Chapter 2

Lattice QCD

2.1 QCD in the Continuum

2.1.1 The QCD Lagrangian

We begin by defining the QCD Lagrangian [14] in Euclidean spacetime in terms of quark fields and gluon fields, and discuss the various transformation properties of those fields and the QCD action. As we will see, the advantage of working in Euclidean spacetime is that oscillating exponentials in time become decaying exponentials. This allows us to interpret the Boltzmann factor e^{-S} (where S is the action in Euclidean spacetime) as a probability distribution, allowing us to use Monte Carlo integration to evaluate path integrals in quantum field theory. Additionally, we will see that this gives us access to the low-lying energy spectrum of a theory by allowing us to fit two-point correlation functions to real decaying exponentials. The QCD Lagrangian $\mathcal{L}_{\mathrm{QCD}}$ can be written in terms of the fermionic part $\mathcal{L}_{\mathrm{F}}$ and pure-gluon part $\mathcal{L}_{\mathrm{G}}$. In a free theory (i.e. without gluons), the fermionic Lagrangian is

$$\mathcal{L}_{\mathrm{F, \ free}} = \sum_f \overline{\psi}_{\alpha c}^{(f)} \left((\gamma_\mu)_{\alpha\beta} \partial_\mu + \delta_{\alpha\beta} m^{(f)} \right) \psi_{\beta c}^{(f)}(x), \tag{2.1}$$

where a sum over repeated indices is implied. ψ and $\overline{\psi}$ are quark and anti-quark fields, respectively, and they carry a flavor index f, a Dirac spin index α, and a color index c. $m^{(f)}$ is the bare mass of the fermion with flavor f, and γ_μ are the Euclidean analogs to the familiar Minkowski γ-matrices satisfying the anticommutation relation $\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}\mathbb{1}$, where $\mu \in \{1, 2, 3, 4\}$ is the Euclidean spacetime index. In its non-interacting form, Eq. (2.1) appears identical to the familiar QED Lagrangian, except with additional indices for color, which do not mix.

In an interacting theory with gluons, the fermionic part of the Lagrangian becomes

$$\mathcal{L}_{\mathrm{F}} = \sum_f \overline{\psi}_{\alpha c}^{(f)} \left((\gamma_\mu)_{\alpha\beta} (\delta_{cd} \partial_\mu + i A_{\mu cd}(x)) + \delta_{\alpha\beta} \delta_{cd} m^{(f)} \right) \psi_{\beta d}^{(f)}(x), \tag{2.2}$$

where we have introduced the gluon fields $A_{\mu cd}$, which carry color indices in addition to spacetime indices. The matrices A_μ are hermitian and traceless in color space. Being matrices, they make QCD a *non-Abelian* gauge theory, as the gluon fields themselves do not commute. The non-Abelian nature of QCD is the main reason why extracting the low-energy physics is difficult. We define the covariant derivative

$$D_\mu(x) = \partial_\mu + i A_\mu(x) \tag{2.3}$$

and the field strength tensor

$$F_{\mu\nu}(x) = -i[D_\mu(x), D_\nu(x)] = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) + i[A_\mu(x), A_\nu(x)]. \tag{2.4}$$

The gauge fields $A_\mu(x)$ are traceless and hermitian and belong to the Lie algebra $\mathfrak{su}(3)$. Therefore, we can write A_μ in terms of the 8 generators of SU(3), denoted by T_i, which form a basis for $\mathfrak{su}(3)$, as

$$A_\mu(x) = \sum_{i=1}^{8} A_\mu^{(i)}(x) T_i. \tag{2.5}$$

We call $A_\mu^{(i)}(x)$ the *color components* of the gauge field, and they are real-valued. The generators satisfy the following properties in addition to being traceless, complex, and hermitian:

$$\mathrm{tr}\,[T_j T_k] = \frac{1}{2} \delta_{jk} \tag{2.6}$$

and

$$[T_j, T_k] = i f_{jkl} T_l, \tag{2.7}$$

where f_{jkl} are known as the *structure constants* and are completely anti-symmetric. Conventionally, we choose to represent $\mathfrak{su}(3)$ using $T_a = \frac{\lambda_a}{2}$, where λ_a are the Gell-Mann matrices [16].

Now we write down the gluonic part of the QCD Lagrangian,

$$\mathcal{L}_{\mathrm{G}} = \frac{1}{2g^2} \mathrm{tr}[F_{\mu\nu}(x) F_{\mu\nu}(x)], \tag{2.8}$$

where g is the QCD coupling parameter. Using the decomposition of the gluon fields into

their color components, we can similarly decompose the field strength tensor into its color components to gain more insight into the gluonic part of the Lagrangian:

$$
\begin{aligned}
F_{\mu\nu}(x) &= \sum_{i=1}^{B} F_{\mu\nu}^{(i)}(x) T_i, \\
F_{\mu\nu}^{(i)}(x) &= \partial_\mu A_\nu^{(i)}(x) - \partial_\nu A_\mu^{(i)}(x) - f_{ijk} A_\mu^{(j)}(x) A_\nu^{(k)}(x).
\end{aligned}
\tag{2.9}
$$

Plugging this into Eq. (2.8), we find that the QCD Lagrangian contains 3- and 4-vertex gluonic self-interactions. This makes QCD and QED fundamentally different: in QED we have linear superposition since the massless vector boson (the photon) does not self-interact, whereas in QCD, linear superposition never applies due to the gluon self-interactions.

