이학박사학위논문

Taste Non-Goldstone Pion Decay Constants and Beyond the Standard Model B-parameters in Lattice QCD with Staggered Fermions

스태거드 페르미온을 이용한 격자 양자색역학에서 파이온 붕괴 상수와 초표준모형 B 파라미터의 계산

2013년 2월

서울대학교 대학원

물리천문학부 물리학전공

윤 보 람

ABSTRACT

In part I, we calculate the next-to-leading order corrections to pion decay constants for the taste non-Goldstone pions using staggered chiral perturbation theory. This is a generalization of the calculation for the taste Goldstone case. New low-energy couplings are limited to analytic corrections that vanish in the continuum limit; the chiral logarithms contain no new couplings. We report results for quenched, fully dynamical, and partially quenched cases of interest in the chiral SU(3) and SU(2) theories. The results can be used to refine existing determinations of decay constants and low energy constants.

In part II, we calculate the beyond the standard model B-parameters using HYP-smeared improved staggered fermions on the MILC asqtad lattices with $N_f = 2 + 1$ flavors. We use three different lattice spacings ($a \approx 0.045,\ 0.06$ and 0.09 fm) to obtain the continuum results. Operator matching is done using one-loop perturbative matching, and results are run to 2 and 3 GeV in the $\overline{\text{MS}}$ scheme. For the chiral and continuum extrapolations, we use SU(2) staggered chiral perturbation theory. We present preliminary results with only statistical errors.

In part III, we give a detailed introduction to the data anlysis including basic probability theory, error anlalysis techniques and least χ^2 fitting method. We also explain how to analyse highly correlated data by applying a number of prescriptions such as diagonal approximation, singular value decomposition (SVD) method and Bayesian method. We propose a brand new method, the eigenmode shift method which allows a full covariance fitting without modifying the covariance matrix.

Keywords: quantum chromodynamics, lattice QCD, chiral perturbation theory, pion decay constant, beyond the standard model bag parameters

Student ID: 2006-20339

Contents

1. Introduction

1.1. Quantum chromodynamics

The standard model (SM) is a theory that describes the interactions concerning the electromagnetic, weak and strong forces between the elementary particles. It has predicted the existence of the charm and top quarks, W and Z bosons and gluon before they are observed by experiments. It also explains lots of physical phenomena with a good precision, despite the simple structure of theory.

In the SM, strong interaction is described by the quantum chromodynamics (QCD), which is a $SU(3)$ non-Abelian gauge theory of color charged quarks and gluons. Modern concept of quarks, which carry color charge, is proposed by Oscar Greenberg [1] and Moo-Young Han with Yoichiro Nambu [2], independently. There are three color degrees of freedom, blue, red and green, and quarks lie in the fundamental representation of the gauge group $SU(3)$. In the SM, there are six species of quarks, denoted by flavors: up (u), down (d), strange (s), charm (c), bottom (b) and top (t). The strong interactions of quarks are mediated by the gluons, which are the quanta of the $SU(3)$ gauge field.

The dynamics of quarks and gluons are described by the gauge invariant QCD Lagrangian,

$$\mathcal{L} = -\frac{1}{2}\mathrm{Tr}(F_{\mu\nu}F^{\mu\nu}) + \sum_f \bar{q}_f(i\not{D} - m_f)q_f, \tag{1.1}$$

where f runs over the six flavors, q_f is the quark field of flavor f, and m_f is the mass of q_f. The field strength is

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial A_\mu - ig[A_\mu, A_\nu], \tag{1.2}$$

where g is the coupling constant in the theory, and the gluon field A_μ is written as

$$A_\mu = \sum_{a=1}^{8} A_\mu^a \frac{\lambda^a}{2}. \tag{1.3}$$

Here λ^a are the Gell-Mann matrices that satisfy

$$\left[\frac{\lambda_a}{2}, \frac{\lambda_b}{2}\right] = if^{abc}\frac{\lambda_c}{2} \tag{1.4}$$

with structure constants f^{abc} of the $SU(3)$ group, and normalized by

$$\mathrm{Tr}(\lambda^a \lambda^b) = 2\delta^{ab}. \tag{1.5}$$

The covariant derivative is

$$D_\mu = \partial_\mu - igA_\mu. \tag{1.6}$$

The slash notation is defined by $\not{D} = \gamma^\mu D_\mu$, where the gamma matrices γ_μ satisfies

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} \tag{1.7}$$

with metric tensor

$$\eta^{\mu\nu} = \mathrm{diag}(1, -1, -1, -1). \tag{1.8}$$

One of the remarkable features of QCD is asymptotic freedom. *Asymptotic freedom* means that quarks and gluons are interact very weakly at the high energy scale. This is a peculiar feature of non-Abelian gauge theory[1]. Asymptotic freedom in QCD is discovered in 1970s by David Politzer [4] and by David Gross and Frank Wilczek [5]. They won Nobel Prize for the discovery in 2004.

What makes QCD interesting is the fact that the coupling constant $\alpha_s = \frac{g^2}{4\pi}$ becomes too large in low energy region ($\Lambda_{\mathrm{QCD}} \approx 220\mathrm{MeV}$) to use the perturbation theory. Without non-perturbative methods, such as lattice QCD, one can calculate only the high energy physics, in which QCD perturbation is valid ($\alpha_s \approx 0.4$ at 1GeV). Two notable non-perturbative phenomena of QCD is the confinement and spontaneous chiral symmetry breaking.

Confinement means that color charged particles, such as quarks, cannot be isolated as a free particle. As a result, quarks are observed only as a bound state in hadrons, which is a color singlet. Since the confiment is a non-perturbative phenomenon, it is not proven analytically. However, it can be shown in the strong coupling limit, and many lattice simulations also show linearly increasing static potential as the two quarks in a $q\bar{q}$ bound state are separated [6, 7].

For massless n_f flavors, QCD Lagrangian has $SU(n_f)_L \times SU(n_f)_R$ chiral symmetry. The chiral symmetry is *spontaneously* broken to $SU(n_f)_V$ by the QCD vacuum, and the spontaneously broken symmetry generates massless Goldstone bosons. In the real world, the chiral symmetry is broken explicitly by the nonzero quark masses. As a result, the Goldstone bosons acquire masses related to the quark masses. For $n_f = 3$, the spontaneous chiral symmetry breaking generates eight pseudosclar mesons (π, K, η). The lightness of the pseudosclar mesons are explained by the small quark masses, m_u, m_d, m_s. As a

[1] Coleman and Gross proved that every asymptotically free renormalizable field theory in four dimension are non-Abelian gauge theory. [3]

non-perturbative phenomenon, the spontaneous chiral symmetry breaking is not proven. However, it explains many physical results related to the hadron spectrum, and lattice QCD simulations show the nonzero chiral condensation $\langle \bar{q}q \rangle \neq 0$, which indicates the spontaneous chiral symmetry breaking. [8, 9, 10]

1.2. Lattice QCD

Lattice QCD is a non-perturbative approach to understand QCD, formulated on a lattice, which is a discretized version of the Euclidean space-time. At present, lattice QCD is the only method to calculate the nonperturbative properties of the QCD from the first principle with controlled error.

