N at this value of α. The square root scaling behavior is evident.

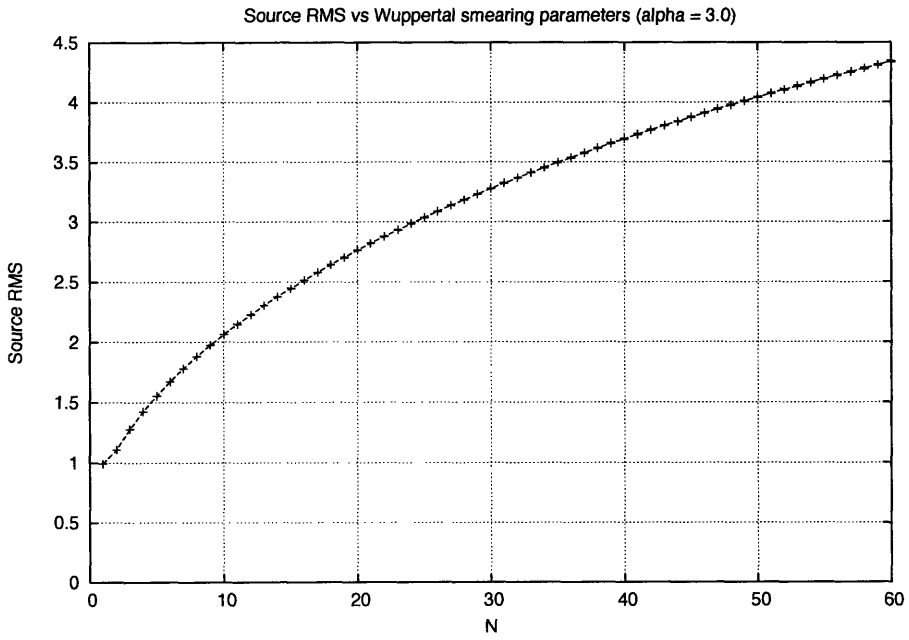


Figure 3-3: RMS (in lattice units) of Wuppertal smeared sources as a function of smearing steps N, with $\alpha = 3.0$ fixed (see Equation (3.3)).

3.2 Momentum projection for quark masses

The purpose of using extended sources is to improve the overlap between J and its ground state. When calculating the mass of this ground state, there is an extra trick that allows us to effectively increase N by orders of magnitude while simultaneously increasing the energy gap between the ground state and the first excited state. Both of these effects allow us to measure bound state masses with few gauge configurations

at high accuracy. The trick is to analytically project the sink onto the zero momentum subspace of states.

As discussed in Section 2.3, we can find the ground state energy of J's ground state by calculating the following quantity:

$$C(\vec{x}, t; \vec{x}', t') := \langle J(\vec{x}, t)\overline{J}(\vec{x}', t')\rangle.$$

As $|t - t'| \to \infty$, we get $C(\vec{x}, t; \vec{x}', t') \to A\exp(-(E_J - E_0)|t - t'|)$, where E_0 is the energy of the vacuum (we can redefine E_0 away by adding a constant shift in energy to the action). By tracking the slope of $\log C(t - t')$, we can extract E_J, which is the ground state energy/mass.

As in the continuum theory, we can Fourier transform to define states of definite momentum:

$$J(\vec{p}, t) := \sum_{\vec{x}} e^{i\vec{p}\cdot\vec{x}} J(\vec{x}, t).$$

If we use a definite momentum sink and insert a complete set of states between it and the source, we can see that only the states of the same momentum in the source survive. In other words, projecting one of the source or sink automatically projects the other one.

Suppose we project the sink onto zero momentum. Fixing $\vec{x}'$ and t', we can define

$$C_0(t - t') := \sum_{\vec{x}} C(\vec{x}, t; \vec{x}', t').$$

This quantity converges faster towards the ground state: presumably, the first zero-momentum excited state of J has a higher energy than the first non-zero momentum excitation of the ground state, so we've increased the relevant energy gap dictating converge rate. Moreover, the sum over all lattice sites $\vec{x}$ implies that C_0 averages many more gauge link paths that does $C(\vec{x}, t; \vec{x}', t)$ alone; effectively, this is equivalent to increasing N by a factor of about L^3, with the corresponding improvement in statistics.

These techniques are used in Section 4.1 to determine the masses of the π and

the ρ at various κ values.

3.3 APE smearing of gauge fields

When smearing a source, the gauge links used introduce a certain amount of noise. We can reduce this noise by *smearing the gauge field*: that is, replacing each gauge link with an average of many nearby gauge paths. As long as the averaging is done in a way that maintains the gauge transformation properties of each gauge link, using this smeared gauge field to generate an extended quark source is just as valid as using the original one, except stochastic fluctuations should be substantially reduced.

APE smearing [11] is a particularly simple iterative scheme of smearing gauge fields. We define the *staple* of a link $U_\mu^\dagger(x)$ in a direction $\hat{\nu}$ as the gauge path that begins at x, goes in the direction $\hat{\nu}$, then along $\hat{\mu}$, then along the $-\hat{\nu}$ direction again. By construction, a staple has the same gauge transformation law as the original link. Denote the stable defined above by $T_{\mu\nu}^\dagger(x)$. It's given by the following formula:

$$T_{\mu\nu}^\dagger(x) := U_{-\nu}^\dagger(x+\hat{\nu}+\hat{\mu})U_\mu^\dagger(x+\hat{\nu})U_\nu^\dagger(x).$$

One iteration of APE smearing replaces each link of the gauge field with a weighted average of itself and its staples. We use a variant of APE smearing in which only spatial links are averaged, and only spatial staples are considered; though not strictly necessary, this restriction avoids mixing links involved in time evolution with those used in source smearing. One iteration of this *spatial APE smearing* performs the following replacement:

$$U_\mu^\dagger(x) \mapsto \mathcal{P}\left[(1-c)U_\mu^\dagger(x) + c\sum_{\substack{\nu=\pm 1 \\ \nu\neq\mu}}^{\pm 3} T_{\mu\nu}^\dagger(x)\right] \quad (\mu \neq 4). \tag{3.4}$$

The parameter c controls the weight of the averaging. Furthermore, since SU(3) is not a vector space, we generally have to project the averaged gauge link back into

SU(3) using the $\mathcal{P}$ operator[1] which we now describe.

Projecting an arbitrary 3×3 matrix V onto SU(3) is done by picking the $U \in \mathrm{SU}(3)$ that minimizes $\|U - V\|^2$ under some norm. The following matrix norm is widely used:

$$\|M\|^2 := \mathrm{Tr}[M^\dagger M]. \tag{3.5}$$

Expressing the above formula in terms of the entries of M, we see

$$\|M\|^2 = \sum_{a,b} M^*_{ab} M_{ab} = \sum_{ab} |M_{ab}|^2.$$

Thus the norm (3.5) is the natural generalization of the standard vector norm to matrices.

We can re-express this minimization more usefully by expanding $\|U - V\|^2$ as follows:

$$\begin{aligned}
\|U - V\|^2 &= \mathrm{Tr}[(U - V)^\dagger (U - V)], \\
&= \mathrm{Tr}[U^\dagger U - U^\dagger V - V^\dagger U + V^\dagger V], \\
&= \mathrm{Tr}[\hat{\mathbb{I}}] - 2\,\mathrm{Re}\,\mathrm{Tr}[U^\dagger V] + \mathrm{Tr}[V^\dagger V].
\end{aligned}$$

The last line follows from U being unitary and $\mathrm{Tr}[M^\dagger] = \mathrm{Tr}[M]^*$ for any matrix M. Since V is fixed during the minimization, projecting V into SU(3) is equivalent to picking a $U \in \mathrm{SU}(3)$ that maximizes $\mathrm{Re}\,\mathrm{Tr}[U^\dagger V]$. Cabibbo and Marinari [6] have found an efficient algorithm, described in Section 3.3.1, to do this maximization.

We now prove an essential property of the $\mathcal{P}$ operator. Let $g \in \mathrm{SU}(3)$. Then the following two equalities hold:

$$\mathcal{P}[gV] = g\mathcal{P}[V]; \tag{3.6a}$$

$$\mathcal{P}[Vg] = \mathcal{P}[V]g. \tag{3.6b}$$

[1] We use the same symbol for the SU(3) projection operator and the path ordering operator; the context usually makes the implied operation clear.

Assume the $U \in \mathrm{SU}(3)$ that maximizes $\mathrm{Re}\,\mathrm{Tr}[U^\dagger V]$ is unique[2]. Set $U' := \mathcal{P}[gV]$, so that U' maximizes $\mathrm{Re}\,\mathrm{Tr}[(U')^\dagger gV] = \mathrm{Re}\,\mathrm{Tr}[(g^\dagger U')^\dagger V]$. By uniqueness of U, we conclude $U = g^\dagger U'$. In other words, we find $U' = gU$, which proves the first statement. The proof of the second statement is analogous.

Equations (3.6a) and (3.6b) guarantee that both V and $\mathcal{P}[V]$ have the same gauge transformation laws. Thus, APE-smeared gauge fields can be used in Wuppertal smearing.

The order in which the replacements in (3.4) is done is important. Two natural choices emerge: (a) perform all the replacements simultaneously; or (b) perform the replacement at every lattice site sequentially. The difference is that in (b), Iteration i at site makes use of the partial results of Iteration i and those of Iteration $i-1$, whereas in (a), Iteration i only makes use of the results of Iteration $i-1$. Empirically, D. Sigaev has found[3] that in scheme (a), the Cabibbo-Marinari algorithm fails to converge for $c \gtrsim 1/3$, whereas in scheme (b), this doesn't happen. For historical reasons, we've used scheme (a) for this thesis.

In Figure 3-4, we calculate the effect of APE smearing with $c = 1/3$ on the source RMS, for various numbers of APE smearing iterations. As expected, APE smearing makes the gauge fields fluctuate less, so there are fewer cancellations along any gauge path, which slightly increases the RMS of the extended quark source.

In Figure 3-5, we show the effect of APE smearing on stochastic fluctuations. To do this, we calculate for each smeared quark source the following quantity:

$$S := \sum_{\vec{r}} |\psi(\vec{r})|^2.$$

Denote by $\overline{S}$ the mean value of S across all gauge fields used, and by σ_S its error. We've plotted $\sigma_S/\overline{S}$ for various sources, under various APE smearing conditions.

[2]This need not be true in certain degenerate cases, for example $V = 0$. However, we may expect it to be true for "typical" V's generated by the APE prescription

[3]Personal communication.

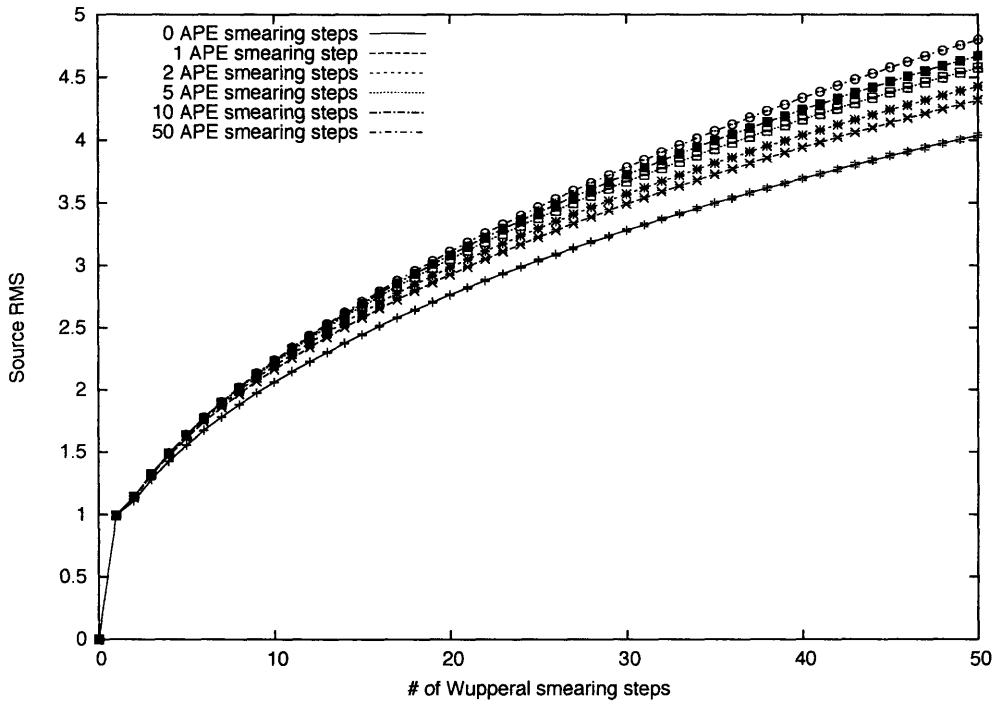


Figure 3-4: Wuppertal-smeared ($\alpha = 3.0$) source RMS when gauge field is APE smeared at $c = 1/3$. The lower curve corresponds to no APE smearing, with subsequently higher curves corresponding to 1–50 APE smearing steps.