2.1.2 Transformation Properties of the Quark and Gluon Fields

QCD is an SU(3) gauge theory, meaning the action is invariant under local SU(3) gauge transformations of the quark and gluon fields. Such a gauge transformation can be associated with a 3×3 unitary color-space transformation $\Omega(x)$ at each spacetime point x, which also has the property $\det[\Omega(x)] = 1$. In order to ensure gauge invariance, we demand that the quark fields transform as

$$
\begin{aligned}
\psi(x) &\to \psi'(x) = \Omega(x)\psi(x) \\
\overline{\psi}(x) &\to \overline{\psi}'(x) = \overline{\psi}(x)\Omega(x)^\dagger,
\end{aligned}
\tag{2.10}
$$

and the gluon fields transform as

$$
A_\mu(x) \to A_\mu'(x) = \Omega(x) A_\mu(x) \Omega(x)^\dagger + \mathrm{i}\left(\partial_\mu \Omega(x)\right) \Omega(x)^\dagger,
\tag{2.11}
$$

where we have suppressed Dirac indices and used matrix/vector notation for the color indices. One can then verify that Eqs. (2.10) and (2.11) leave the action $S = \int d^4 x \{\mathcal{L}_{\mathrm{F}} + \mathcal{L}_{\mathrm{G}}\}$ invariant.

2.2 Discretizing QCD on the Lattice

To facilitate our nonperturbative computations, we must formulate QCD on a spacetime lattice. We consider our lattice Λ as an isotropic four-dimensional grid of points, indexed by a 4-tuple of integers n such that $x = an$, where a is the lattice spacing. That is, we can write

$$
\Lambda = \{n = (n_1, n_2, n_3, n_4)\},
\tag{2.12}
$$

where n_i are integers. We will later also consider anisotropic lattices. Since a lattice has finite spacing, that implies a momentum cutoff, which also acts as an ultraviolet regulator for the theory. That is, the i component momentum is restricted to $p_i \in (-\frac{\pi}{a}, \frac{\pi}{a}]$. Additionally, we will work with periodic boundary conditions, which discretizes the momentum $\boldsymbol{p} = \frac{2\pi \boldsymbol{n}}{L}$, where $\boldsymbol{n}$ is a vector of integers and L is the length of the lattice.

2.2.1 Free Fermions

We begin with a naive approach to discretizing fermions on the lattice, which will then be improved upon. The continuum action for one free fermion with a bare mass m can be written as

$$S_F^0[\psi, \overline{\psi}] = \int d^4x \, \overline{\psi}(x) \left(\gamma_\mu \partial_\mu + m \right) \psi(x). \tag{2.13}$$

A first pass at discretizing Eq. (2.13) involves replacing the integral by a sum, replacing x by n, and rewriting the derivative terms as simple finite differences, like so,

$$\int d^4x \rightarrow a^4 \sum_n \tag{2.14}$$

$$\psi(x) \rightarrow \psi(n) \tag{2.15}$$

$$\overline{\psi}(x) \rightarrow \overline{\psi}(n) \tag{2.16}$$

$$\partial_\mu \psi(x) \rightarrow \frac{1}{2a}(\psi(n + \hat{\mu}) - \psi(n - \hat{\mu})). \tag{2.17}$$

We immediately find that such a discretization scheme breaks local SU(3) gauge invariance, due to terms of the form $\overline{\psi}(n)\psi(n + \hat{\mu})$.

2.2.2 Gauge Fields and Link Variables

Though we expect that gauge invariance for the free fermion action is restored in the limit $a \rightarrow 0$, we wish to maintain exact gauge invariance for finite a in the entire discretized action. In determining how to discretize the gluon fields, we are guided by this requirement. Under a local gauge transformation $\Omega(n)$,

$$\overline{\psi}(n)\psi(n + \hat{\mu}) \rightarrow \overline{\psi}(n)\Omega(n)^\dagger \Omega(n + \hat{\mu})\psi(n + \hat{\mu}), \tag{2.18}$$

which is obviously not gauge-invariant. We can fix this issue by introducing new fields $U_\mu(n)$, called *link variables* [45], which transform under a gauge transformation as

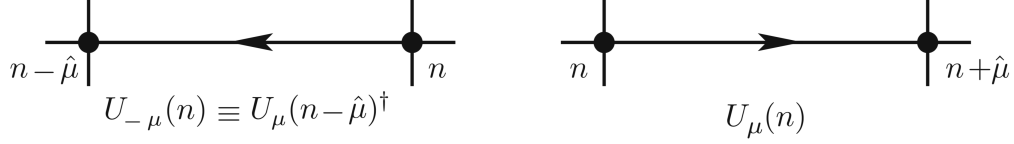$$U_\mu(n) \rightarrow \Omega(n)U_\mu(n)\Omega(n + \hat{\mu})^\dagger. \tag{2.19}$$

5

Figure 2.1: Link variables, depicted as linking two adjacent lattice sites. Figure from Ref. [15].

These link variables are matrix-valued, and are elements of the gauge group SU(3). They are also directionally oriented, and are defined such that $U_{-\mu}(n) \equiv U_\mu(n - \hat{\mu})^\dagger$. They can be visually represented as living on the links between lattice sites, as depicted in Fig. 2.1. We then propose the following form for the naively discretized interacting fermionic action for one flavor as follows, which one can verify to be gauge-invariant:

$$S_F[\psi, \overline{\psi}, U] = a^4 \sum_{n \in A} \overline{\psi}(n) \left(\sum_{\nu=1}^{4} \gamma_\mu \frac{U_\mu(n)\psi(n + \hat{\mu}) - U_{-\mu}(n)\psi(n - \hat{\mu})}{2a} + m\psi(n) \right). \quad (2.20)$$