Using lattice, QCD can be regularized preserving the gauge invariance. Usually the lattice QCD uses hypercubic lattice with lattice spacing a, which plays a role as an ultraviolate regulator. Quarks are put on the lattice sites, and the gluons (gauge fields) are placed on the links between the sites. The gauge transformation of the quark fields is

$$q(x) \to V(x)q(x), \qquad \bar{q}(x) \to \bar{q}(x)V^\dagger(x), \tag{1.9}$$

where $V(x)$ is an element of $SU(3)$. The gauge fields $U_\mu(x)$ are introduced as parallel transporters connecting x and $x + a\hat{\mu}$:

$$U_\mu(x) = \mathcal{P} \exp\left[ig \int_0^a A_\mu(x + s\hat{\mu})ds\right] \tag{1.10}$$

$$= 1 + iagA_\mu(x + \frac{a}{2}\hat{\mu}) + \cdots, \tag{1.11}$$

where $\mathcal{P}$ is path-ordering. $U_\mu(x)$ are called the gauge links. Under gauge transformations $V(x)$, it transforms as

$$U_\mu(x) \to V(x)U_\mu(x)V^\dagger(x). \tag{1.12}$$

Using the gauge links, quark bilinears defined over different lattice sites can be written in a gauge invariant form. For example,

$$\bar{q}(x)U_\mu(x)q(x + a\hat{\mu}) \tag{1.13}$$

is gauge invariant. Another important gauge invariant object is the trace of the Wilson loop. The Wilson loop is a product of the gauge links over a closed loop on the lattice. For example, plaquette U_P is the Wilson loop defined on the smallest loop,

$$U_P(x) = U_\mu(x)U_\nu(x + a\hat{\mu})U_\mu^\dagger(x + a\hat{\nu})U_\nu^\dagger(x), \tag{1.14}$$

whose trace is gauge invariant.

Lattice version of QCD action is written on the discrete Euclidean space time with finite lattice spacing a. The integration over the space time is replaced by the summation over the lattice sites,

$$\int d^4x \quad \rightarrow \quad a^4 \sum_x, \tag{1.15}$$

and the lattice action converges to the continuum version of QCD action as $a \rightarrow 0$. The simplest lattice gauge action is the Wilson plaquette action [11],

$$S_g = \frac{\beta}{3} \sum_P \mathrm{ReTr}(1 - U_P), \tag{1.16}$$

where $\beta = 6/g^2$, trace is performed over the color indices, and $\sum_P$ is the summation over the all possible plaquettes of a fixed orientation. For small a, it approaches to the continuum gauge action

$$S_g \quad \rightarrow \quad \int d^4x \frac{1}{2} \mathrm{Tr}(F_{\mu\nu}^2), \tag{1.17}$$

up to $\mathcal{O}(a^2)$. Details of the Wilson plaquette action and improved version of gauge action will be explained in Sec. 2.1.

On the lattice, fermions are introduced as Grassmann fields $q(x)$ and $\bar{q}(x)$, defined on the lattice sites. The direct discretization of the continuum fermion action yields the "naive" lattice fermion action:

$$S_f^{\mathrm{naive}} = a^4 \sum_x \bar{q}(x) \Big[\gamma_\mu \nabla_\mu + m \Big] q(x), \tag{1.18}$$

with the discretized version of the covariant derivative

$$\nabla_\mu q(x) = \frac{1}{2a} \big[U_\mu(x) q(x + a\hat{\mu}) - U_\mu^\dagger(x - a\hat{\mu}) q(x - a\hat{\mu}) \big]. \tag{1.19}$$

However, the "naive" fermion action possesses the infamous *fermion doubling problem*: it describes 16 equivalent fermions in the continuum limit. It gets complicated when one considers the chiral symmetry because the Nielsen-Ninomiya theorem [12] states that one cannot formulate a ultralocal[2] and unitary lattice fermion action that has a correct continuum limit, exact chiral symmetry ($\gamma_5 D + D\gamma_5 = 0$) and no doublers. In order to avoid the fermion doubling problem, various methods have been proposed. They will be explained in Sec. 2.2.

The quantum theory of the lattice QCD is defined by the path integral for-

[2]Here "ultralocal" denotes an action whose coupled fields are only finite number of lattice spacings apart.

mulation. The partition function of lattice QCD is

$$Z = \int \mathcal{D}[\bar{q}, q]\mathcal{D}[U] e^{-S_g[U] - \sum_f \bar{q}_f \left(D[U] + m_f\right) q_f}, \tag{1.20}$$

where f denotes flavors, $\mathcal{D}[\bar{q}, q]$ is the product of measures of all anticommuting Grassman fields

$$\mathcal{D}[\bar{q}, q] = \prod_{x,f} [d\bar{q}_f(x) dq_f(x)], \tag{1.21}$$

and $\mathcal{D}[U]$ is the product of invariant $SU(3)$ Haar measures

$$\mathcal{D}[U] = \prod_{x,\mu} dU_\mu(x). \tag{1.22}$$

The integration over the $\bar{q}$ and q can be performed by hand, which yields

$$Z = \int \mathcal{D}[U] e^{-S_g[U]} \prod_f \det\left(D[U] + m_f\right). \tag{1.23}$$

The expectation value of an observable $\mathcal{O}(U, q, \bar{q})$ is

$$\langle \mathcal{O}(U, q, \bar{q}) \rangle = \frac{1}{Z} \int \mathcal{D}[\bar{q}, q]\mathcal{D}[U] \mathcal{O}(U, q, \bar{q}) e^{-S_g[U] - \sum_f \bar{q}_f \left(D[U] + m_f\right) q_f} \tag{1.24}$$

$$= \frac{1}{Z} \int \mathcal{D}[U] \mathcal{O}(U) e^{-S_g[U]} \prod_f \det\left(D[U] + m_f\right), \tag{1.25}$$

where q and $\bar{q}$ in the observable $\mathcal{O}(U, q, \bar{q})$ are replaced by the quark propagator, $\left(D[U] + m_f\right)^{-1}$, in Eq. (1.25). In the lattice simulation, the integration in Eq. (1.25) is carried out by the Monte Carlo method. The first step is the generation of Markov chain of gauge links $U_\mu^{(i)}(x)$ for $i = 1, 2, 3, \cdots, n$, that are distributed according to the probability distribution

$$p(U^{(i)}) = \frac{1}{Z} e^{-S_g[U^{(i)}]} \prod_f \det\left(D[U^{(i)}] + m_f\right). \tag{1.26}$$

Next step is the *measurement*, which is taking average over the gauge configuration,

$$\langle \mathcal{O}(U, q, \bar{q}) \rangle = \frac{1}{n} \sum_{i=1}^{n} \mathcal{O}\left(U^{(i)}\right). \tag{1.27}$$

Since the expectation value of an observable is estimated by the average over statistical samples, the lattice simulation results come with the statistical errors.

In general, the generation of the Markov chain of gauge links requires huge computational cost. Most expensive part of the generation of gauge links is the evaluation of the determinant, $\det\left(D[U]+m_f\right)$. It is done using Φ algorithm [13] on the hybrid Monte Carlo (HMC) algorithm [14], which combinines the molecular dynamics based (MD) algorithms [15, 16] with the Metropolis accept-reject steps [17].