3.3.1 The Cabibbo-Marinari algorithm

In their paper [6], Cabibbo and Marinari provide a recipe for generating random elements of SU(N) with the following Boltzmann distribution:

$$p(U) \sim \exp(-\beta \operatorname{Re} \operatorname{Tr}[V^\dagger U]), \tag{3.7}$$

where V is a fixed, arbitrary $N \times N$ matrix, when we know how to do this analytically only for SU(2). By looking at the standard Wilson gauge action (2.13), we see this is a crucial step in generating Boltzmann-distributed gauge fields $U_\mu(x)$ for use in the gauge Monte Carlo integral; in that case, V is the sum of all the staples of a particular

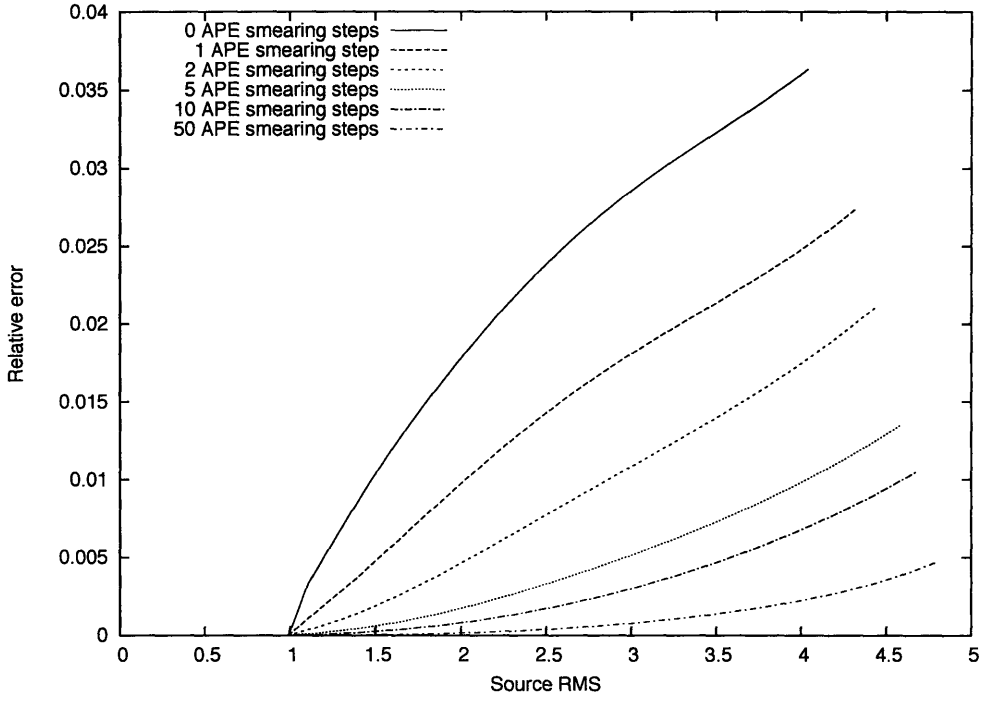


Figure 3-5: Relative error $\sigma_S/\overline{S}$ (see text) in source as a function of source RMS and number of APE smearing steps ($\alpha = 3.0$, $c = 1/3$).

gauge link. When $\beta \to \infty$, the probability of U maximizing $\mathrm{Re\,Tr}[V^\dagger U]$ tends to 1. Their algorithm thus provides a concrete way to project any matrix onto SU(N).

The main idea is to define a set $F := \{\mathrm{SU}(2)_1, \ldots, \mathrm{SU}(2)_m\}$ of m SU(2) subgroups of SU(N). Denote by α some element of one of the subgroups of F. We require the set F to be large enough that no element of SU(N) is invariant under left multiplication by some α. For $N = 3$, the following three subgroups suffice:

$$\mathrm{SU}(2)_1 := \begin{pmatrix} a_1 & a_2 & 0 \\ a_3 & a_4 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathrm{SU}(2)_2 := \begin{pmatrix} a_1 & 0 & a_2 \\ 0 & 1 & 0 \\ a_3 & 0 & a_4 \end{pmatrix}, \quad \text{and} \quad \mathrm{SU}(2)_3 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & a_1 & a_2 \\ 0 & a_3 & a_4 \end{pmatrix},$$

where

$$\begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \in \mathrm{SU}(2).$$

Let $\alpha_i \in \mathrm{SU}(2)_i$. Further, let $U^{(0)} := U$ and

$$U^{(k)} := \alpha_k \alpha_{k-1} \cdots \alpha_1 U.$$

Suppose we chose α_1, then α_2, ..., then α_m, such that α_i is distributed according to

$$p(\alpha_i) \sim \exp(-\beta \operatorname{Re} \operatorname{Tr}[V^\dagger \alpha_i U^{(i-1)}]) = \exp(-\beta \operatorname{Re} \operatorname{Tr}[(V(U^{(i-1)})^\dagger)^\dagger \alpha_i]). \tag{3.8}$$

Cabibbo and Marinari's main result is that if $U \in \mathrm{SU}(N)$ is distributed according to (3.7), and the α_i are picked as above, then

$$U' := U^{(m)} = \alpha_m \cdots \alpha_1 U \tag{3.9}$$

is also distributed according to (3.7). Thus, starting with any U (say $U = 1$), we can generate a sequence of Boltzmann distributed elements of $\mathrm{SU}(N)$.

We're interested in the $\beta \to \infty$, in which choosing α_i according to (3.8) is equivalent to minimizing $\operatorname{Re} \operatorname{Tr}[M^\dagger \alpha_i]$. This operation we can do analytically. For simplicity, we take $M, \alpha \in \mathrm{SU}(2)$, though the derivation extends trivially to $\alpha_i \in \mathrm{SU}(2)_i$.

Any $\alpha \in \mathrm{SU}(2)$ can be written as the following linear combination:

$$\alpha := \beta \hat{\mathbb{I}} + \vec{\beta} \cdot \vec{\sigma},$$

where $\beta \in \mathbb{R}$ and $\vec{\beta} \in \mathbb{R}^3$, and, α being unitary, satisfy

$$\beta^2 + |\vec{\beta}|^2 = 1. \tag{3.10}$$

The vector $\vec{\sigma}$ consists of the three Pauli matrices. Recall that these matrices are traceless and Hermitian, and that $\operatorname{Tr}[\sigma_i \sigma_j] = 2\delta_{ij}$. Furthermore, by using complex

prefactors, we can represent *any* 2×2 matrix as follows:

$$M := N\,\hat{\mathbb{I}} + \vec{N} \cdot \vec{\sigma},$$

where

$$N = \tfrac{1}{2}\,\mathrm{Tr}[M] \quad \text{and} \quad \vec{N} = \tfrac{1}{2}\,\mathrm{Tr}[M\vec{\sigma}].$$

Within this setup, we can see that

$$\mathrm{Re}\,\mathrm{Tr}[M^\dagger \alpha] = \mathrm{Re}[2N^*\beta + 2\vec{N}^* \cdot \vec{\beta}].$$

The right hand side is maximized when $(\beta, \vec{\beta}) = k\,\mathrm{Re}(N, \vec{N})$, with k chosen to satisfy (3.10).

The result derived in the previous two paragraphs allows us to pick α_i matrices distributed according to (3.8) as $\beta \to \infty$. By applying the Cabibbo-Marinari result (3.9), we can maximize $\mathrm{Re}\,\mathrm{Tr}[U^\dagger V]$ when $U \in \mathrm{SU}(N)$, in particular when $N = 3$.

3.4 Averaging multiple timeslices

Another straightforward statistics improvement technique we can implement is to average the values of some operator over as many timeslices as possible, rather than just considering it at one timeslice. The limiting factor is the speed at which the ground state is filtered out by J. For instance, consider the following matrix element:

$$\rho(\vec{x}, t) = \frac{\langle J(\vec{0}, t_{\mathrm{snk}})\rho(\vec{x}, t)\overline{J}(\vec{0}, t_{\mathrm{src}})\rangle}{\langle J(\vec{0}, t_{\mathrm{snk}})\overline{J}(\vec{0}, t_{\mathrm{src}})\rangle}. \tag{3.11}$$

For $t_{\mathrm{src}} \ll t \ll t_{\mathrm{snk}}$, we expect $\rho(\vec{x}, t)$ to be independent of t. Thus, averaging over many such t is equivalent to multiplying the number of gauge fields by some small factor.

In Section 4.2, we evaluate matrix elements with $t_{\mathrm{src}} = 11$ and $t_{\mathrm{snk}} = 20$. After showing reasonable convergence of the ground state at intermediate timeslices, we average the calculated matrix element over $t = 15$ and $t = 16$, doubling our statistics.

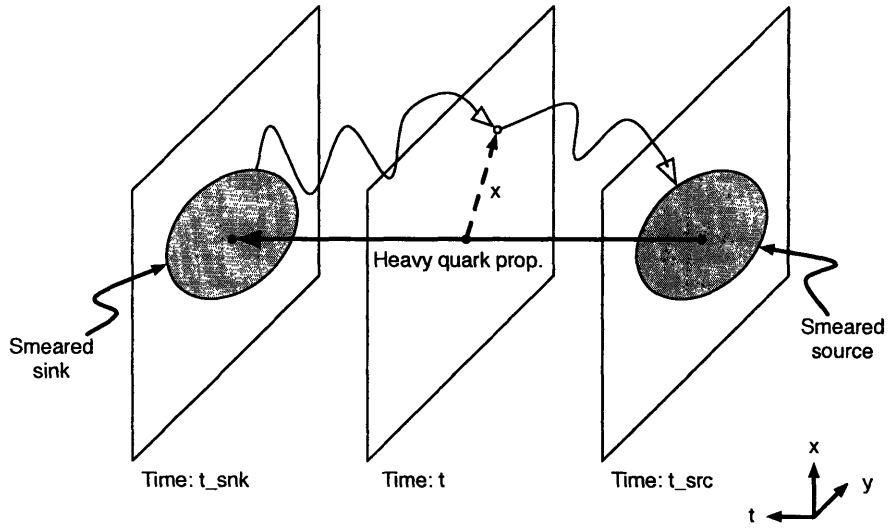


Figure 3-6: Schematic calculation of the density expectation at $\vec{x}$ of the light quark in a heavy-light meson, measured from the position of the heavy quark.