In taking the limit $a \to 0$, the fermionic action will contain $\mathcal{O}(a)$ lattice artifacts. Now that we have identified a candidate for a gauge-invariant discretized interacting fermion action, it remains to relate the link variables $U_\mu(n)$ to the continuum gauge fields $A_\mu(x)$. In the continuum, there is an object which satisfies the same gauge transformation properties as the link variables (see Eq. (2.19)), and that object is the gauge transporter,

$$G(x, y) = P \exp \left(i \int_{\mathcal{C}_{xy}} A \cdot ds \right), \quad (2.21)$$

where $\mathcal{C}_{xy}$ is a curve connecting two points x and y, and P denotes path-ordering. The link variables approximate this gauge transporter, i.e. $U_\mu(n) = G(n, n + \hat{\mu}) + \mathcal{O}(a)$, and are given by

$$U_\mu(n) = \exp\left(iaA_\mu(n) \right). \quad (2.22)$$

One can verify that in the continuum limit $a \to 0$, the action in Eq. (2.20) reduces to its continuum form given by integrating the Lagrangian in Eq. (2.2). Note that we have gone from using $\mathfrak{su}(3)$ Lie algebra-valued fields $A_\mu(x)$ to variables $U_\mu(x)$ which are elements of the Lie group SU(3). This will become relevant in defining the path integral on the lattice.

2.2.3 Fermion Doubling

We define the *Dirac Matrix $M[U]$* for a theory with one fermion flavor by rewriting the fermion action in Eq. (2.20) as

$$S_F[\psi, \overline{\psi}, U] = \sum_{n,m \in \Lambda} \sum_{a,b,\alpha,\beta} \overline{\psi}(n)_{\alpha a} M(n|m)_{\alpha\beta;ab} \psi(m)_{\beta b}, \tag{2.23}$$

implying that

$$M(n|m)_{\alpha\beta;ab} = a^4 \sum_{\mu=1}^{4} (\gamma_\mu)_{\alpha\beta} \frac{U_\mu(n)_{ab}\delta_{n+\hat{\mu},m} - U_{-\mu}(n)_{ab}\delta_{n-\hat{\mu},m}}{2a} + m\delta_{\alpha\beta}\delta_{ab}\delta_m. \tag{2.24}$$

The free quark propagator on the lattice is found by inverting the Dirac matrix when the gauge fields are set to $U_\mu(n) = \mathbb{1}$. In momentum space, the Dirac matrix is given by

$$\widetilde{M}(p) = a^4 \left(m + i \sum_{\mu=1}^{4} \gamma_\mu \frac{\sin(p_\mu a)}{a} \right) \tag{2.25}$$

and its inverse is

$$\widetilde{M}^{-1}(n|m) = a^{-4} \frac{m - ia^{-1} \sum_\mu \gamma_\mu \sin(p_\mu a)}{m^2 + a^{-2} \sum_\mu \sin(p_\mu a)^2}. \tag{2.26}$$

If we examine this propagator in the massless case, $m = 0$, we find a pole at $p^2 = 0$, as we would expect. However, we also find 15 additional poles at the corners of the first Brillouin zone, i.e. at $p = (\pi/a, 0, 0, 0), (0, \pi/a, 0, 0), \ldots, (\pi/a, \pi/a, \pi/a, \pi/a)$. These extra poles are known as the *doublers*. They are not present in the continuum (just take $a \to 0$ in the denominator of the propagator to see this), but on the lattice, we must deal with these unphysical poles. Wilson found a clever way to deal with these doublers [46] by adding an additional term, now known as the *Wilson term*, to the Dirac matrix. In momentum space, the Dirac matrix with the added Wilson term is (c.f. Eq. (2.25))

$$\widetilde{M}(p) = a^4 \left(m + i \sum_{\mu=1}^{4} \gamma_\mu \frac{\sin(p_\mu a)}{a} + \sum_{\mu=1}^{4} \frac{(1 - \cos(p_\mu a))}{a} \right). \tag{2.27}$$

The Wilson term vanishes at $p = (0, 0, 0, 0)$, but acts as an extra mass term at each corner of the Brillouin zone. These mass terms scale as $\frac{1}{a}$, so for small lattices, they will decouple from the theory and not affect the low energy spectrum. Inverting the Fourier transform, the Wilson term turns out to be a discretization of $-\frac{a}{2}\partial_\mu\partial_\mu$, which vanishes in the continuum limit.

While the Wilson term solves our fermion doubling problem, it introduces a new problem. Since the Wilson term is proportional to $\mathbb{1}$ in Dirac space, it does not anticommute with $\gamma_5 = \gamma_4 \gamma_1 \gamma_2 \gamma_3$, which implies that it violates chiral symmetry. (See Section 7.2 in Ref. [15] for more details.) This problem of chiral symmetry breaking is not a defect of only Wilson's solution to fermion doubling, but of any solution that attempts to add terms to free the theory of doublers, as shown in the famous "no-go" theorem of Nielson and Ninomiya [37].

2.2.4 The Wilson Gauge Action

After constructing a discretized version of the fermion action, it remains to construct a discretized version of the gauge action. The first step is to identify a gauge-invariant object constructed only of link variables. An ordered product of link variables

$$P[U] = U_{\mu_0}(n_0)\, U_{\mu_1}(n_0 + \hat{\mu}_0) \ldots U_{\mu_{k-1}}(n_1 - \hat{\mu}_{k-1}) \equiv \prod_{(n,\mu) \in \mathcal{P}} U_\mu(n) \tag{2.28}$$

along a path $\mathcal{P}$ on the lattice is the lattice analog to the continuum gauge transporter. This product shares the same transformation properties as the link variables and the continuum gauge transporter, i.e.