1.3. Recent progress of the lattice calculation

In this section, we briefly review recent lattice results for the low-energy particle physics.

To control the finite volume effect, the volume of the lattice should stay in a certain size[3]. Thus the number of lattice sites increases as the lattice spacing decreases. As a results, the lattice simulation cannot be performed with arbitrary small lattice spacing, but it is restricted by the computational cost. Recently, the lattice spacing arrived at 0.045fm [6, 18], which corresponds to the momentum cutoff of $\pi/a \approx 14$GeV.

Most of the current lattice simulations use the same masses for the two light quark: up and down ($m_u = m_d$). Many simulations are done with three dynamical quark flavors: up, down and strange, and $m_u = m_d \neq m_s$, which is denoted by the notation $N_f = 2+1$. Recently, it has just begun to use four dynamical quark flavors including the charm quark, $N_f = 2+1+1$, with highly improved staggered quarks (HISQ) [19, 20, 21] and Wilson twisted mass quarks [22, 23, 24, 25].

In the lattice simulations, most computation time is used to obtain the inverse of Dirac operation $(D+m)$, whose computational cost increases rapidly as the quark mass m decreases. Thus, most current simulations are done with the light quark masses m_u, m_d heavier than their physical values; $M_\pi \gtrsim 200$MeV. The extrapolation is controlled by the chiral perturbation theory (ChPT). However, the light quark masses in the lattice simulation are getting lighter, and some of the recent simulations are performed at the physical light quark masses $m_{u,d}^{\rm phys}$ [26, 27].

Here is the summary of the recent main lattice results. Since many different groups calculate the same observables, it is needed to consider an "average" value. The Flavianet Lattice Averaging Group (FLAG) [28] works on the averaging the lattice results. Here we report their results.

The lattice determination of running quark masses in $\overline{\rm MS}$ scheme at 2GeV is

[3] A rule of thumb for the required spatial extent L is $M_\pi L \gtrsim 4$, since the finite volume correction to the hadron masses are roughly proportional to $e^{-M_\pi L}$.

[6, 27, 29, 30, 31, 32, 33, 34, 35, 36, 37]

$$m_s^{\overline{\rm MS}}(2{\rm GeV}) = 94 \pm 3{\rm MeV}, \tag{1.28}$$

$$m_{ud}^{\overline{\rm MS}}(2{\rm GeV}) = 3.43 \pm 0.11{\rm MeV}, \tag{1.29}$$

$$\frac{m_s}{m_{ud}} = 27.4 \pm 0.4, \tag{1.30}$$

where $m_{ud} = (m_u + m_d)/2$.

The form factor for the semileptonic decay $K^0 \to \pi^-$ at zero momentum transfer, $f_+(0)$ can be directly estimated on the lattice [38, 39, 40, 41]:

$$f_+(0) = 0.959 \pm 0.008. \tag{1.31}$$

For f_K/f_π, the direct results from the lattice are [22, 42, 43]:

$$\frac{f_K}{f_\pi} = 1.193 \pm 0.005. \tag{1.32}$$

From the experimental input for $|V_{us} f_+(0)|$ and $\dfrac{|V_{us}| f_K}{|V_{ud}| f_\pi}$, the lattice determination of $f_+(0)$ and f_K/f_π gives a way to test the unitarity of the CKM matrix [28]:

$$|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 1.002 \pm 0.015. \tag{1.33}$$

One another important lattice determination is the B_K parameter of the $K^0 - \bar{K}^0$ mixing [37, 44, 45, 46, 47, 48]:

$$\hat{B}_K = 0.738 \pm 0.020. \tag{1.34}$$

Besides the results qouted in this section, there are lots of researches using the lattice QCD in hadron spectroscopy, QCD thermodynamics, BSM physcis, computational algorithms, and so on.

1.4. Summary of this thesis

This thesis is composed of three parts: (I) Decay constants in staggered chiral perturbation theory, (II) Kaon mixing matrix elements from BSM operators, and (III) Art of data analysis.

1.4.1. Decay constants in staggered chiral perturbation theory

Staggered fermion formulation is a popular method to put quark fields on the lattice. It preserves part of the chiral symmetry such that it is protected from additive mass renormalization but suffers from a lattice artifact: taste symmetry breaking.

Staggered chiral perturbation theory (SChPT) is designed to describe lattice data generated using staggered fermions. SChPT provides the fitting functional form to use for the chiral extrapolation to physical quark masses as well as to remove the lattice artifacts originating from the taste symmetry breaking. It has been successfully applied to many physical observables in lattice QCD such as pion masses, pion decay constants, semileptonic form factors and kaon mixing parameters.

In part I, we calculate the next-to-leading order corrections to pion decay constants for the taste non-Goldstone pions using staggered chiral perturbation theory. This is a generalization of the calculation for the taste Goldstone case [49]. New low-energy couplings are limited to analytic corrections that vanish in the continuum limit; the chiral logarithms contain no new couplings. We report results for quenched, fully dynamical, and partially quenched cases of interest in the chiral SU(3) and SU(2) theories. The general results for the decay constants are given compactly by Eq. (4.13) with Eqs. (4.45) and (4.46); they reduce to those of Ref. [49] in the taste Goldstone sector.

The results can be used to refine existing determinations of decay constants, quark masses and the Gasser-Leutwyler low energy constants. It also can be used to test the unitarity of the CKM matrix by obtaining the $|V_{us}|$ from the ratio f_K/f_π and the semileptonic decay $K \to \pi\ell\nu$ [28, 50]. f_π in the lattice also can be used to determine the lattice spacing [51].

1.4.2. Kaon mixing matrix elements from BSM operators

In the standard model (SM), the indirect CP violation in the neutral kaon system is described by a single four-fermion operator, and its contribution to ε_K is parametrized by B_K. Recently, the SWME collaboration reported that the theoretical estimate of ε_K in the SM has a 3σ tension with the experimental value in the exclusive V_{cb} channel [52]. Hence, the contribution from the beyond the standard model (BSM) physics draws more attention at present.

There are four additional four-fermion operators which come from the BSM physics. These corrections can be parametrized into B_i $(i = 2, 3, 4, 5)$ parameters. Since the mixing of neutral kaon is well known in the experiments, calculating the BSM B-parameters B_i can give strong constraints to the BSM physics.

In part II, we calculate the beyond the standard model B-parameters using HYP-smeared improved staggered fermions on the MILC asqtad lattices with $N_f = 2 + 1$ flavors. We use three different lattice spacings ($a \approx 0.045,\ 0.06$ and 0.09 fm) to obtain the continuum results. Operator matching is done using one-loop perturbative matching, and results are run to 2 and 3 GeV in the $\overline{\text{MS}}$ scheme. For the chiral and continuum extrapolations, we use SU(2) staggered chiral perturbation theory. We present preliminary results with only statistical errors.