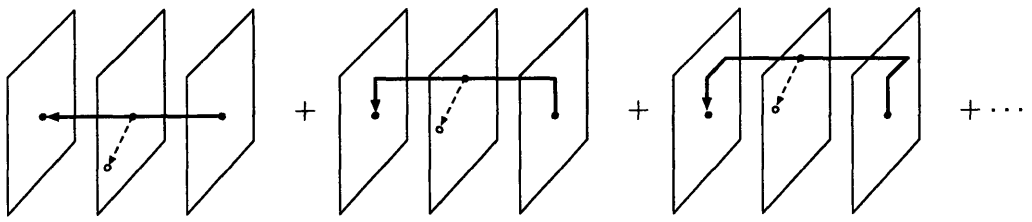
3.5 Multiple heavy lines for heavy quark matrix elements

When calculating matrix elements of a source with heavy quarks, stochastic errors are grossly larger than when dealing with light quarks alone. The origin of this difference can be seen as follows. Figure 3-6 visualizes the propagation of quarks that the numerator of (3.11) gives rise to when J is a heavy-light meson. The heavy propagator, shown as a thick black line, is a chain of gauge links (see Equation (2.25)). The fluctuations in this one path completely dominate the error in the matrix element.

The solution to the heavy-line problem is to average over many "equivalent" heavy lines. We've explored two ways of implementing this solution: averaging over several displaced heavy line and smearing the links that make up the heavy line. Here we discuss the former method; Section 3.6 is devoted to the latter.

The choice of position $\vec{h}$ of the heavy line on the lattice is completely arbitrary, as long as the position $\vec{x}$ of the light-quark density operator is measured from $\vec{h}$. With this insight, we may displace the heavy line a few lattice units away from $\vec{0}$ and calculate a new value of $\rho(\vec{r})$, effectively increasing the number of lattices in the

Monte Carlo gauge integration by about one order of magnitude. We need a way to connect the displaced heavy line to the origin of the light quark propagator to make the entire matrix element gauge invariant: the Wuppertal smearing matrix $S(\vec{h}, \vec{0})$, with a low number of Wuppertal smears, provides one such connection. Schematically, the averaging looks as follows:



The dashed arrow represents $\vec{x}$. Each term is a copy of (3.11), not just its numerator. Notice that, at the measurement timeslice, all the terms in the above sum look like shifts of each other; thus, by translational invariance, they should be equal.

An alternative way to understand these heavy line shifts is to regard the heavy-line as fixed and the above sum as averaging many copies (3.11) with displaced light quarks. This perspective highlights an important fact: displacing the heavy line by too much distance is unhelpful, because the overlap of the sources with the heavy-light meson ground state becomes negligible.

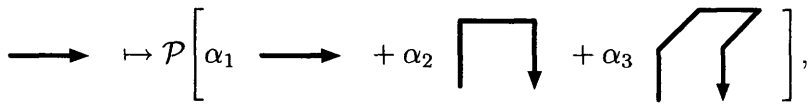
In Section 4.2, we compare the results of computing a matrix element with a single heavy line and with many displaced heavy lines.

3.6 HYP smearing of heavy quark lines

Averaging over heavy lines as described above is a clever way of averaging many equivalent observables together to lower statistical fluctuations. Nevertheless, it's a faithful rendition of a heavy-light meson: the heavy quark produces a single line of gauge links. A complementary approach, used by [1], is to apply so-called *HYP smearing* to the time links. This scheme replaces each time link with a weighted average of paths close to the link, resulting in an order of magnitude reduction in

errors. However, the heavy quark interpretation of the calculation is then only valid in the $a \to 0$ limit.

HYP smearing, shorthand for hypercubic smearing, was first introduced in [13]. The prescription to follow is given symbolically below:



where $\mathcal{P}$ is the SU3 projector defined in Section 3.3. The idea is similar in spirit to APE smearing, but using only links less than 2 lattice units away to smear. Thus, the smoothing is much more localized. More formally, we have

$$U_\mu^\dagger(x) \mapsto \mathcal{P}\Big[\alpha_1 U_\mu^\dagger(x) + \alpha_2 \sum_{\nu\neq\mu} U_\nu(x+\hat{\mu})U_\mu^\dagger(x+\hat{\nu})U_\nu^\dagger(x) + \alpha_3 \sum_{\substack{\nu\neq\mu \\ \eta\neq\mu,\nu}} U_\nu(x+\hat{\mu})U_\eta(x+\hat{\mu}+\hat{\nu})U_\mu^\dagger(x+\hat{\nu}+\hat{\eta})U_\eta^\dagger(x+\hat{\nu})U_\nu^\dagger(x)\Big].$$

The parameters α_i are chosen to minimize fluctuations, in some sense. We use the values $\alpha_1 = 0.75$, $\alpha_2 = 0.60$ and $\alpha_3 = 0.30$; these were found in [13] to minimize fluctuations in the average plaquette.

Chapter 4

Measurements

In this chapter we summarize a number of lattice measurements. First, to confirm the implementation of the methods described in Chapters 2 and 3, we reproduce published results on the masses of the π and the ρ, in the a regime of heavy up and down quarks. We then proceed to measure the density correlator in a heavy-light meson system; we show a progression of results as various smoothing techniques and corrections are applied, culminating with an especially clean measurement of the density operator shown in Figure 4-11. Finally, we extend our measurements to the case of a density-density correlator in a heavy-heavy-light baryon. This last result is compared to the recent literature, and then its possible interpretation is discussed.

Throughout, we use the lattices described in Section 2.8.

4.1 Lattice Mass Measurements

To measure the mass of a source J, we calculate the following correlation function:

$$C(t) := \sum_{\vec{y}} \langle J(\vec{y}, t) \overline{J}(\vec{x}, t_{\text{src}}) \rangle.$$

Here, $\vec{x}$ and t_{src} are fixed. As explained in Section 3.2, the sum projects J onto zero momentum, increasing the energy gap between the ground and first excited state, as well as massively improving the statistics of the measurement by averaging over

many more gauge paths than if $\vec{y}$ were fixed.

In the region where the ground state has been filtered out, we expect

$$C(t) = A \exp[-(E_J - E_0)(t - t_{\text{src}})]. \tag{4.1}$$

By shifting E_0 to 0, we can eliminate any reference to it. We can then define the following quantity, called an *effective mass*:

$$m_{\text{eff}}(t) = \log \frac{C(t+1)}{C(t)}.$$

In the time regime where (4.1) holds, a graph of m_{eff} vs. t looks flat; we say it "plateaus." By fitting the measured m_{eff} to a flat line in this region, we can extract E_J. Outside this region, on the other hand, the speed at which m_{eff} tends to a flat line indicates the size of the energy gap between the ground and first excited states.

We calculated E_J for a two sources: a pseudoscalar π, with $J = \overline{q}\gamma^5 q$; and a ρ, which has $J = \overline{q}\gamma^\mu q$ (all four μ's are equivalent, we've used $\mu = 1$). All quarks had the same mass: we set κ to 0.135, 0.153, 0.155 and 0.1558. As shown below, these κ values correspond to pion masses of 2.37 GeV, 821 MeV, 577 MeV and 454 MeV, respectively; the physical pion mass is 135 MeV [10]. Choosing these particular values for κ allowed us to numerically compare our results with [3]. Our sources were Wuppertal smeared, with $\alpha = 2.5$ and $N = 40$, and were placed at $(\vec{x}, t_{\text{src}}) = (\vec{0}, 6)$. This increases the amount of time during which (4.1) holds while decreasing the effect of images induced by the hard-wall boundary in the time dimension.

In Figures 4-1 and 4-2, we show a few representative plots of m_{eff} vs. t. The onset of plateaus is evident. Also shown are the fits to straight lines used to extract source masses. We chose to fit $m_{\text{eff}}(t)$ to a straight line in the range of 11–24. Empirically, this looked separated enough away from $t_{\text{src}} = 6$ for the ground state to have filtered out, and far enough away from the hard wall at $t = 32$ that image effects are negligible.

Table 4.1 shows our results for the mass measurements, compared to the literature results from [3]. In particular, taking $a = 0.101 \pm 0.002$fm as in [2], we see that at $\kappa = 0.153$, the pion has a mass of 821 ± 17 MeV.

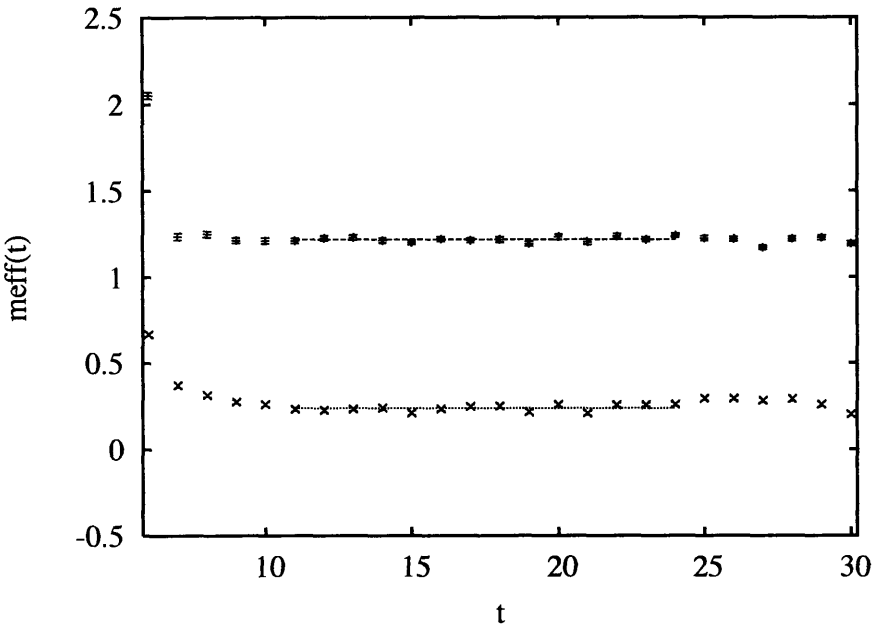


Figure 4-1: Effective mass plots for π at $\kappa = 0.135$ (top) and $\kappa = 0.1558$ (bottom). Best fits to flat lines in range $t = 11$–24 are also shown.

4.2 Density Correlators in Heavy-Light Mesons

One of the primary objectives of this thesis was to calculate the following observable for a heavy-light meson source:

$$\rho(\vec{x}) := \frac{\langle J(\vec{0}, t_{\text{snk}})\rho(\vec{x}, t)\overline{J}(\vec{0}, t_{\text{src}})\rangle}{\langle J(\vec{0}, t_{\text{snk}})\overline{J}(\vec{0}, t_{\text{src}})\rangle}. \tag{4.2}$$

	π		ρ	
κ	Thesis	Gupta et al.	Thesis	Gupta et al.
0.135	1.2175 ± 0.0025	1.217 ± 0.001	1.2294 ± 0.0027	1.229 ± 0.001
0.153	0.4245 ± 0.0031	0.421 ± 0.000	0.5075 ± 0.0059	0.504 ± 0.001
0.155	0.2989 ± 0.0043	0.296 ± 0.001	0.426 ± 0.010	0.422 ± 0.002
0.1558	0.2377 ± 0.0051	0.233 ± 0.001	0.396 ± 0.014	0.386 ± 0.003

Table 4.1: Mass measurements of π and ρ at various κ. All results are in lattice units ($1a = 0.101\text{fm} = 1.95\text{GeV}$).