$$P[U] \to \Omega(n_0)\, P[U] \Omega(n_1)^\dagger. \tag{2.29}$$

If we choose a path that forms a closed loop $\mathcal{L}$, then such a product transforms under a gauge transformation as

$$\prod_{(n,\mu) \in \mathcal{L}} U_\mu(n) \to \Omega(n_0) \prod_{(n,\mu) \in \mathcal{L}} U_\mu(n) \Omega(n_0)^\dagger, \tag{2.30}$$

since the color rotations $\Omega(n)$ cancel out at every intermediate point due to unitary. We can then use the cyclicity of the trace to form a gauge-invariant object out of Eq. (2.30). Hence,

$$\text{tr} \left[\prod_{(n,\mu) \in \mathcal{L}} U_\mu(n) \right] \tag{2.31}$$

is gauge-invariant. The shortest closed loop is called the *plaquette*, which is defined as a product of four link variables,

$$U_{\mu\nu}(n) = U_\mu(n) U_\nu(n + \hat{\mu}) U_{-\mu}(n + \hat{\mu} + \hat{\nu}) U_{-\nu}(n + \hat{\nu}). \tag{2.32}$$

With these definitions, we now write down Wilson's formulation of the gauge action on the lattice, which is a sum over all possible plaquettes on the lattice (only counting one orientation for each plaquette):

$$S_G[U] = \frac{2}{g^2} \sum_{n \in \Lambda} \sum_{\mu < \nu} \mathrm{Re} \, \mathrm{tr} \left[\mathbb{1} - U_{\mu\nu}(n) \right]. \tag{2.33}$$

In the continuum limit $a \to 0$, it can be shown that this reduces to the continuum version given by integrating the Lagrangian in Eq. (2.8). In taking the limit, the action will contain $\mathcal{O}(a^2)$ lattice artifacts, that is,

$$S_G[U] = \frac{2}{g^2} \sum_{n \in \Lambda} \sum_{\mu < \nu} \mathrm{Re} \, \mathrm{tr} \left[\mathbb{1} - U_{\mu\nu}(n) \right] = \frac{a^4}{2g^2} \sum_{n \in A} \sum_{\mu,\nu} \mathrm{tr} \left[F_{\mu\nu}(n)^2 \right] + \mathcal{O} \left(a^2 \right). \tag{2.34}$$

The Wilson action is just one way to formulate a discretized gauge action. We will discuss different discretization strategies in order to reduce the discretization error to higher orders in a.

2.2.5 Improving the Discretized Action

In our naively discretized theory, the fermionic action action contains $\mathcal{O}(a)$ lattice artifacts and the gauge action contains $\mathcal{O}(a^2)$ artifacts. By adding terms to the action which cancel these artifacts but which vanish in the continuum limit, we can improve discretization errors to higher orders in a. This is called *Symanzik improvement* [41]. For example, we make use of a *clover improved* [40] action, which involves adding a term to the Wilson action,

$$c_{\mathrm{sw}} a^5 \sum_{n \in \Lambda} \sum_{\mu < \nu} \bar{\psi}(n) \frac{1}{2} \sigma_{\mu\nu} \widehat{F}_{\mu\nu}(n) \psi(n), \tag{2.35}$$

where c_{sw} is a real coefficient, called the SheikholeslamiWohlert coefficient,

$$\widehat{F}_{\mu\nu}(n) = \frac{-\mathrm{i}}{8a^2} \left(Q_{\mu\nu}(n) - Q_{\nu\mu}(n) \right), \tag{2.36}$$

$$\sigma_{\mu\nu} \equiv \left[\gamma_\mu, \gamma_\nu \right] / 2\mathrm{i}, \tag{2.37}$$

and

$$Q_{\mu\nu}(n) \equiv U_{\mu,\nu}(n) + U_{\nu,-\mu}(n) + U_{-\mu,-\nu}(n) + U_{-\nu,\mu}(n). \tag{2.38}$$

The clover term is named as such because it is constructed from four adjacent plaquettes, as depicted in Fig. 2.2.
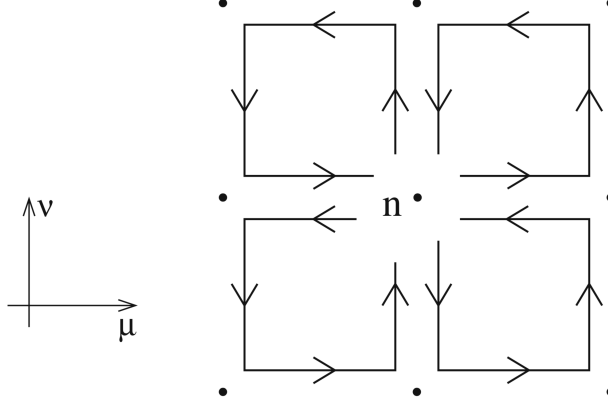
9

Figure 2.2: Depiction of the clover term, built out of four plauqettes. Figure from [15].

In addition to Symanzik improvements, there are so-called *tadpole* improvements [22] that can be made to the gauge action. These seek to correct contributions from QCD tadpole diagrams that appear at low energy, when the coupling is large. For a small coupling g, we may reasonably expand the link variables as $U_\mu(x) \equiv e^{iagA_\mu(x)} \rightarrow 1 + iagA_\mu(x)$. It turns out that, however, that higher order terms are suppressed by powers of g^2, and so when the coupling is large, there will be contributions from QCD tadpoles. For a detailed presentation, see [22]. In order to deal with these tadpoles, we rescale the link variables as $U \rightarrow \frac{U}{u}$, where

$$u = \left\langle \frac{1}{3} \operatorname{Re} \operatorname{Tr} U_{\mu\nu} \right\rangle^{1/4}. \tag{2.39}$$

u is both a parameter of the action and an observable, so we must iteratively tune u in the action such that the u we observe is the same as the u we put into the action.