	2GeV	3GeV
B_K	0.5383(66)	0.5199(64)
B_2	0.6245(30)	0.5524(26)
B_3	0.5032(22)	0.4174(19)
B_4	1.0698(56)	1.0222(54)
B_5	0.8432(48)	0.8450(47)

1.4.3. Art of data analysis

Least χ^2 fitting is a popular method to determine unknown parameters in a theory from the numerical data such as lattice QCD results. One caveat is that the least χ^2 fitting works only if the fitting function is precise enough. This requirement becomes a troublesome problem if there is high correlation between data points because the high correlation generates small eigenvalues in the covariance matrix [53, 54]. When some eigenvalues of a covariance matrix are small, even a tiny error of fitting function can produce large χ^2 and spoil the fitting procedure. In some cases, the fitting function cannot be precise enough due to theoretical uncertainty such as truncated higher order terms in a perturbative series expansion.

In part III, we give a detailed introduction to the data anlysis including basic probability theory, error anlalysis techniques and least χ^2 fitting method. We also explain how to analyse highly correlated data by applying a number of prescriptions such as diagonal approximation, singular value decomposition (SVD) method and Bayesian method. We proposed a brand new method, the eigenmode shift (ES) method which allows a full covariance fitting without modifying the covariance matrix.

The basic idea of the ES method is following. The trouble comes from the

inexact fitting function that has small error in the direction of eigenmodes corresponding to the small eigenvalues. Thus, ES method tunes the fitting function $f(x)$ to $f'(x)$, in the troublesome eigenmodes v_k direction,

$$f'(x) = f(x) + \sum_k \eta_k v_k, \tag{1.35}$$

with small tuning parameters η_k constrained by the Bayesian method. If η_k are not constrained, the ES method is the same as the SVD mehtod. However, the ES method is rigorously based on the Bayesian method and is subject to the probability interpretation, while the SVD method is not. In addition, in the ES method, we modify the fitting function by the shift parameter η, which we can monitor and give us an estimate of how much we are changing the fitting function.

2. QCD on the Lattice

Lattice QCD is a non-perturbative approach to understand QCD, formulated on a lattice, which is a discretized version of the Euclidean space-time. In order to formulate QCD on the lattice, we define discretized version of derivatives and replace integrals by summations. We restrict the lattice version of QCD actions to have a exact gauge symmetry and converge to the continuum QCD action in $a \to 0$ limit, so that they mimic continuum QCD. In this chapter, we discuss the lattice actions for the gluons and fermions.

2.1. Gluons on the lattice

As discussed in Sec. 1.2, the simplest lattice gauge action is the Wilson plaquette action [11],

$$S_g = \frac{\beta}{3} \sum_P \mathrm{ReTr}(1 - U_P), \tag{2.1}$$

where $\beta = 6/g^2$, plaquette U_P is defined in Eq. (1.14), trace is performed over the color indices, and $\sum_P$ is the summation over the all possible plaquettes of a fixed orientation. For small a, it approaches to the continuum gauge action

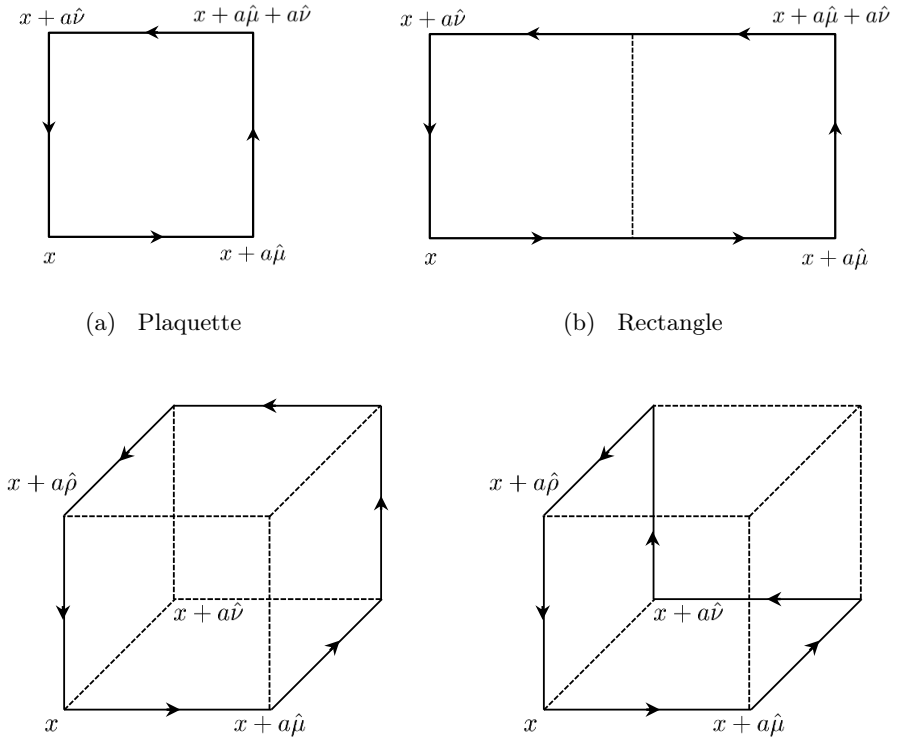
$$S_g \quad \to \quad \int d^4x \frac{1}{2} \mathrm{Tr}(F_{\mu\nu}^2), \tag{2.2}$$

up to $\mathcal{O}(a^2)$; i.e. the Wilson plaquette action has $\mathcal{O}(a^2)$ discretization error.

There is a systematic way to reduce the discretization error in the lattice action – the *Symanzik improvement program* [55, 56], which removes the discretization error order by order perturbatively. Using the Symanzik improvement program, the discretization error in Wilson plaquetten action can be reduced. Adding $6a$-perimeter loops to the $4a$-perimeter plaquette, the gluon action can be written in the following form,

$$S_g^{\mathrm{Imp}} = \frac{\beta}{3} \Bigg[c_{\mathrm{pl}} \sum_{\mathrm{pl}} \mathrm{ReTr}(1 - U_{\mathrm{pl}}) + c_{\mathrm{rt}} \sum_{\mathrm{rt}} \mathrm{ReTr}(1 - U_{\mathrm{rt}}) \tag{2.3}$$

$$+ c_{\mathrm{pg}} \sum_{\mathrm{pg}} \mathrm{ReTr}(1 - U_{\mathrm{pg}}) + c_{\mathrm{ch}} \sum_{\mathrm{ch}} \mathrm{ReTr}(1 - U_{\mathrm{ch}}) \Bigg], \tag{2.4}$$



(a) Plaquette (b) Rectangle

(c) Parallelogram (d) Chair

Figure 2.1: $4a$ and $6a$-perimeter Wilson loops

where pl, rt, pg and ch stand for plaquette, rectangle, parallelogram and chair, respectively. Each of them represents Wilson loops whose shape is shown in Fig. 2.1.