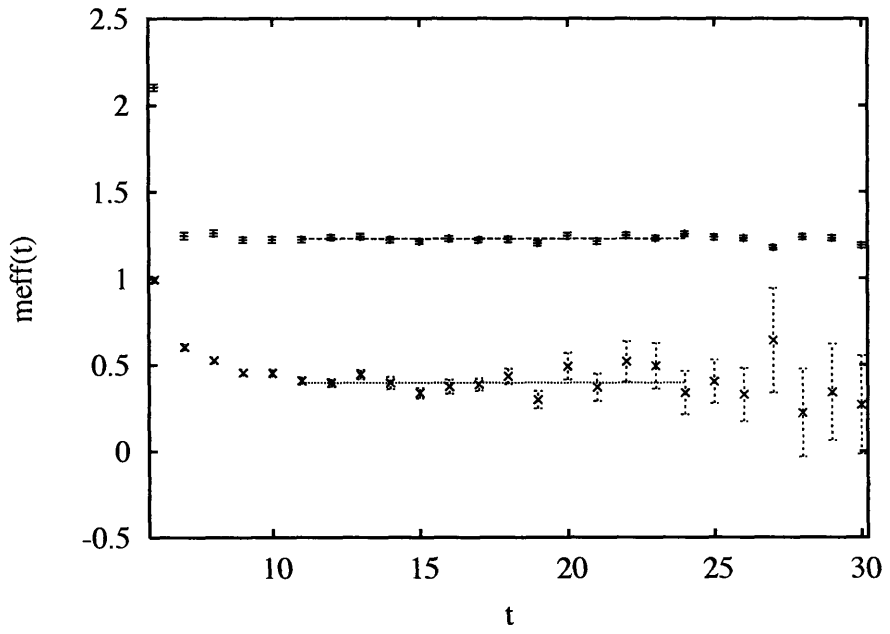


Figure 4-2: Effective mass plots for ρ at $\kappa = 0.135$ (top) and $\kappa = 0.1558$ (bottom). Best fits to flat lines in range $t = 11$–24 are also shown.

The details of the calculation are described in Section 2.6. For the light quark, we use $\kappa = 0.153$. Since we need to calculate new sets of propagators for different values of t_{src} and t_{snk}, we've arbitrarily fixed $t_{\text{src}} = 11$.

There is considerable latitude in choosing the arbitrary parameters of the above matrix element:

1. Where to place t with respect to t_{src} and t_{snk}: so as to not priviledge either source or sink, we want $t \approx \frac{1}{2}(t_{\text{src}} + t_{\text{snk}})$. When this average is not an integer, there are two symmetric t values: by averaging ρ on these two middle timeslices, we can double our statistics without inducing large systematic errors.

2. How far away to place t_{snk} from t_{src}: Ideally, we'd place them as far away as possible while still staying away from the time walls of the lattice (to avoid image effects). In practice, the error bars on the measured ρ's explode as $t_{\text{snk}} - t_{\text{src}}$ grows. Thus, we need a compromise: we want $t_{\text{snk}} - t_{\text{src}}$ as small as possible, while still ensuring that the ground state has been filtered out by time t.

3. How to smear the sources: The purpose is two-fold. First, we want to maximize the overlap between J and the ground state; that way we can keep $t_{\text{snk}} - t_{\text{src}}$ low. Secondly, by smearing we are sampling over many similar lattice paths; up to a certain point, this sampling will reduce error bars on sample points.

Figure 4-3 shows the result of a naïve evaluation of (4.2). For this case, we've used $t_{\text{snk}} = 18$. The light quarks were smeared with Wuppertal smearing at $\alpha = 3.0$ for $N = 8$ times with an APE-smeared gauge field ($c = 1/3$, $N = 10$), which corresponds to a source RMS of 2.0. Each point on the graph shows a particular value of $\rho(\vec{x})$, with $r := |\vec{x}|$, at $t = 14$. The error bars have been suppressed for clarity: they are of the same general size as the spread in the data values.

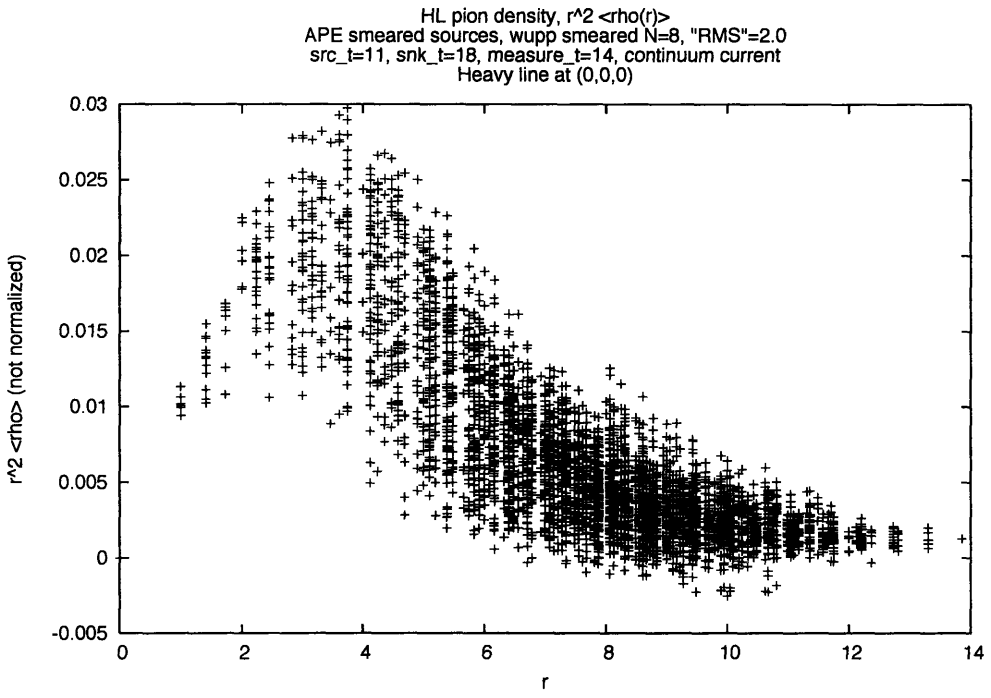


Figure 4-3: Naïve evaluation of density correlator; see text.

With this calculation complete, we then aimed at using the techniques described in Chapter 3 to reduce the errors on the estimate of ρ as much as possible. Figure 4-4 shows the same calculation after a step in the direction of reducing errors. In this

figure, we average the data points about the two central timeslices $t = 14$ and $t = 15$, as described in Section 3.4.

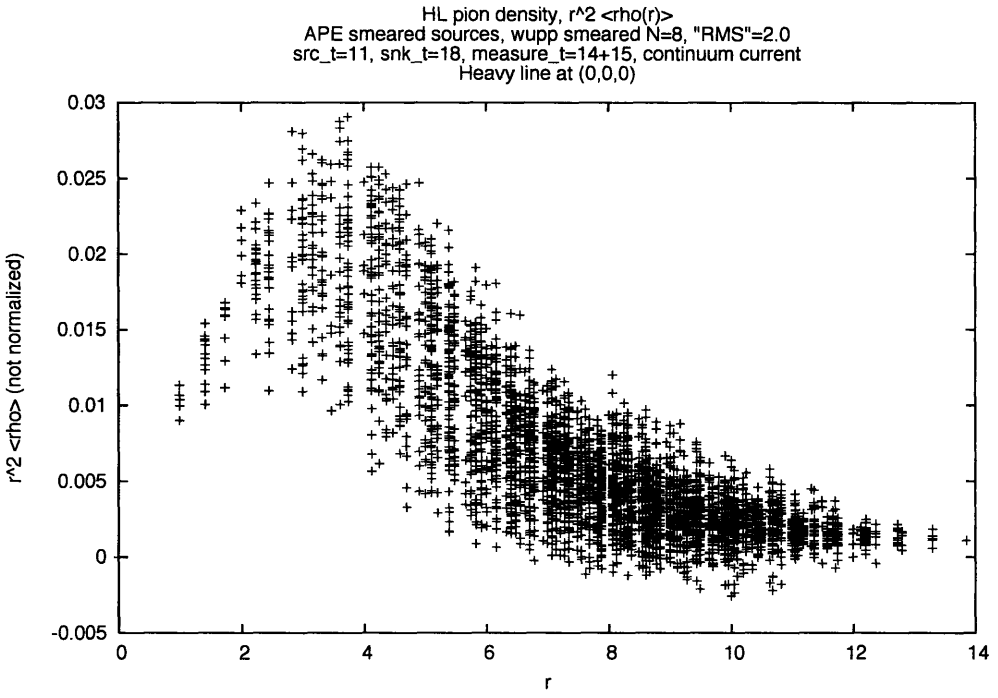


Figure 4-4: Density correlator, averaged over two central timeslices; see text.

The next obvious step is to average over many heavy lines, as described in Section 3.5. Figure 4-5 shows the same plot as Figure 4-4, but averaged over 7 heavy lines: $(0,0,0)$ and all lines exactly 1 unit away from the origin. The improvement is evident. We start seeing convergence towards a single curve. Strong image effects that were previously masked by statistical noise are now becoming evident, as seen by the peak at $r = 8$ (halfway between the source and its nearest periodic image).

A question that arises naturally at this point is whether or not $t_{\mathrm{snk}} = 18$ is high enough for the ground state to have filtered out. If the ground state were filtered out by $t = 14$, the shape of the ρ should be relatively independent of the source used. We evaluated ρ for three sources of varying sizes: they were Wuppertal smeared 8, 19 and 35 times, to make their RMSs 2.0, 3.0 and 4.0 lattice units, respectively.

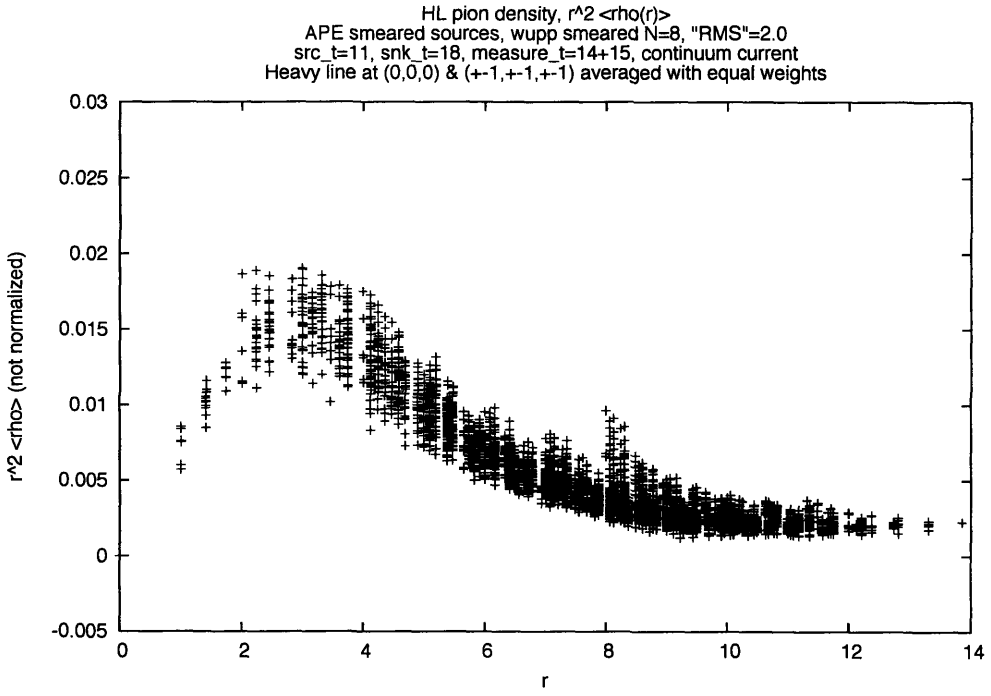
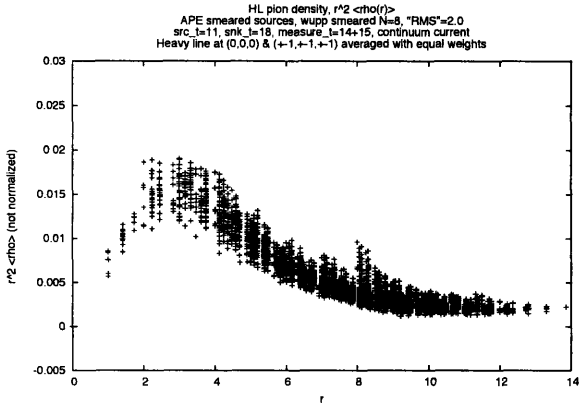


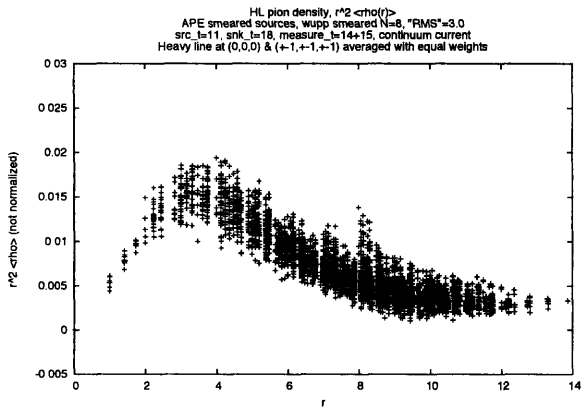
Figure 4-5: Averaging over many heavy lines.