In practice, our lattices are anisotropic [11, 23]. The motivation for this is that the primary observables we measure are temporal correlators whose signal-to-noise ratios decrease as time separation increases. To combat this issue, we could use a finer lattice in order to sample the correlators at more time separations. The computational cost of refining the lattice in every direction is large, however, so we choose to refine only the temporal direction. That is, we choose an anisotropy $\xi \equiv \frac{a_s}{a_t} > 1$, where a_s is the spatial lattice spacing and a_t is the temporal lattice spacing. ξ is in fact a quantity that undergoes renormalization, so the anisotropy enters into the gauge action as a parameter γ_g and into the fermionic action as a parameter γ_f. Both of these parameters must be tuned to produce the desired renormalized anisotropy.

Finally, we write down the action we use in this work. The gauge action is a Symanzik-improved Lüscher-Weisz action [25, 24], and uses tadpole improvement coefficients from [35,

36, 23, 11]:

$$S_G^\xi[U] = \frac{\beta}{3\gamma_g} \left[\sum_{x, i \neq j} \left(\frac{5}{6u_s^4} \Omega_{\mathcal{P}_{ij}}(x) - \frac{1}{12u_s^6} \Omega_{\mathcal{R}_{ij}}(x) \right) + \sum_{x,i} \left(\frac{4}{3u_s^2 u_t^2} \Omega_{\mathcal{P}_{it}}(x) - \frac{1}{12u_s^4 u_t^2} \Omega_{\mathcal{R}_{it}} \right) \right].$$
(2.40)

Here, P is a plaquette, $\Omega_P = \mathrm{Re}\,\mathrm{Tr}(1 - P)$, R is a 2×1 planar Wilson loop, u_s and u_t are the aforementioned tadpole improvement coefficients, β is the inverse gauge coupling, and γ_g is the bare anisotropy parameter for the gauge sector. Notice that only improvements in spatial links have been included. Lattice gauge-field actions which involve more than two time slices often lead to temporal correlators which do not fall exponentially in a simple way. This action has $\mathcal{O}(a_s^4, a_t^2, g^2 a_s^2)$ discretization errors. The fermionic part of the action is given by

$$S_F^\xi[U, \bar{\psi}, \psi] = \sum_x \overline{\hat{\psi}}(x) \frac{1}{\tilde{u}_t} \left\{ \tilde{u}_t \hat{m}_0 + \gamma_t \hat{W}_t + \frac{1}{\gamma_f} \sum_s \gamma_s \hat{W}_s \right.$$
$$\left. - \frac{1}{2} \left[\frac{1}{2} \left(\frac{\gamma_g}{\gamma_f} + \frac{1}{\xi} \right) \frac{1}{\tilde{u}_t \tilde{u}_s^2} \sum_s \sigma_{ts} \hat{F}_{ts} + \frac{1}{\gamma_f} \frac{1}{\tilde{u}_s^3} \sum_{s<s'} \sigma_{ss'} \hat{F}_{ss'} \right] \right\} \hat{\psi}(x).$$
(2.41)

Here, γ_f is the bare anisotropy for the gauge sector, $\xi = \frac{a_s}{a_t}$ is the renormalized anisotropy, γ_μ and $\sigma_{\mu\nu} = \frac{1}{2}[\gamma_\mu, \gamma_\nu]$ are Dirac gamma matrices, $\tilde{u}_s$ is the spatial tadpole improvement factor, and $\tilde{u}_t$ is the temporal tadpole improvement factor. Hats denote dimensionless quantities that relate to dimensionful quantities by scaling with appropriate factors of the lattice spacing. $\hat{\psi} = a_s^{3/2} \psi$ where ψ is the quark field, $\hat{m}_0 = m_0 a_t$ where m_0 is the bare quark mass, and $\hat{F}_{\mu\nu} = a_\mu a_\nu F_{\mu\nu} = \frac{1}{4} \mathrm{Im}\,(\mathcal{P}_{\mu\nu}(x))$ where $F_{\mu\nu}$ is the field strength tensor. $\hat{W}$ is the *Wilson operator*, and is given by

$$\hat{W}_\mu \equiv \hat{\nabla}_\mu - \frac{1}{2} \gamma_\mu \hat{\Delta}_\mu,$$
(2.42)

where

$$\hat{\nabla}_\mu = a_\mu \nabla_\mu, \; \hat{\Delta}_\mu = a_\mu^2 \Delta_\mu.$$
(2.43)

This action has $\mathcal{O}\left(g^2 a_s, g^2 a_t, a_s^2, a_t^2\right)$ discretization errors. It is also important to mention that the link variables in the fermionic action are *stout-smeared* [34], which we will discuss further in Chapter 3. The use of smeared link variables in the action leads to better chiral behavior.

2.2.6 Parameter Tuning

Anisotropy

The parameters we put into our action are bare parameters. We obtain these parameters by choosing a set of observables, and tuning the bare parameters in our action such that their desired values are produced. For example, we desire a renormalized anisotropy $\xi \approx 3.5$, and must tune the values γ_g and γ_f to obtain this. In order to tune γ_g, we measure the following quantities [20]:

$$
\begin{aligned}
R_{ss}(x, y) &= \frac{W_{ss}(x, y)}{W_{ss}(x + 1, y)}, \\
R_{st}(x, t) &= \frac{W_{st}(x, t)}{W_{st}(x + 1, t)},
\end{aligned}
\tag{2.44}
$$

where $W_{\mu\nu}(x_\mu, x_\nu)$ is the expectation value of a Wilson loop of size x_μ and x_ν and oriented in the μ and ν directions, and then demand that $R_{ss}(x, y) = R_{st}(x, \xi t)$. In order to tune γ_f, we impose a continuum dispersion relation for mesons on the lattice,

$$
a_t^2 E^2(\boldsymbol{p}) = a_t^2 m^2 + \frac{a_s^2 \boldsymbol{p}^2}{\xi^2}.
\tag{2.45}
$$

We find that in order to attain $\xi \approx 3.5$, we set $\gamma_g = 4.3$ and $\gamma_f = 3.4$.