To converge to a proper continuum limit, the coefficients should satisfy the following relation,

$$c_{\rm pl} + 8c_{\rm rt} + 8c_{\rm pg} + 16c_{\rm ch} = 1. \tag{2.5}$$

At one-loop order, the following combination of the coefficient cancel the leading

discretization error in the Wilson plaquette action,

$$c_{\rm pl} = \frac{5}{3} + 0.2370g^2, \tag{2.6}$$

$$c_{\rm rt} = -\frac{1}{12} - 0.02521g^2, \tag{2.7}$$

$$c_{\rm pg} = -0.00441g^2, \tag{2.8}$$

$$c_{\rm ch} = 0, \tag{2.9}$$

which are found by Lüscher and Weisz [57]. Usually, the improved lattice gluon action with Eq. (2.9) is called the Lüscher and Weisz action.

2.2. Fermions on the lattice

QCD action is composed of two parts: one is the gluon action, and the other is the fermion action. To formulate QCD on the lattice, we need discretized version of QCD actions. The gluon action can be easily discretized as discussed in Sec. 2.1, but the fermion action is not easy to be discretized. As introduced in Sec. 1.2, the naive way of discretization yields the infamous fermion doubling problem. Hence, we need a special treatment in order to place fermions on the lattice. In this section, we describe the fermion doubling problem, and discuss about the lattice fermion actions.

2.2.1. Fermion doubling

The direct discretization of the continuum fermion action gives the "naive" lattice fermion action:

$$S_f^{\rm naive} = a^4 \sum_x \bar{q}(x)\Big[\gamma_\mu \nabla_\mu + m\Big] q(x), \tag{2.10}$$

with the discretized version of the covariant derivative

$$\nabla_\mu q(x) = \frac{1}{2a}\big[U_\mu(x) q(x + a\hat{\mu}) - U_\mu^\dagger(x - a\hat{\mu}) q(x - a\hat{\mu})\big]. \tag{2.11}$$

The equation of motion for the naive fermion action is

$$\left(\nabla_\mu \gamma^\mu + m\right)\psi = 0. \tag{2.12}$$

If $\psi(x)$ is a solution of Eq. (2.12),

$$\begin{aligned}\psi'(x) &= i\gamma_5\gamma_\rho(-1)^{x_\rho/a}\psi(x) \\ &= i\gamma_5\gamma_\rho e^{ix_\rho\pi/a}\psi(x),\end{aligned} \tag{2.13}$$

is also a solution, which can be easily shown using

$$\nabla_\mu \gamma^\mu \left(\gamma_5 \gamma_\rho (-1)^{x_\rho/a} \right) = \gamma_5 \gamma_\rho (-1)^{x_\rho/a} \nabla_\mu \gamma^\mu. \tag{2.14}$$

It is more intuitive in the momentum space,

$$\psi(x) = \int \frac{d^4p}{(2\pi)^4} \tilde{\psi}(p) e^{ip\cdot x}, \tag{2.15}$$

$$\begin{aligned} \psi'(x) &= i\gamma_5\gamma_\rho \int \frac{d^4p}{(2\pi)^4} \tilde{\psi}(p) e^{-i\pi \frac{x_\rho}{a}} e^{ip\cdot x} \\ &= i\gamma_5\gamma_\rho \int \frac{d^4p}{(2\pi)^4} \tilde{\psi}(p) e^{ip\cdot x - i\pi \frac{x_\rho}{a}}. \end{aligned} \tag{2.16}$$

Comparing above two equations, one would find that the low momentum mode of $\psi(x)$ is equivalent to $\psi'(x)$ having momentum near $p_\rho \simeq \pi/a$. This is called the "doubler".

In general, we can define a transformation of $\psi(x)$ that generates another solution of the equation of motion, as follows:

$$\psi(x) \to \psi'(x) = \prod_\rho (i\gamma_5\gamma_\rho)^{\xi_\rho} e^{ix\cdot \frac{\xi\pi}{a}} \psi(x), \tag{2.17}$$

where ξ is a vector with one or more components are 1 and others are 0. In four dimension, there are 16 choice of ξ (i.e. $\xi = (1,0,0,0)$), and 15 additional solutions of the equation of motion. In other words, there are 15 doublers in the naive fermion action in four dimension. This is the *fermion doubling problem*: the naive fermion action describes 16 equivalent fermions in the continuum limit.

There is another way of looking the fermion doubling problem. Let us consider the quark propagator of the naive fermion action,

$$D(p) = \frac{-i \sum_\mu \gamma_\mu \left[\frac{1}{a} \sin(p_\mu a) \right] + m}{\sum_\nu \left[\frac{1}{a} \sin(p_\nu a) \right]^2 + m^2}. \tag{2.18}$$

As $\frac{1}{a}\sin(p_\mu a)$ converges p_μ in the limit of $a \to 0$, the propagator has proper continuum limit. Problem is the pole structure of the propagator: if p gives a pole of the propagator Eq. (2.18), then $p' = p + \pi\xi$ also gives a pole, where ξ is a vector with one or more components are 1 and others are 0. As a results, the naive fermion action describes 16 fermions, and it is called the fermion doubling problem.

The fermion doubling problem gets complicated when one considers the chiral symmetry because the Nielsen-Ninomiya theorem [12] states that one can-

not formulate a ultralocal[1] and unitary lattice fermion action that has a correct continuum limit, exact chiral symmetry ($\gamma_5 D + D\gamma_5 = 0$) and no doublers. In order to avoid the fermion doubling problem, various methods have been proposed.

2.2.2. Wilson fermions

One simple prescription to the fermion doubling problem is adding dimension five terms to the naive fermion action $S_f^{\rm naive}$. One of the dimension five terms is the Wilson term,

$$S_f^W = -\frac{ar}{2}\sum_x \bar{\psi}(x)\Box\psi(x)\,, \tag{2.19}$$

where $\Box$ corresponds to the lattice version of covariant second derivative, defined by

$$\Box\psi(x) \equiv \frac{1}{a^2}\sum_\mu \Big[U_\mu(x)\psi(x+a\hat{\mu}) + U_\mu^\dagger(x-a\hat{\mu})\psi(x-a\hat{\mu}) - 2\psi(x)\Big]\,. \tag{2.20}$$

As the Wilson term is a irrelevant operator suppressed by a, the Wilson fermion action $S_f^{\rm naive} + S_f^W$ has proper continuum limit. Unlike the naive fermion action, the Wilson fermion action has no doublers. Let us see the quark propagator of the Wilson fermion action,

$$\frac{1}{a}D(p) = \frac{-i\sum_\mu \gamma_\mu \sin(p_\mu a) + [ma + 2r\sum_\nu \sin^2(ap_\nu/2)]}{\sum_\nu \sin^2(p_\nu a) + [ma + 2r\sum_\nu \sin^2(ap_\nu/2)]^2}\,. \tag{2.21}$$

For small a, the Wilson term dose not affect the propagator when the momentum is small, $p \simeq (0,0,0,0)$. When one of the momentum component in p approaches to $\frac{\pi}{a}$, however, the Wilson term makes the effective mass of the fermion in the propagator heavy. In this way, it removes the doubler modes.

Unfortunately, the Wilson fermion action suffers from the explicit chiral symmetry breaking. In the chiral limit ($m \to 0$), QCD Lagrangian is invariant under the following chiral transformation

$$\psi \to e^{i\gamma_5\theta}\psi\,, \qquad \bar{\psi} \to \bar{\psi}e^{i\gamma_5\theta}\,. \tag{2.22}$$

This chiral symmetry plays an important role in phenomenology of QCD. The Wilson term, however, dose not invariant under this chiral transformation. As a result, the Wilson fermion action has problems including additive mass renormalization of quarks.