The resulting plot sequence is shown in Figure 4-6; it is evident the curves are *not* independent of the source shape.

(a)



(b)



(c)

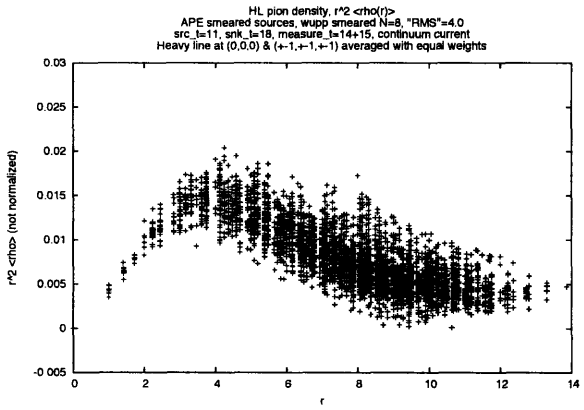


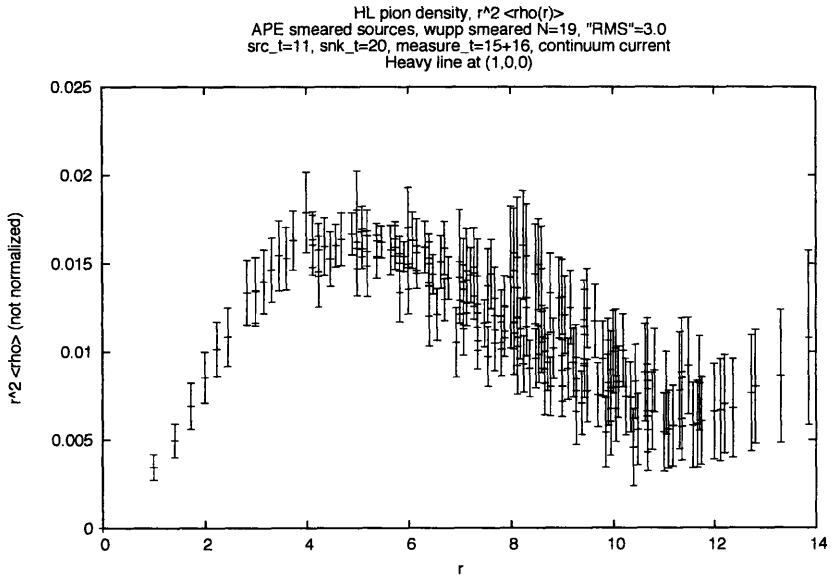
Figure 4-6: A value for $t_{\mathrm{snk}} - t_{\mathrm{src}}$ of 7 is not high enough: the calculated matrix elements depend on the shape of the source, not just its quantum numbers. The sources used have different RMS (Wuppertal smearing iterations): (a) 2.0 (8); (b) 3.0 (19); and (c) 4.0 (35).

Though not shown explicitly here, we explored using $t_{\text{snk}} = 20$ and $t_{\text{snk}} = 22$. At the larger value, error bars started becoming unmanageable, whereas those for $t_{\text{snk}} = 20$ weren't much worse than in the $t_{\text{snk}} = 18$ case. Thus, we decided to select $t_{\text{snk}} = 20$ and continue with the exploration of observables; that this source-separation distance suffices will be shown at the end of the section. From now on, the two timeslices that are averaged are $t = 15$ and $t = 16$. Because we also want a source that matches the actual meson ground state as closely as possible, we decided to use the $N = 19$ source, with an RMS of 3.0 (an RMS of 4.0 would have been more appropriate, but the increased error bars of that source far outweighed the benefit of better source/ground-state overlap).

We can improve our statistics for free by averaging $\rho(\vec{x})$ over cubically equivalent lattice sites. For example, because of cubical symmetry, we know $\rho(0,0,1)$, $\rho(0,1,0)$ and $\rho(1,0,0)$ are relatively independent statistical samples of the same mean value. An arbitrary point will in general be in a family of $8 \times 6 = 48$ cubically equivalent points. We call these *sibling points* of the original point. The siblings of a point correspond to all straight angle rotations and axis reflections that preserve a cubic lattice; for some special points, some sibling points are equal. By averaging over cubically equivalent points, we can increase the effective number of lattices used by a factor of 48. We perform this average from now on.

There is one more complementary technique that we can use to lower errors: HYP smearing of the heavy quark links. As described in Section 3.6, the calculation with smeared heavy quark links becomes equivalent to the calculation with unsmeared links in the continuum limit. We applied HYP smearing without averaging over many heavy lines at first; the dramatic contrast between HYP smearing and no HYP smearing is evidenced in Figure 4-7.

(a)



(b)

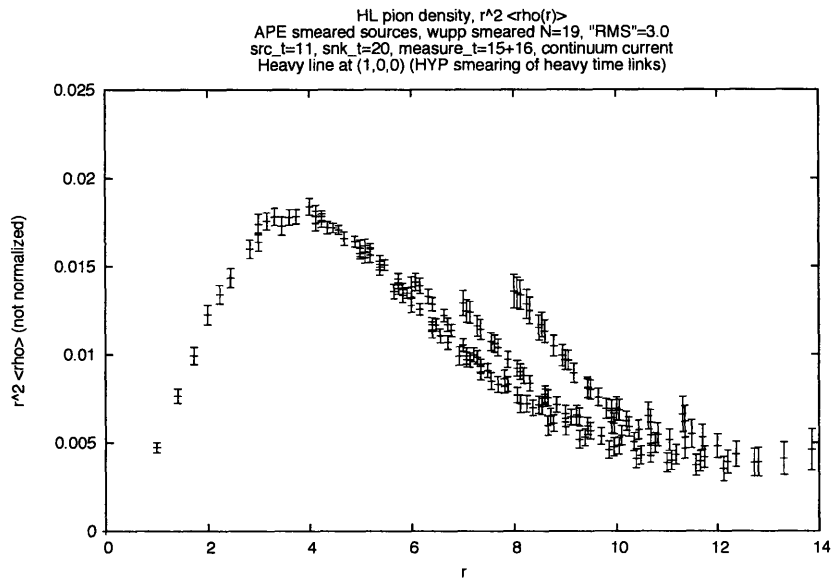


Figure 4-7: Density correlator source at $t_{\mathrm{src}} = 11$ and sink at $t_{\mathrm{snk}} = 20$, central two timeslices averaged; sources at RMS= 3.0. (a) No HYP smearing of time links; (b) Time links are HYP smeared.

Combining HYP smearing and heavy-line averaging, we get an extremely accurate measurement of $\rho(\vec{x})$ on the lattice, as shown in Figure 4-8. The two most serious artifacts that remain are due to the lattice. At high r, we have the periodic images of the sources contributing to the measurement of ρ. At low r, "Cartesian" paths along a single direction to certain lattice points unduly favor quark propagation among them; the small outliers in this region, in fact, reflect the breaking of rotational symmetry at distances on the scale of the lattice spacing. Both artifacts are addressed below.

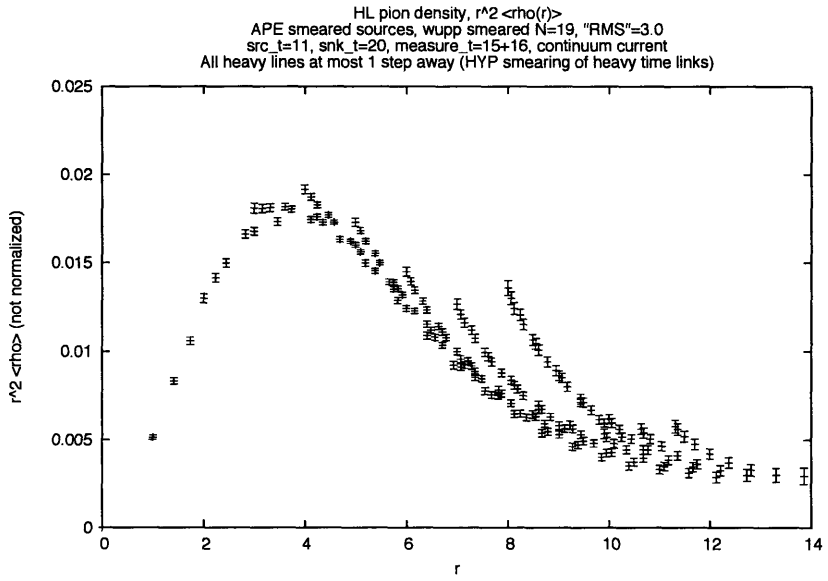
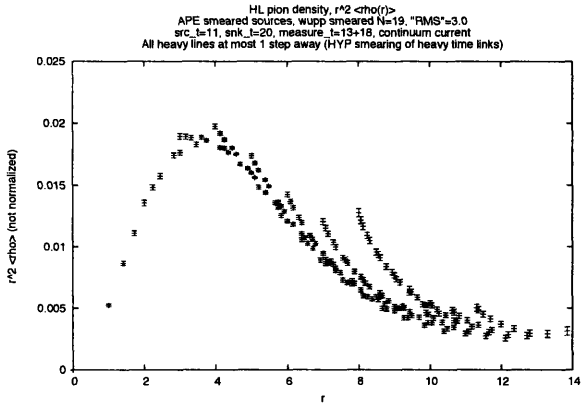


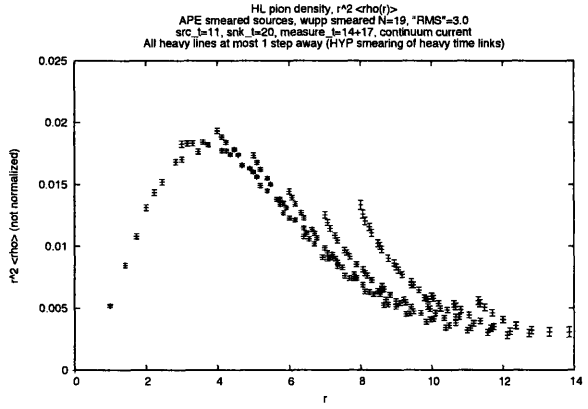
Figure 4-8: Density correlator source at $t_{\text{src}} = 11$ and sink at $t_{\text{snk}} = 20$, central two timeslices averaged; sources at RMS= 3.0, HYP smearing of time links; averaged over 7 heavy lines. No image correction.

Before proceeding, however, we want to verify that having $t_{\text{snk}} = 20$ indeed is large enough to ensure that the ground state has been filtered out. To this end, we calculated ρ with $t = 14$ & 17 and with $t = 13$ & 18. In particular, we should see the plot of ρ stabilize as t is brought further away from the source or sink. This is exactly what is shown in Figure 4-9.