Gauge Coupling and Lattice Spacing

The renormalization scheme we use requires that the coupling and masses in the lattice QCD action vary (or run) with the ultraviolet cutoff such that observables tend to their physical values as the cutoff is removed. The lattice spacing acts as a cutoff by eliminating very small wavelengths from the theory. The cutoff is proportional to the inverse lattice spacing, limiting momenta to the first Brillouin zone. We tune the inverse coupling β in our action to obtain a desired lattice spacing. Determining the lattice spacing using a physical observable is known as *setting the scale*. One way to set the scale is by choosing a particle, say the Ω, and requiring that its mass measured on the lattice should equal its physical mass. That is, we calculate

$$
a_t = \frac{a_t m}{m_{\text{Phys}}},
\tag{2.46}
$$

where $a_t m$ is the quantity we directly measure in a lattice calculation. In our calculations, this yields $a_t \approx 0.034$ fm, giving $a_s \approx 0.12$ fm, and we obtain these values by choosing $\beta = 1.5$.

Quark Masses

As we will discuss further in Chapter 3, we work in a theory of $N_f = 2{+}1$ QCD, meaning there are three flavors of quarks, two of which have degenerate masses. These are the degenerate up and down quarks (collectively referred to as the *light* quarks), and the strange quark. An important flaw of our calculations is that we work in a theory of an unphysically heavy pion. This is necessary for two reasons. First, finite volume corrections to the theory are exponentially suppressed by a factor of $e^{-m_\pi L}$, where m_π is the mass of the pion (in general, it is the mass of the lightest state in the theory) and L is the spatial extent of the lattice. It is necessary that the *correlation length*, $m_\pi L$ is > 1, but in practice, a good rule of thumb is to require $m_\pi L > 4$ or so. Second, inverting Dirac matrices (which we will soon see to be an important step of our calculation) becomes much more computationally difficult at lower pion masses, and increases the odds that the Dirac matrices will become ill-conditioned.

In order to tune the light and strange quark masses, we aim to set the following ratios to their physical values [23]:

$$l_\Omega = \frac{9m_\pi^2}{4m_\Omega^2}, \quad s_\Omega = \frac{9\left(m_K^2 - m_\pi^2\right)}{4m_\Omega^2}, \tag{2.47}$$

where m_K is the mass of the kaon, m_π is the mass of the pion, and m_Ω is the mass of the Ω baryon. These ratios are chosen because first-order chiral perturbation theory tells us that they should be proportional to their associated quark masses [23]. s_Ω is set to its physical value, and l_Ω is lowered to a value that is unphysical but allows calculations to remain possible. Using this method of tuning, the bare quark masses are set to be $a_t m_l = -0.0860$ for the light quarks, and $a_t m_s = -0.0743$ for the strange quark. With these quark masses, our pion mass is approximately 230 MeV.

2.3 Extracting Energies from Two-Point Correlation Functions

In order to do spectroscopy, the main observables we calculate on the lattice are time-ordered two-point correlation functions of the form

$$C(t) = \langle 0| \, \mathcal{O}(t + t_0)\overline{\mathcal{O}}(t_0) \, |0\rangle, \tag{2.48}$$

where $|0\rangle$ is the vacuum state, $\overline{\mathcal{O}}(t_0)$ is an operator that creates a state out of the vacuum at time t_0, and $\mathcal{O}(t + t_0)$ is an operator that annihilates such a state at a later time $t + t_0$.

Let $|n\rangle$ denote the nth energy eigenstate of the theory with energy E_n, ordered such that $E_n < E_{n+1}$, and let us shift the energies such that the vacuum energy $E_0 \equiv 0$. Then, using a spectral decomposition and working in the Heisenberg picture in imaginary time $t \to -it$, we can see

$$
\begin{aligned}
C(t) &= \sum_{n=0}^{\infty} \langle 0 | e^{H(t+t_0)} \mathcal{O}(0) e^{-H(t+t_0)} | n \rangle \langle n | e^{Ht_0} \overline{\mathcal{O}}(0) e^{-Ht_0} | 0 \rangle \\
&= \sum_{n=0}^{\infty} e^{E_0(t+t_0)} \langle 0 | \mathcal{O}(0) | n \rangle e^{-E_n(t+t_0)} e^{E_0 t_0} \langle n | \overline{\mathcal{O}}(0) | 0 \rangle e^{E_n t_0} \quad (2.49) \\
&= \sum_{n=0}^{\infty} \langle 0 | \mathcal{O}(0) | n \rangle \langle n | \overline{\mathcal{O}}(0) | 0 \rangle e^{-E_n t}.
\end{aligned}
$$

Here, we have made the assumption that $t \ll T$, where T is the temporal length of the lattice, which is necessary to prevent temporal wrap-around effects. It is now clear that by calculating $C(t)$ on the lattice, we have access to the energy spectrum of the theory. In principle, with infinite statistics, we could fit $C(t)$ to obtain the entire spectrum. In practice, however, we can only reliably extract the lowest non-zero energy in this sum by fitting $C(t)$ to single- or two-exponential functions. As we will discuss further in Chapter 5, we can circumvent this issue by constructing a *matrix* of correlators

$$
C_{ij}(t) = \langle 0 | \mathcal{O}_i(t+t_0) \overline{\mathcal{O}}_j(t_0) | 0 \rangle \quad (2.50)
$$

and using a variational method to find *principal correlators* $C^{(N)}(t)$ satisfying

$$
C^{(N)}(t) \xrightarrow[t \to \infty]{} A e^{-E_N t}. \quad (2.51)
$$

2.4   Monte Carlo Path Integration

A fundamental equation for calculating observables on the lattice is

$$
\langle \mathcal{O} \rangle_T = \frac{1}{Z_T} \int \mathcal{D}[\psi, \overline{\psi}, U] \mathcal{O}[\psi, \overline{\psi}, U] e^{-S[\psi, \overline{\psi}, U]}, \quad (2.52)
$$

where $\mathcal{O}$ denotes an operator corresponding to a desired observable, T denotes the time extent of the lattice, S denotes the Euclidean action, and Z_T is the partition function, given by

$$
Z_T = \int \mathcal{D}[\psi, \overline{\psi}, U] e^{-S[\psi, \overline{\psi}, U]}. \quad (2.53)
$$