[1] Here "ultralocal" denotes an action whose coupled fields are only finite number of lattice spacings apart.

2.2.3. Staggered fermions

Another prescription to the fermion doubling problem is the Staggered fermion formalism proposed by Kogut and Susskind [58]. Let us introduce a new single component fermion fields χ and $\bar{\chi}$,

$$\psi(x) = \Gamma(x)\chi(x)\,, \quad \bar{\psi}(x) = \bar{\chi}(x)\Gamma^{\dagger}(x)\,, \tag{2.23}$$

where

$$\Gamma(x) = \gamma_1^{(x_1/a)}\,\gamma_2^{(x_2/a)}\,\gamma_3^{(x_3/a)}\,\gamma_4^{(x_4/a)}\,. \tag{2.24}$$

This procedure is called the *spin diagonalization*. Using the following relations,

$$\Gamma^{\dagger}(x)\Gamma(x) = \mathbb{1}\,, \tag{2.25}$$

$$\Gamma^{\dagger}(x)\gamma_\mu\Gamma(x+a\mu) = \eta_\mu(x)\mathbb{1}\,, \tag{2.26}$$

$$\eta_\mu(x) \equiv (-1)^{(x_1+\cdots+x_{\mu-1})/a}\,, \tag{2.27}$$

one can rewrite the naive fermion action in Eq. (2.10) as

$$S_f^{\rm stag} = \sum_x \bar{\chi}(x)\left[\sum_\mu \eta_\mu(x)\,\nabla_\mu\, + m\right]\chi(x)\,. \tag{2.28}$$

This is the staggered fermion action. Note that the spin diagonalization allows us to remove three out of four Dirac components and write the fermion action in terms of single component fermions fields. As a result, the doubling is reduced from 16 to 4.

Staggered fermion has the chiral symmetry so that it is protected from additive mass renormalization. However, there still remain the fermion doubling problem. The remaining four fermion degrees of freedom is called the *tastes*. If the fermions of different tastes are independent, the action is invariant under a rotation in the taste space. It is called the taste symmetry. As described in Sec. 2.2.1, however, the fermion of a taste can transform into the fermion of another taste through the large momentum exchange near the lattice cutoff $k \simeq \pi/a$. As a result, the taste symmetry is broken in finite lattice spacing.

In the staggered quark formalism, there are two sources of the $\mathcal{O}(a^2)$ discretization errors: (1) discrete derivative and (2) taste symmetry breaking generated by taste exchange interactions. The former can be reduced by adding the Naik term to the discrete derivative [59], and the later can be reduced by fattening (also called the smearing or the blocking) the gauge links. Gauge link smearing replaces a gauge link (thin link) by a smeared link (fat link), which is an average over the near gauge links in a gauge invariant way. Taking average over paths, it suppresses high momentum transfers, and reduces the taste symmetry breaking.

One of the fattening methods is the Fat7 smearing [60]. AsqTad action [61],

which is an $\mathcal{O}(a^2)$ and tadpole improved [62] staggered fermion action, uses the Fat7 smearing. Highly improved staggered quark (HISQ) also uses the Fat7 smearing. In order to reduce the taste symmetry breaking effect, HISQ uses the Fat7 smearing twice, and projects the first smeared links to $U(3)$ group. The unitary projection also gives the same improvement effect as the tadpole improvement.

Another famous smearing method is the HYP smearing [63]. It constructs the fat link using the links within the hypercubes attached to the original thin link, so that it preserves the short distance properties. It also uses unitary projection on each steps of the smearing. HYP-smeared staggered fermions has smaller taste symmetry breaking error than the AsqTad action [64]. The size of taste symmetry breaking in HYP-smeared staggered fermions is similar to that in the HISQ [65].

Part I

Decay Constants in Staggered Chiral Perturbation Theory

3. Chiral Perturbation Theory

3.1. Introduction to chiral perturbation theory

Although QCD is a good candidate of the fundamental theory describing strong interactions, it is very hard to calculate the low energy dynamics of QCD from the first principle. Based on the symmetries of QCD, chiral perturbation theory provides us a systematic way to understand the low energy properties of QCD, such as masses and decay constants of hadrons, as an effective field theory.

The QCD Lagrangian in Euclidean space is

$$\mathcal{L} = \sum_f \bar{q}_f (i\not{D} + m_f) q_f + \frac{1}{2}\mathrm{Tr}(F_{\mu\nu}F^{\mu\nu})\,, \tag{3.1}$$

where f runs over the six flavors, q_f is the quark field of flavor f, m_f is the mass of q_f, and $F_{\mu\nu}$ is field strength of the gauge fields. Besides the local gauge symmetry, the QCD Lagrangian is also invariant under the global rotation for the flavor index $\mathrm{U}(n_f)_L \times \mathrm{U}(n_f)_R$ in the chiral limit ($m_f \to 0$). Afterwards, we consider only the three light quarks u, d and s, which are lighter than the hadronic scale 1 GeV, in order to account the massless quark limit. In quantum level, however, $\mathrm{U}(1)_A$ of the $\mathrm{U}(3)_L \times \mathrm{U}(3)_R$ is not a symmetry [66, 67, 68]. Thus, the global symmetry of QCD in chiral limit is $\mathrm{SU}(3)_L \times \mathrm{SU}(3)_R \times \mathrm{U}(1)_V$.

The symmetry of $\mathrm{U}(1)_V$ is realized in nature as a baryon number conservation. For the low energy hadron spectrum, however, only the $\mathrm{SU}(3)_V$ symmetry is realized, approximately. It can be considered as a consequence of the *spontaneous symmetry breaking*. The ground state of the system may not invariant under the full symmetry of Hamiltonian. Such symmetries are said to be spontaneously broken.

According to the Goldstone's theorem, if a continuous symmetry is broken spontaneously, there are massless particles called the Goldstone bosons, and the number of Goldstone bosons is the same as the number of generators of the broken symmetry [69, 70]. If we assume the spontaneous symmetry breaking

$$\mathrm{SU}(3)_L \times \mathrm{SU}(3)_R \;\to\; \mathrm{SU}(3)_V\,, \tag{3.2}$$

there should be eight massless particles corresponding to the $\mathrm{SU}(3)_A$ symmetry. The octet of the pseudoscalar mesons (π, K, η) are the good candidates of the Goldstone bosons. Their masses are very small in comparison with the 1^-

vector mesons, and the finite masses can be considered as a consequence of the explicit symmetry breaking due to nonzero quark masses.

As a non-perturbative phenomenon, the spontaneous chiral symmetry breaking is not proven. However, it explains many physical results related to the hadron spectrum. Furthermore, lattice QCD simulations show the non-vanishing scalar quark condensation $\langle \bar{q}q \rangle \neq 0$ in the chiral limit [8, 9, 10], which is a sufficient condition for the spontaneous chiral symmetry breaking in QCD [71].