(a)



(b)



(c)

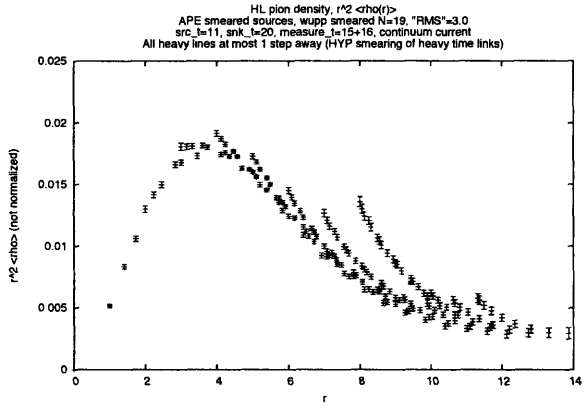


Figure 4-9: With $t_{\mathrm{snk}} - t_{\mathrm{src}}$ set to 9, we see the calculated matrix elements become almost independent of the shape of the source, as they should be if we've allowed enough time between the source/sink and the density operator for excited states to have been filtered out. The density insertions are (a) 2, (b) 3 and (c) 4 timeslices from the source.

Image effects [5] are easy to understand in theory, but harder to correct for in practice. The principle is shown in Figure 4-10 for the 1-dimensional case. In gauge theories, the density of a periodic source and its image is equal to the sum of the two individual densities (in essence, the cross-term is an infinite Polyakov line, with infinite energy; see [5] for the details).

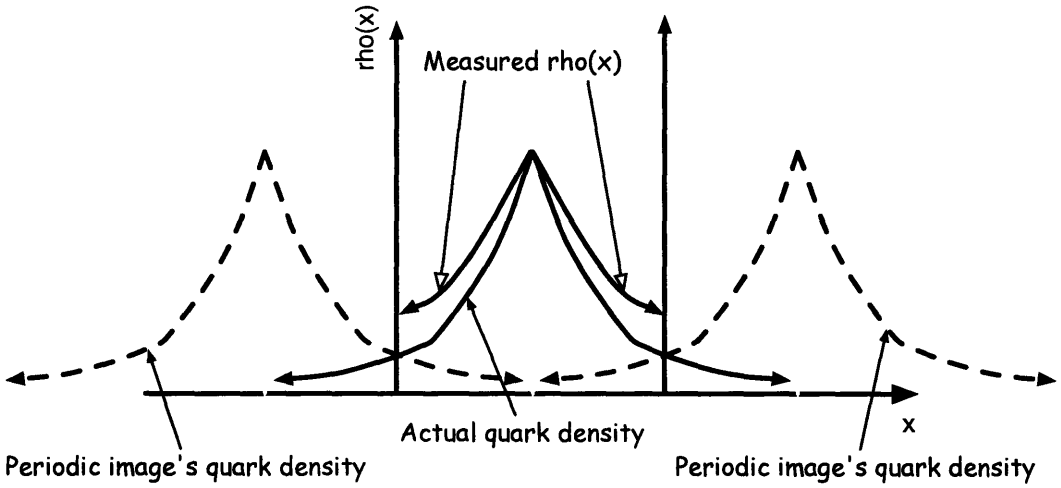


Figure 4-10: Principle behind image effects and their correction.

The contribution of periodic images to the density at a particular point is usually at a region far away from the image's source location. It is known that $\rho(\vec{x})$ decays exponentially as $|\vec{x}| \to \infty$. By assuming this limit is valid for some region Ω within the lattice whose points have high r (for example, $r > 8$), we can calculate what $\rho(\vec{y})$ with $\vec{y} \in \Omega$ should look like. Then by varying the parameters such as the decay rate of $\rho(\vec{r})$ at large r, we can fit the data in Ω to a phenomenological model. Finally, we can use that model to calculate the effect on density of periodic images throughout the lattice, and subtract that effect in a post-processing step.

This exact strategy is implemented below on the lattice, with $\rho(\vec{r}) = A\exp(-m_0 r)$ assumed for $r > 8$. After correcting for the images found with this procedure, we get our final measurement of $\rho(\vec{x})$, shown in Figure 4-11.

The progression from Figure 4-3 to Figure 4-11 shows the cumulative effect of the calculational improvements introduced in this work. These improvements make possible the measurement of diquark correlations in the next section.

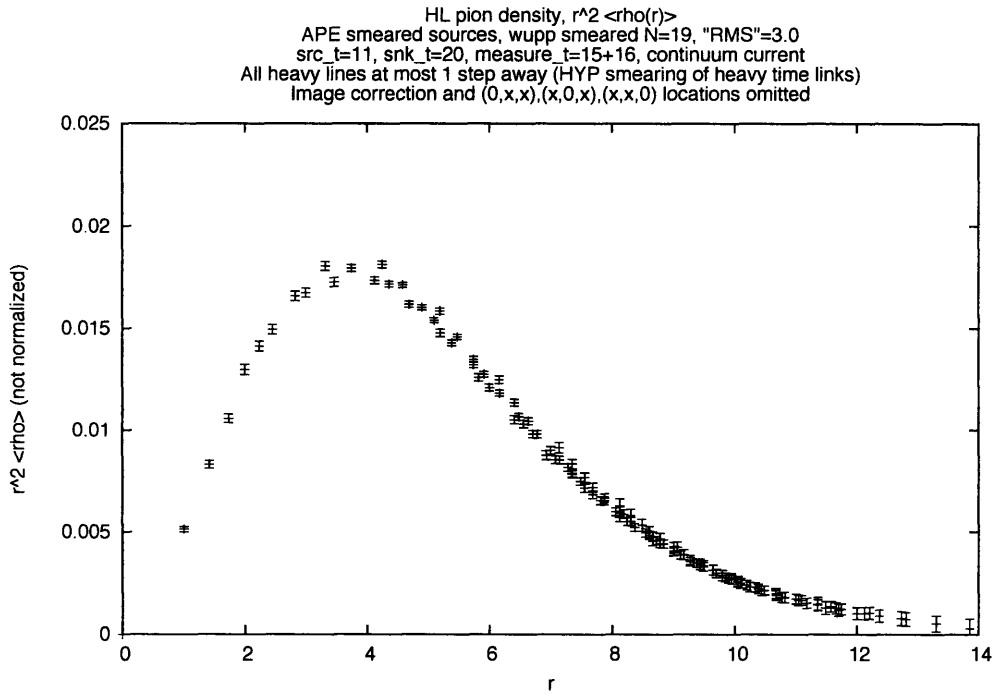


Figure 4-11: Final measurement of $\rho(\vec{r})$. Source at $t_{\text{src}} = 11$ and sink at $t_{\text{snk}} = 20$, central two timeslices averaged; sources at RMS= 3.0, HYP smearing of time links; averaged over 7 heavy lines; Cartesian paths close to origin omitted; image corrected.

4.3 Density-density correlator in heavy-light-light baryon

4.3.1 Diquarks in theory

Within larger hadrons, we call *diquarks* any pairs of quarks that interact strongly. An extended discussion on diquarks can be found in [14]. In gross terms, the most attractive channel for two-quark interactions, after those that induce confinement and chiral symmetry breaking, is thought to be the color $\overline{\mathbf{3}}$, flavor antisymmetric, spin zero channel; this phenomenon is observed in the gluon exchange diagram between two heavy quarks, for instance, and in interactions mediated by instantons. Thus, we expect the two quarks in the diquark $(\bar{q}C\gamma^5 q)$ to interact, but the interaction can't

be calculated analytically.

A simplification used by Jaffe and Wilczek [15] is to treat the diquark as nearly point-like, so that it behaves like an antiquark. They then use ideas from spectroscopy models of meson based on flux tubes, and apply them to hadrons, treating them as quark-antiquark (diquark) pairs. Of course, diquarks are not pointlike objects. Here, we look at exactly how correlated the two quarks in a diquark really are.

4.3.2 Diquarks in practice

In this section, we finally calculate one of the matrix elements that tell us about diquarks: the density-density correlator. We'll define it as follows:

$$\rho\rho(\vec{x},\vec{y}) := \frac{\langle J(\vec{0},t_{\text{snk}})\rho_u(\vec{x},t)\rho_d(\vec{y},t)\overline{J}(\vec{0},t_{\text{src}})\rangle}{\langle J(\vec{0},t_{\text{snk}})\overline{J}(\vec{0},t_{\text{src}})\rangle}.$$

For this initial exploration, we've used a single source, namely $J = (uC\gamma^5 d)s$; the diquark in parenthesis is commonly known as the "good" diquark, in opposition to the "bad" diquark $(uC\gamma^\mu d)$. Although we were unaware of it when we began this work, Alexandrou, de Forcrand and Lucini [1] have performed a similar calculation, albeit in a more restricted geometry. Our work aims to extend their calculation to remove this restriction, and so explore diquarks in a broader regime[1].

From the experimentation with the heavy-light meson, we decided to use the following conditions for the calculations throughout:

- Wuppertal smeared sources, with $\alpha = 3.0$ and $N = 19$, using an APE-smeared ($c = 1/3$, $N = 10$) gauge field for smearing.
- A source and sink locations of $t_{\text{src}} = 11$ and $t_{\text{snk}} = 20$, respectively.
- Data taken at timeslices $t = 15$ and $t = 16$, then averaged.
- HYP smearing of the heavy time links.

[1] Recently, the de Forcrand collaboration has presented work similar in nature to ours

As explained in Section 2.6, we take the mass of the u and d quarks to be equal. In this case, the source is symmetric under the exchange of u and d; thus, we expect the following equation to hold:

$$\rho\rho(\vec{x},\vec{y}) = \rho\rho(\vec{y},\vec{x}).$$

Indeed, we used this condition, among others, to check the validity of our codes.

The density-density correlator allows us to study a diquark in the color field of an infinitely massive quark: we cannot study a diquark in isolation because the structure is not a color singlet. It is important, then, to separate the effects of the polarizing heavy quark field from interactions between the two quarks. For this purpose, define the single-quark density as in Section 4.2, namely

$$\rho_u(\vec{x}) := \frac{\langle J(\vec{0},t_{\text{snk}})\rho_u(\vec{x},t)\overline{J}(\vec{0},t_{\text{src}})\rangle}{\langle J(\vec{0},t_{\text{snk}})\overline{J}(\vec{0},t_{\text{src}})\rangle}.$$

The subscript is there to distinguish between different quark flavor densities. For our particular source, the manifest u/d symmetry ensures $\rho_u(\vec{x}) = \rho_d(\vec{x})$. We then define the following *correlation function*:

$$C(\vec{x},\vec{y}) := \frac{\rho\rho(\vec{x},\vec{y}) - \rho_u(\vec{x})\rho_d(\vec{y})}{\rho_u(\vec{x})\rho_d(\vec{y})}.$$

If the two light quarks had independent dynamics, then we'd have $\rho\rho(\vec{x},\vec{y}) = \rho_u(\vec{x})\rho_d(\vec{y})$, so $C(\vec{x},\vec{y}) = 0$. It follows that $C(\vec{x},\vec{y})$ provides a convenient, dimensionless measure of how the behaviour of the two light quarks is correlated. A more convenient expression for $C(\vec{x},\vec{y})$ is given below:

$$C(\vec{x},\vec{y}) = \frac{\rho\rho(\vec{x},\vec{y})}{\rho_u(\vec{x})\rho_d(\vec{y})} - 1. \tag{4.3}$$

To set the stage, we first calculate $\rho_u(\vec{r})$ and plot $r^2\rho_u(\vec{r})$ in Figure 4-12. For this initial exploration, we used a single heavy line and did no image correction. The former results in larger error bars, while the latter results in systematic errors.

Nevertheless, as is argued in [1], because both ρ_u and $\rho\rho$ suffer from the same kinds of image effects, these effects cancel to a large extent in the correlation function.