It turns out that $\langle \mathcal{O} \rangle_T = \langle 0| \mathcal{O} |0 \rangle$ only in the $T \to \infty$ limit, but in practice, T is large enough that $\langle \mathcal{O} \rangle_T \approx \langle 0| \mathcal{O} |0 \rangle$. In the above equations, we are integrating over all configurations of the quark fields, antiquark fields, and link variables representing the gauge fields. It is important to note that the quark and antiquark fields are treated as independent variables. The action can be decomposed as

$$S[\psi, \overline{\psi}, U] = \overline{\psi} M[U] \psi + S_G[U], \tag{2.54}$$

where $M[U]$ is the Dirac matrix which is a functional of the gauge links, and the gauge action S_G is a functional only of the link variables. Integration over the quark fields can be readily done analytically, as the integrals reduce to Gaussian integrals. Wick's theorem gives

$$\langle \mathcal{O} \rangle_T = \frac{\int \mathcal{D}[U] f \left(M^{-1}[U] \right) \det M[U] e^{-S_G[U]}}{\int \mathcal{D}[U] \det M[U] e^{-S_G[U]}}, \tag{2.55}$$

where $f \left(M^{-1}[U] \right)$ is some function of the inverse of the Dirac matrix, given by the relevant Wick contractions.

While the fermions are integrated out analytically, it is intractable to analytically integrate out the gauge fields. This is where the need for numerical integration arises. Numerically integrating over the gauge fields is nontrivial, however, since the integral in Eq. (2.55) is of very high dimension due to the fact that we must integrate over a link variable at every single lattice site. The way to proceed is by use of Monte Carlo integration, which is immune to the so-called curse of dimensionality that precludes other standard numerical methods.

The strategy of Monte Carlo integration proceeds as follows. Consider a very high dimensional integral

$$I_f = \int_{\mathcal{V}} \mathcal{D} U p(U) f(U), \tag{2.56}$$

where $p(U)$ is a probability density over the volume $\mathcal{V}$, and $f(U)$ is some integrable function of the variables U. By factoring out a probability density in the integrand, we ensure much better efficiency in Monte Carlo sampling compared to using uniformly random samples of the integrand, since it allows us to focus on sampling points from the integrand that most contribute to the result of the integral. This technique is called *importance sampling*. If we can randomly sample the variables U with probability density $p(U)$ thereby generating a large ensemble $\{U_1, U_2, ..., U_{N_U}\}$ of N_U configuartions, then the integral I_f can be approximated as

$$I_f \approx \frac{1}{N_U} \sum_{k=1}^{N_U} f\left(U_k \right), \tag{2.57}$$

by the law of large numbers. For large N_U and independently generated U_k, then by the central limit theorem, the standard deviation in the approximation of I_f is given by

$$\sigma_{I_f} = \sqrt{\frac{V(f(U))}{N_U}}, \tag{2.58}$$

where $V(f(U))$ is the variance of $f(U)$ with respect to the probability density $p(U)$. This is where the advantage of working in Euclidean spacetime becomes apparent: the real Boltzmann factor $e^{-S_G[U]}$ functions as a probability distribution (up to a normalization constant). It turns out that it is also necessary to include the fermion determinant $\det M$ in the probability distribution, as including it as a part of the observable itself can lead to large statistical uncertainties due to the very large fluctuations about the mean value. [15] Including the Dirac matrix determinant and the normalization factor, the probability distribution for calculating observables is

$$p(U) = \frac{\det M[U]e^{-S_G[U]}}{\int \mathcal{D}[U']\det M[U']e^{-S_G[U']}}. \tag{2.59}$$

The validity of using $p(U)$ as probability distribution is not obvious, as it requires that $\det M$ is real and positive semi-definite. Fortunately, for an even number of mass-degenerate fermions (an approximation we use for the u and d quarks), this can be shown analytically (see Ref. [15], 8.2.1). When we add in a strange quark, there is no such analytic guarantee, but we observe empirically that the mass of the strange quark is high enough that the fermion determinant is always found to be positive.

In order to proceed in generating a set of gauge configurations $\{U_i\}$, we make use of Markov chains. A Markov chain is defined as a stochastic process that generates a sequence of states in which the probability of transitioning from one state to another depends only on the current state of the system [31]. For us, a state is defined by a gauge field configuration. A certain class of Markov chains (those which are *irreducible* and *aperiodic*) have the property that they tend to a *stationary* distribution provided that they satisfy a condition called *detailed balance* defined by

$$T(U' \leftarrow U)p(U) = T(U \leftarrow U')p(U'), \tag{2.60}$$

where $T(U' \leftarrow U)$ is the probability of transitioning to a state U' if the current state is U. In this case, after a certain number of steps along the Markov chain (a process known as *thermalization* or bringing the chain into *equilibrium*), then all future states will be distributed according to $p(U)$.