3.1.1. Chiral Effective Lagrangian

Here we will discuss about the effective chiral Lagrangian for mesons. At low-energy, only the pseudoscalar octet (π, K, η) will show up. The chiral effective Lagrangian is written in terms of the octet members, and it should have the $\mathrm{SU}(3)_L \times \mathrm{SU}(3)_R$ symmetry in the chiral limit. It also should mimic the QCD spontaneous symmetry breaking $\mathrm{SU}(3)_L \times \mathrm{SU}(3)_R \to \mathrm{SU}(3)_V$.

Let us consider a exponential representation,

$$U(x) = \exp\left(i\frac{2\phi(x)}{f}\right), \tag{3.3}$$

$$\phi(x) = \sum_{a=1}^{8} \lambda_a \phi_a(x) \equiv \begin{pmatrix} \frac{1}{\sqrt{2}}\pi_0 + \frac{1}{\sqrt{6}}\eta & \pi^+ & K^+ \\ \pi^- & -\frac{1}{\sqrt{2}}\pi_0 + \frac{1}{\sqrt{6}}\eta & K^0 \\ K^- & \overline{K^0} & -\frac{2}{\sqrt{6}}\eta \end{pmatrix}, \tag{3.4}$$

where f is a constant with mass dimension, which will be the tree-level decay constant, and λ_a are the SU(3) generators. The SU(3) transformation rule of $U(x)$ fields is

$$U(x) \to RU(x)L^\dagger, \tag{3.5}$$

where $L, R \in \mathrm{SU}(3)$. This choice of $U(x)$ realizes the spontaneous symmetry breaking: the ground state $U_0 = 1$ $(\phi(x) = 0)$ is invariant under the vector transformation, but it is not invariant under the axial transformation.

Using this building block, we will expand the Lagrangian in powers of the momentum, because we are concerning only the low momentum (low energy) physics. The leading order Lagrangian invariant under the chiral symmetry is

$$\frac{f^2}{8}\mathrm{Tr}(\partial_\mu U \partial^\mu U^\dagger). \tag{3.6}$$

Let us consider the mass term. In QCD, the mass term is

$$\bar{q}_R M q_L + \bar{q}_L M^\dagger q_R, \tag{3.7}$$

where the mass matrix is

$$M = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_d & 0 \\ 0 & 0 & m_s \end{pmatrix} \tag{3.8}$$

The mass term explicitly breaks the chiral symmetry. To implement the chiral symmetry breaking term into the chiral effective Lagrangian, let us hypothetically "promote" the mass matrix to a field [72]. If the mass matrix transforms under the chiral transformation as

$$M \to RML^\dagger\,, \tag{3.9}$$

then the mass term Eq. (3.7) would be invariant. We find the mass term of the chiral effective Lagrangian that invariant under the Eq. (3.9). The lowest order mass term is

$$\frac{1}{4}\mu f^2 \mathrm{Tr}(MU^\dagger + UM^\dagger)\,, \tag{3.10}$$

where μ is a constant with mass dimension, describing the chiral quark condensate. This mass term precisely mimics the explicit chiral symmetry breaking of the QCD Lagrangian. Since the squared masses of pions are proportional to the quark masses, we take the order of the quark mass as $\mathcal{O}(p^2)$.

Combining Eq. (3.6) and (3.10), we obtain the leading order chiral effective Lagrangian at order $\mathcal{O}(p^2)$. The $\mathcal{O}(p^4)$ Lagrangian can be obtained by applying the above construction to next-to-leading order. For details, see Ref. [73].

3.2. Staggered chiral perturbation theory

In the lattice simulations, most computing time is used to obtain the inverse of the Dirac operation $(D+m)$, more precisely, the solution x of the Dirac operation $(D+m)x = h$. Since $(D+m)$ is a huge and sparse matrix, the conjugate gradient (CG) algorithm (see Sec. 11.2) or similar iterative methods are used to obtain the solution. The computational cost of these matrix inversion algorithms is determined by the condition number of the matrix, where the condition number for a normal matrix[1] is the ratio of the largest and the smallest eigenvalues. Since the smallest eigenvalue of $(D+m)$ is about m, the computational cost of the lattice simulations increases rapidly as the quark mass decreases. As a result, most of the current lattice simulations are done with light (u and d) quark masses larger than their physical values. The results should be extrapolated to the physical quark masses, and it is done by the chiral perturbation theory.

[1] A complex square matrix M is normal if $M^\dagger M = MM^\dagger$.

Staggered chiral perturbation theory (SChPT) is the effective field theory that describes the chiral limit of lattice QCD with staggered fermions. It is designed to describe lattice data generated using staggered fermions, which have an exact chiral symmetry at nonzero lattice spacing. SChPT provides the fitting functional form to use for the chiral extrapolation to physical quark masses as well as to remove the lattice artifacts originating from the taste symmetry breaking. It has been successfully applied to many physical observables in lattice QCD such as pion masses, pion decay constants, semileptonic form factors, and kaon mixing parameters [6, 44, 46, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89].

SChPT was first developed by Lee and Sharpe for single flavor at $\mathcal{O}(a^2)$ [90] and generalized by Aubin and Bernard for multiple flavors [49, 91]. SChPT was extended to next order by Sharpe and Van de Water [92].

3.2.1. Chiral Lagrangian for staggered quarks

In this subsection, we briefly review the chiral Lagrangian for staggered quarks at $\mathcal{O}(a^2)$. The first step is determining the Symanzik effective action for lattice QCD with staggered quarks at low-energy limit ($p \ll 1/a$). Integrating out the high momentum modes, the discretization effects are described by higher dimension operators suppressed by the powers of a.

The Symanzik effective action has the form

$$S = \int d^4x (\mathcal{L}_4 + a^2 \mathcal{L}_6 + \cdots), \tag{3.11}$$

where $\mathcal{L}_4$ is the continuum QCD Lagrangian and $\mathcal{L}_6$ is the sum of dimension 6 operators that satisfy the lattice symmetries. Note that there is no dimension 5 operators (or $\mathcal{L}_5$) that satisfy all of the lattice symmetries (mostly killed by the axial symmetry).

There are three kinds of operators that contribute to $\mathcal{L}_6$: gluonic operators, fermion bilinear operators and four-fermion operators [90]. The gluonic operators are

$$\sum_{\mu\nu\rho} \mathrm{Tr}\left[D_\mu F_{\mu\nu} D_\rho F_{\rho\nu}\right] \tag{3.12}$$

$$\sum_{\mu\nu\rho} \mathrm{Tr}\left[D_\mu F_{\nu\rho} D_\mu F_{\nu\rho}\right] \tag{3.13}$$

$$\sum_{\mu\nu} \mathrm{Tr}\left[D_\mu F_{\mu\nu} D_\mu F_{\mu\nu}\right], \tag{3.14}$$

and the fermion bilinear operators are

$$\bar{q}(\not{D})^3 q \tag{3.15}$$

$$\sum_\mu \bar{q}\left(D_\mu^2 \not{D} + \not{D} D_\mu^2\right) q \tag{3.16}$$

$$\sum_\mu \bar{q} D_\mu \not{D} D_\mu q \tag{3.17}$$

$$\sum_\mu \bar{q}(\gamma_\mu \otimes 1) D_\mu^3 q \tag{3.18}$$

$$m\bar{q}(\not{D})^2 q \tag{3.19}$$

$$\sum_\mu m\bar{q} D_\mu^2 q \tag{3.20}$$

$$m^2 \bar{q} \not{D} q \tag{3.21}$$

$$m^3 \bar{q} q. \tag{3.22}$$

Most of the gluonic operators and the fermion bilinear operators are observed by field redefinition, renormalization of the coupling g and redefinition of the mass. The others yield operators in the chiral Lagrangian at higher order.