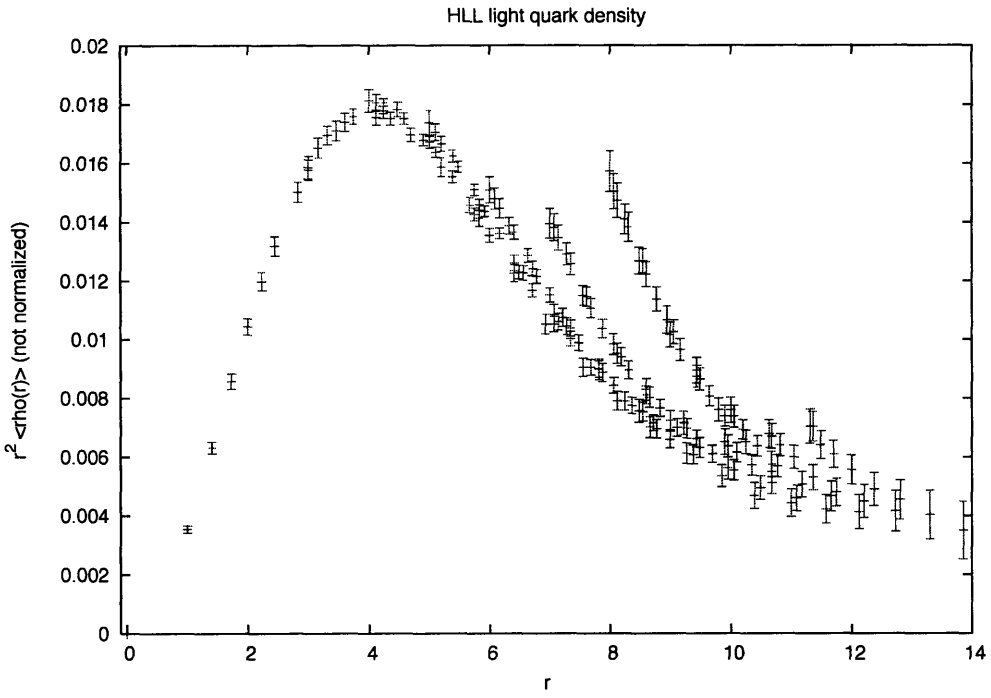
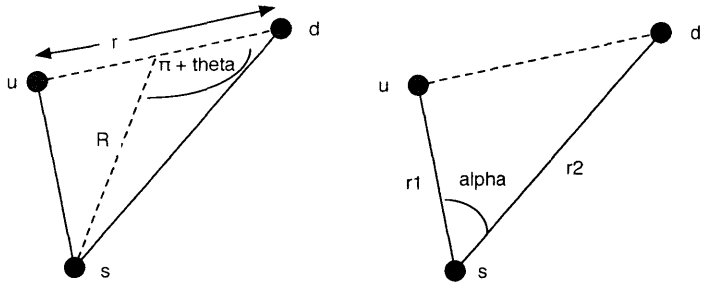


Figure 4-12: Single-quark density correlator for heavy-light-light baryon.

Throughout, all values of $\rho\rho(\vec{x},\vec{y})$ and $C(\vec{x},\vec{y})$ are averaged over cubically equivalent pairs of points. As before, there are up to 48 of these (96 counting exchanging $\vec{x}$ and $\vec{y}$, but as discussed before, this exchange yields no new information), again corresponding to cubical rotations and reflections.

Cartesian coordinates are ill-suited to describing $C(\vec{x},\vec{y})$: as it stands, the function has six degrees of freedom, yet symmetry considerations dictate there be only three. Thus, we use the following two coordinate systems; the first one is more natural to described a two-particle body around the heavy quark, while the second one (from [1]) is more useful for describing angular correlations as measured from the heavy quark.



As a first measurement, Figure 4-13 shows $\rho\rho(R, r, \theta = 0)$ for various R values, as a function of r. Setting $\theta = 0$ corresponds to making perpendicular the line joining the two light quarks and that from the center of this line to the heavy quark. We expect $\rho\rho$ to decay exponentially as a function of r in that case, and indeed it does. We expect pronounced image effects when $\sqrt{R^2 + (r/2)^2} \approx 7$ or 8, that is, when $\vec{x}$ or $\vec{y}$ are close to the edge of the lattice. For $R = 2, 3, 4$, and 5, that means there should be an anomaly around r = 13–15, 12–14, 11–14, and 10–12, respectively. These ranges coincide roughly with the observed "bumps" in 4-13.

Next, we reproduce the qualitative result of [1] when it comes to observing attractive behavior in the "good" diquark channel (see their Figure 2; where we write α, they write θ). They calculated $C(r_1 = r_2, \alpha)$ for $r_1 = 5.1$ and plot it versus $\cos\alpha$. Our equivalent plots for $r_1 = r_2 = 4$ and 6 are shown in Figure 4-14. The conclusion to be derived is that there is clear evidence for the existence of an attractive interaction between the two light quarks, as C rises sharply above 0 as $\cos\alpha \to 1$.

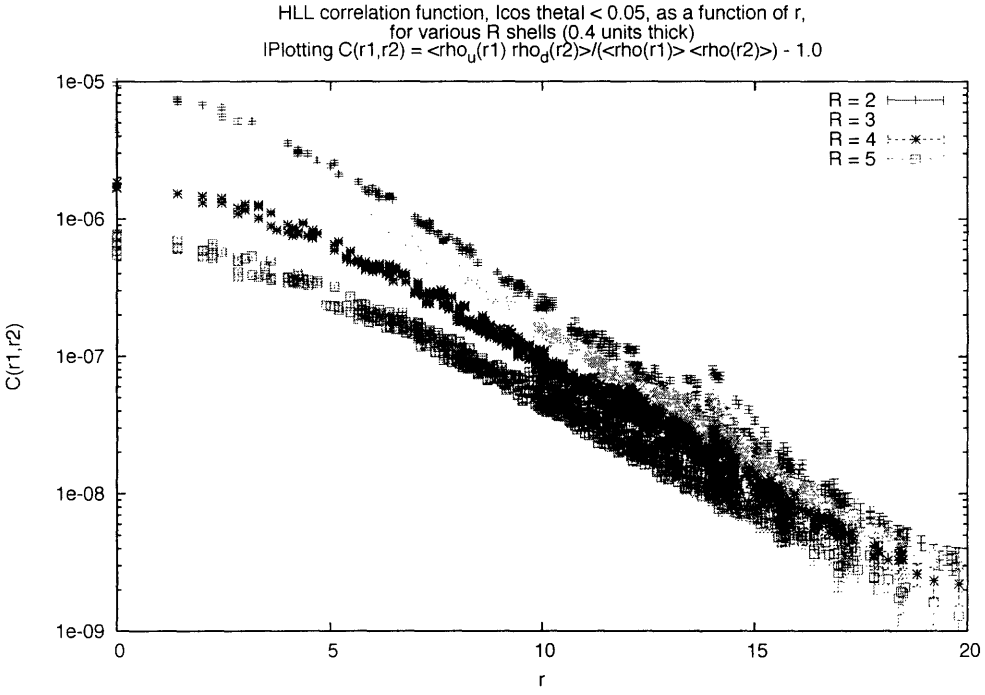


Figure 4-13: Consistency check for $\rho\rho(R, r, \theta = 0)$ calculation.

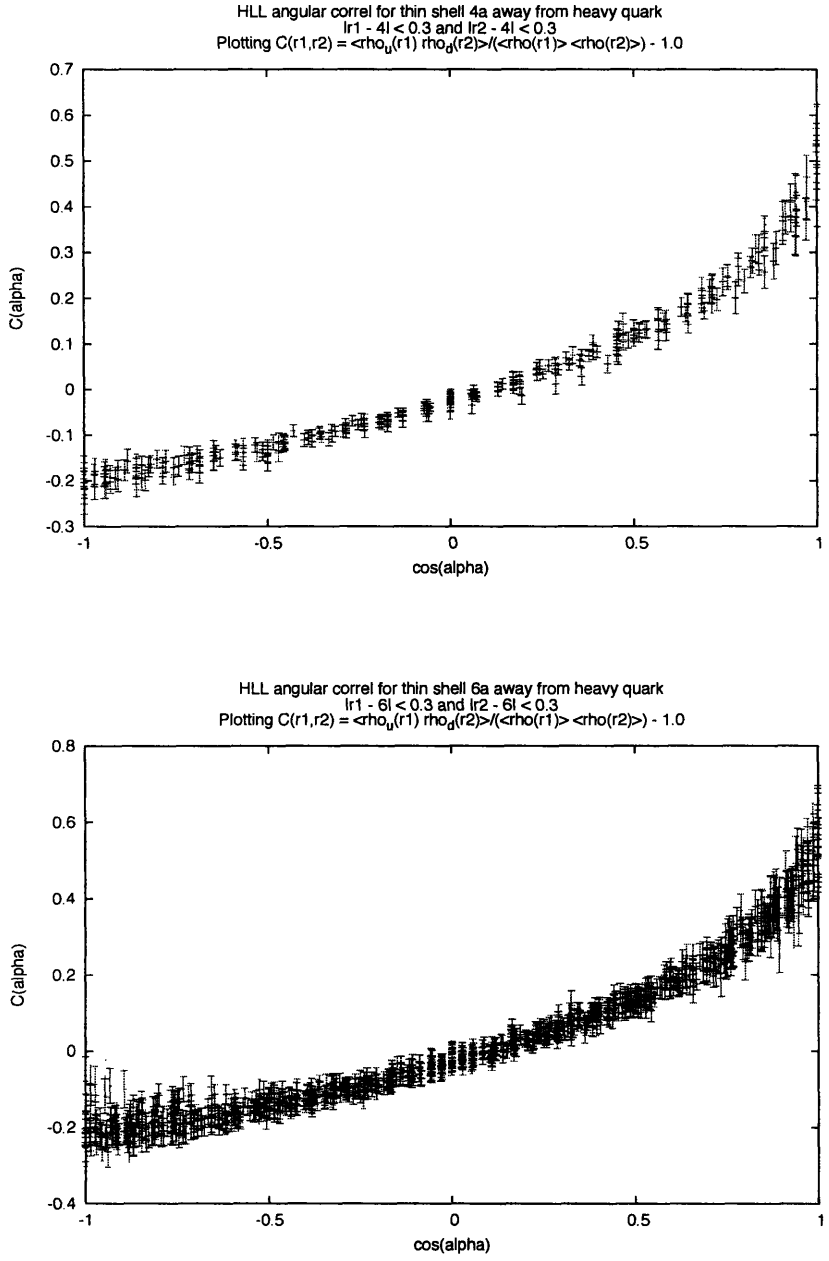


Figure 4-14: Angular correlation function $C(r_1 = r_2 = r, \alpha)$ for $3.7 \leq |r| \leq 4.3$ (top); and $5.7 \leq |r| \leq 6.3$ (bottom).

Now we can go beyond the work in [1]. The correlation function $C(-\vec{r}, \vec{r})$ has only one degree of freedom and serves as another rough measure to observe the behavior of the diquark. However, very few pairs of points $(-\vec{r}, \vec{r})$ are available in the lattice. Instead, we used two pairs $\vec{r}_1$ and $\vec{r}_2$ whose center of mass is within a unit of the heavy line, which greatly enhances the number of points for which $C(-\vec{r}, \vec{r})$ is available while only compromising the analysis minimally. The results are shown in Figure 4-15(a). In part (b), we additionally exclude any data point with Cartesian paths to them, to mitigate some of the worse image effects. The results are shown in Figure 4-16.

The figures clearly show again that the light quarks like to be together at short distances, and have a corresponding lower probability of being separated by a large spatial distance r.