A popular method of defining such a Markov chain is the Metropolis-Hastings method [28],

16

which relies on an accept-reject method of proposing new configurations by making small local changes to the current configuration. These small changes are necessary because large changes can cause the acceptance probability to be too low. The Metropolis-Hastings method works well in both pure gauge theories and in the quenched approximation where sea-quark effects are neglected and the fermion determinants are set to one. In a full theory, however, such a local updating scheme is not useful, since the fermion determinant is a non-local quantity.

2.4.1 Hybrid Monte Carlo

The need for a global updating scheme with a reasonable acceptance rate is solved by the Hybrid Monte Carlo Method [19] (HMC) algorithm, which applies in the case of an even number of degenerate quarks, as in our case with the u and d quarks (ignore the strange quark for a moment). The HMC proceeds by representing the fermion determinant as an integral over so-called *pseudofermion* fields ϕ and $\phi^\dagger$, which are complex-valued (rather than Grassmann-valued) fields with the same indices as the fermionic fields:

$$\det M[U] = \int \mathcal{D}\phi^\dagger \mathcal{D}\phi \, e^{-\phi^\dagger M^{-1}[U]\phi}. \tag{2.61}$$

The fermion determinant can be factored as

$$\det M^{(u)}[U] \det M^{(d)}[U] = \left(\det M^{(l)}[U]\right)^2 = \int \mathcal{D}\phi^\dagger \mathcal{D}\phi \, \exp\left[\phi^\dagger \left(M^{(l)\dagger}[U]M^{(l)}[U]\right)^{-1} \phi\right], \tag{2.62}$$

where the u quark Dirac matrix $M^{(u)}$ and the d quark Dirac matrix $M^{(d)}$ are equal collectively referred to as the *light* quark Dirac matrix $M^{(l)}$. Here we have made use of γ_5-hermiticity of M, which is general feature of most discretized Dirac matrices [15],

$$\gamma_5 M \gamma_5 = M^\dagger, \tag{2.63}$$

and the fact that $\gamma_5^2 = 1$. Using this factorization of the determinant, we can replace the action in Eq. (2.59) by an effective action,

$$S_{\text{eff}}[U, \phi, \phi^\dagger] = S_G[U] + \phi^\dagger \left(M^{(l)\dagger}[U]M^{(l)}[U]\right)^{-1} \phi, \tag{2.64}$$

where we are now sampling the fields ϕ and $\phi^\dagger$ in addition to U in the Monte Carlo integration. The next step of the HMC is to construct a fictitious Hamiltonian and evolve the system according to Hamilton's equations of motion. This involves introducing a ficti-

tious field π which is viewed as the conjugate momentum field to U. Then, using the clever expression of unity,

$$1 = \int \mathcal{D}\pi \exp\left\{-\frac{1}{2}\pi^\dagger\pi\right\}, \tag{2.65}$$

we insert it into an integral over the gauge fields and pseudofermion fields,

$$
\begin{aligned}
&\int \mathcal{D}\pi \exp\left[-\frac{1}{2}\pi^\dagger\pi\right] \int \mathcal{D}U \mathcal{D}\phi^\dagger \mathcal{D}\phi \exp(-S_{\text{eff}}) \\
&= \int \mathcal{D}\pi \mathcal{D}U \mathcal{D}\phi^\dagger \mathcal{D}\phi \exp\left[-\frac{1}{2}\pi^\dagger\pi - S_{\text{eff}}\right] \\
&= \int \mathcal{D}\pi \mathcal{D}U \mathcal{D}\phi^\dagger \mathcal{D}\phi \exp[-H],
\end{aligned}
\tag{2.66}
$$

defining the fictitious Hamiltonian as

$$H = \frac{1}{2}\pi^\dagger\pi + S_{\text{eff}}[U, \phi, \phi^\dagger]. \tag{2.67}$$

The system is then evolved forward as in a molecular dynamics simulation using Hamilton's equations of motion. In principle, H should be conserved in this process, but in practice, there will be some errors due to a finite time step. We solve this by adding in an accept-reject step at the end of the time evolution by introducing an acceptance probability,

$$P_{\text{accept}} = \min(1, e^{-\delta H}). \tag{2.68}$$

The pseudofermion fields also need to be refreshed, which can be done by producing a normally-distributed vector χ with variance $\frac{1}{2}$ and then calculating $\phi = M^\dagger\chi$.

2.4.2   Rational Hybrid Monte Carlo

Introducing the Hybrid Monte Carlo method involved neglecting the strange quark. The generalization of HMC to include the strange quark is known as the Rational Hybrid Monte Carlo (RHMC) method [9]. Just as before, the fermion determinant for the strange quark field is expressed as an integral over pseudofermion fields:

$$\det M_s = \det\left(M_s^\dagger M_s\right)^{\frac{1}{2}} = \int \mathcal{D}\phi^\dagger \mathcal{D}\phi \exp\left(\phi^\dagger\left(M_s^\dagger M_s\right)^{-1/2}\phi\right). \tag{2.69}$$

This is only valid if $\det M_s \geq 0$. As stated before, this is not analytically guaranteed, but it is generally true due to the large strange quark mass. We proceed just as in the HMC method for the light quarks, but whereas in the HMC method we had one M for

each pseudofermion, we now have $(M^\dagger M)^{\frac{1}{4}}$ applied to each field ϕ. This fourth-root can be estimated with a rational approximation,

$$\left(M^\dagger M\right)^{\frac{1}{4}} \approx \alpha_0 1 + \sum_i \frac{\alpha_i}{M^\dagger M + \beta_i}, \tag{2.70}$$

where α_i and β_i are coefficients that specify the approximation [9]. Refreshing the pseudofermion fields is done similarly as before, by generating a normally-distributed vector χ with variance $\frac{1}{2}$ and calculating $\phi = (M^\dagger M)^{\frac{1}{4}} \chi$.