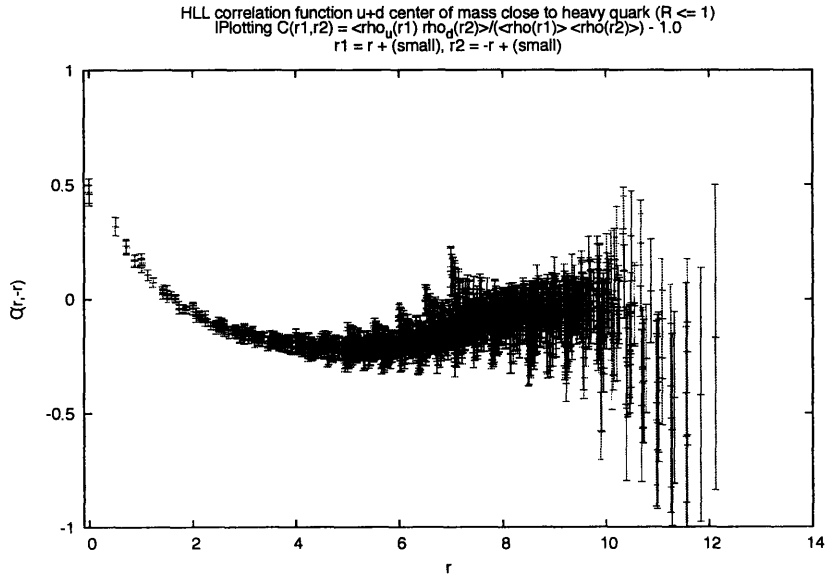


Figure 4-15: Correlation functions of (nearly) diametrically opposite light quarks

Figure 4-17 shows that $C(R, r, \theta)$ is only weakly dependent on θ. In other words, it appears the polarization of the diquark pair induced by its interaction with the heavy quark is of the order of the noise of the calculation. With this in mind, we can eliminate one degree of freedom and simplify our analysis of the diquark. In particular, after fixing R, we can observe the decay rate in r of $C(R, r)$ to look at the

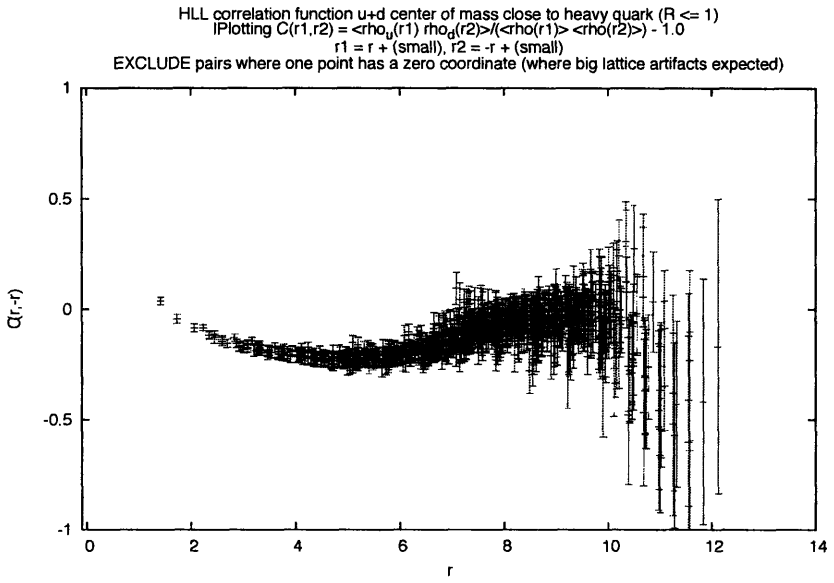


Figure 4-16: Same as Figure 4-15, but excluding coordinates along Cartesian paths.

spatial extent of a diquark. This idea is realized in Figure 4-18. There, we plot $C(r)$ for $R = 4.0 \pm 0.2$ and a small fraction of θ space. When θ is set to other values, the plots look qualitatively similar but see to have stronger image effects.

4.4 Conclusion

The results described above highlight that diquarks do form inside a model baryon; moreover, in real QCD, they're not point-like, but rather, have spatial extent of the same order of magnitude as the RMS radius of the light quark distribution. An immediate improvement to our present analysis is to incorporate image correction and heavy-line averaging, thus increasing statistics and reducing systematic errors. It's then one extra simple step to redo these calculations with the "bad" diquark sources to have a basis for comparison. The successful result of this research motivates further study in these directions, which we do plan perform and publish in the near future.

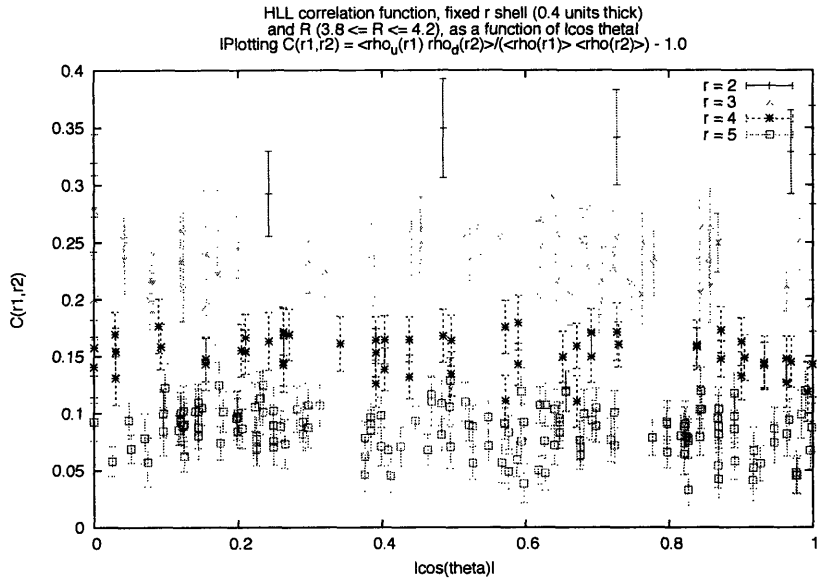


Figure 4-17: $C(R, r, \theta)$ plotted with respect to θ at various r values, with $R = 4.0 \pm 0.2$. Notice the lack of θ-dependence.

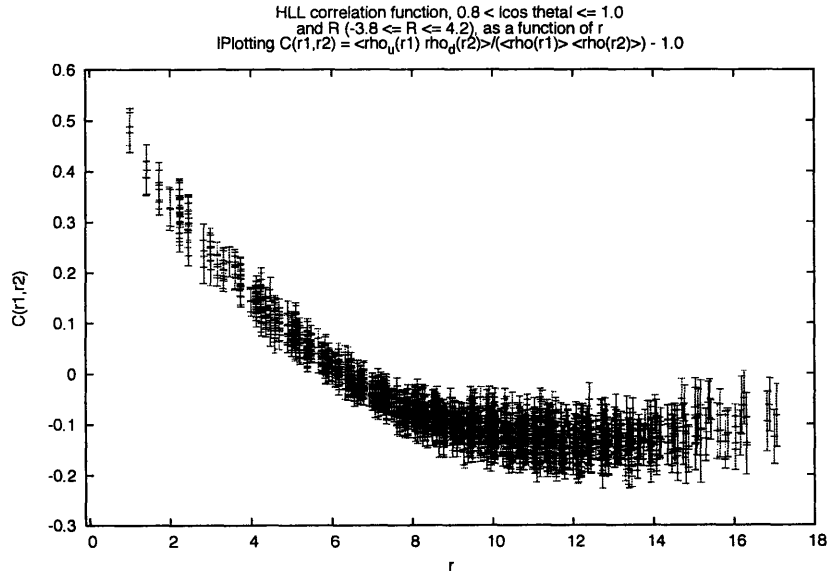


Figure 4-18: $C(R, r, \theta)$ plotted with respect to r at $|\cos\theta| > 0.8$, with $R = 4.0 \pm 0.2$. Plots for other θ slices are very similar but seem to have stronger image effects.

Bibliography

[1] Alexandrou, C., De Forcrand, P., and Lucini, B., *Searching for Diquarks in Hadrons*, Proc. of Science (LAT2005) (2005), 053.

[2] Bali, G.S., and Schilling, K., *Static quark-antiquark potential: Scaling behavior and finite-size effects in SU3 lattice gauge theory*, Phys. Rev. D **46** (1992), 2636.

[3] Bhattacharya, T., Gupta, R., Kilcup, G., and Sharpe, S., *Hadron Spectrum with Wilson fermions*, Phys. Rev. D **53** (1996), 6486–6508.

[4] Brown, L.S., "Quantum Field Theory," Cambridge University Press, 1992.

[5] Burkardt M., Grandy J.M., and Negele J.W., *Calculation and Interpretation of Hadron Correlation Functions in Lattice QCD*, Annals of Physics **238** (1995), 441-472.

[6] Cabibbo, N., and Marinari, E., *A New Method for Updating SUN Matrices in Computer Simulations of Gauge Theories*, Phys. Lett. B **119** (1982), 387–390.

[7] Creutz, M., "Quarks, Gluons and Lattices," Cambridge University Press, 1983.

[8] Edwards, R.G., and Joó B., *The Chroma Software System for Lattice QCD*, Nucl. Phys. B (Proc. Suppl) **140** (2005), 832–834.

[9] Efron, B., "The Jackknife, the Bootstrap and Other Resampling Plans," Society for Industrial and Applied Mathematics, 1982.

[10] Eidelman, S. et al., *Review of Particle Physics*, Phys. Lett. B **592** (2004), 1–5.

[11] Falcioni, M., Paciello, M., Parisi, G., and Taglienti, B., *Again on SU3 glueball masses*, Nucl. Phys. **B251** (1985), 624–632.

[12] Gusken, S.. *A study of smearing techniques for hadron correlation functions,* Nucl. Phys. Proc. Suppl. **17** (1990), 361364.

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

[14] Jaffe, R.L., *Exotica*, Nucl. Phys. Proc. Suppl. **142** (2005), 343-355.

[15] Jaffe, R.L., and Wilczek, F., *Diquarks and Exotic Spectroscopy*, Phys. Rev. Lett. **91** (2003), 232003.

[16] Leinweber, *Visualizations of Quantum Chromodynamics*, `http://www.physics.adelaide.edu.au/~dleinweb/VisualQCD/Nobel/index.html`

[17] Lubicz, V., *Phenomenology of the Standard Model from Lattice QCD*, `hep-ph/0010171`.

[18] Luscher, M., *Construction of a self-adjoint, strictly positive transfer matrix for Euclidean lattice gauge theories*, Comm. Math. Phys. **54** (1977), 283–292.

[19] McClendon, C., *Optimized Lattice QCD Kernels for a Pentium 4 Cluster*, Jlab preprint, JLAB-THY-01-29, `http://www.jlab.org/~edwards/qcdapi/reports/dslash_p4.pdf`

[20] Michael, C., *Fitting correlated data*, Phys. Rev D **49** (1994), 2616–2619.

[21] Negele, J.W., *Understanding Hadron Structure Using Lattice QCD*, `hep-lat/9804017`.

[22] Negele, J. W., and Orland, H., "Quantum Many-Particle Systems," Perseus Books Group, 1998.

[23] NERSC (National Energy Research Scientific Computing Center), `http://qcd.nersc.gov/configs/OSU/Q60a/`.

[24] Peskin, M.E., and Schroeder, D.V., "An Introduction to Quantum Field Theory," Westview Press, 1995.

[25] Renner, D.B., "Exploring Proton Structure using Lattice QCD," MIT Ph. D. thesis, 2004.

[26] Wilczek, F., *Opportunities, challenges, and fantasies in lattice QCD*, Nucl. Phys B (Proc. Suppl) **119** (2003), 3–12.

[27] Wilson, K.G., *Confinement of Quarks*, Phys. Rev. D **10** (1974), 24452459