Taste symmetry is broken by only the four-fermion operators [90]. The generic form of the four-fermion operators is

$$O(s,t,s',t') = \bar{q}_i(\gamma_s \otimes \xi_t) q_i \bar{q}_j (\gamma_{s'} \otimes \xi_{t'}) q_j, \tag{3.23}$$

where i and j are flavor indices, s denotes spin and t denotes taste. Constrained by the lattice symmetries, $\xi_t = \xi_{t'}$ and $\gamma_s = \gamma_{s'}$ [6]. Thus we can specify an operator using the notation [Spin $\times$ Taste].

There are two choices that construct the four-fermion operators: (a) the spin and taste indices are separately sumed over, and (b) the spin and taste indices are coupled. The former are called type-A operators and the later are called type-B operators. The type-A operators are

$$\begin{array}{llll} [S\times A], & [S\times V], & [A\times S], & [V\times S], \\ [P\times V], & [P\times A], & [V\times P], & [A\times P], \\ [T\times V], & [T\times A], & [V\times T], & [A\times T], \end{array} \tag{3.24}$$

and the type-B operators are

$$[T_\mu \times V_\mu], \quad [T_\mu \times A_\mu], \quad [V_\mu \times T_\mu], \quad [A_\mu \times T_\mu]. \tag{3.25}$$

Here S denotes scalar (1), P denotes pseudosclar (γ_5), V denotes vector (γ_μ), A denotes axial vector ($\gamma_{\mu 5}$) and T denotes tensor ($\gamma_{\mu\nu} \equiv \gamma_\mu \gamma_\nu$). An example

of the type-A operators is

$$[V \times P] = \sum_{\mu} \bar{q}_i(\gamma_\mu \otimes \xi_5) q_i \bar{q}_j(\gamma_\mu \otimes \xi_5) q_j, \tag{3.26}$$

and an example of type-B operators is

$$[V_\mu \times T_\mu] = \sum_{\mu<\nu} \Big[\bar{q}_i(\gamma_\mu \otimes \xi_{\mu\nu}) q_i \bar{q}_j(\gamma_\mu \otimes \xi_{\nu\mu}) q_j - \bar{q}_i(\gamma_\mu \otimes \xi_{\mu\nu}\xi_5) q_i \bar{q}_j(\gamma_\mu \otimes \xi_5\xi_{\nu\mu}) q_j \Big] \tag{3.27}$$

Note that the type-B operators are not invariant for SO(4) rotation of spin and taste, as the μ index is repeated 4 times (See Eq. (3.27)). They are invariant only under 90° rotations.

Next step is finding the chiral Lagrangian for the QCD with staggered quarks. Before writing down the chiral Lagrangian, we need to specify the power counting scheme. Considering $N_f = 3$ flavors, the full chiral symmetry is $SU(12)_L \times SU(12)_R$. Nonzero quark mass breaks the chiral symmetry and gives masses to the Goldstone pions. Including lattice artifacts (taste symmetry breaking, parameterized by the four fermion operators), the chiral symmetry is also broken by the a^2 terms and it shifts the pion masses. According to the data analysis in the MILC coarse lattice ($a \approx 0.12$fm) with asqtad staggered quarks, the taste-symmetry breaking effects in the pion mass spliting is the order of the pion mass, which is proportional to the quark mass [6]. As a result, the standard power counting is

$$\mathcal{O}(p^2/\Lambda_\chi^2) \approx \mathcal{O}(m_q/\Lambda_\chi) \approx \mathcal{O}(a^2\Lambda_\chi^2)\,. \tag{3.28}$$

On finer lattice or with improved staggered quarks, the a^2 could be smaller than the quark mass or the pion mass.

Here, we consider the 4+4+4 theory, in which there are three flavors and four tastes per flavor. The exponential parameterization of the PGB fields is a 12×12 matrix,

$$\Sigma = \exp\left(i\frac{\phi}{f}\right), \tag{3.29}$$

where the PGB fields are

$$\phi = \sum_a \phi^a \otimes T^a, \tag{3.30}$$

$$\phi^a = \begin{pmatrix} U_a & \pi_a^+ & K_a^+ \\ \pi_a^- & D_a & K_a^0 \\ K_a^- & \bar{K}_a^0 & S_a \end{pmatrix}, \tag{3.31}$$

$$T^a \in \{\xi_5,\ i\xi_{\mu 5},\ i\xi_{\mu\nu}(\mu<\nu),\ \xi_\mu, \xi_I\}. \tag{3.32}$$

Here a runs over the 16 PGB tastes, and the T^a are 4×4 generators of $U(4)_T$; ξ_I is the identity matrix. Under a chiral transformation, Σ transforms as

$$\mathrm{SU}(12)_L \times \mathrm{SU}(12)_R :\ \Sigma \to L\Sigma R^\dagger \tag{3.33}$$

where $L,\ R \in SU(12)_{L,R}$.

The leading order (LO) chiral Lagrangian is can be written by

$$\mathcal{L}_{\mathrm{LO}} = \mathcal{L}^4_{\mathrm{LO}} + a^2\mathcal{L}^6_{\mathrm{LO}} \tag{3.34}$$

where $\mathcal{L}^4_{\mathrm{LO}}$ is the continuum chiral Lagrangian.

The taste symmetry breaking effect is described by the $a^2\mathcal{L}^6_{\mathrm{LO}}$. To represent the matching chiral Lagrangian terms for the type-B operators in the Symanzik effective action Eq. (3.25), it is required to introduce derivatives since the type-B operators break SO(4) rotation symmetry in spin and taste. Thus the type-B operators become next-to-leading order (NLO) terms in the staggered chiral perturbation theory, and they do not contribute to $\mathcal{L}^6_{\mathrm{LO}}$. As a result, the LO chiral Lagrangian has SO(4) taste symmetry, accidentally.

Including the contribution of type-A operators, the LO chiral Lagrangian is

$$\begin{aligned}\mathcal{L}_{\mathrm{LO}} =& \frac{f^2}{8}\mathrm{Tr}(\partial_\mu\Sigma\partial_\mu\Sigma^\dagger) - \frac{1}{4}\mu f^2\mathrm{Tr}(M\Sigma + M\Sigma^\dagger) \\ &+ \frac{2m_0^2}{3}(U_I + D_I + S_I)^2 + a^2(\mathcal{U} + \mathcal{U}'),\end{aligned} \tag{3.35}$$

where M is the mass matrix,

$$M = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_d & 0 \\ 0 & 0 & m_s \end{pmatrix} \otimes \xi_I. \tag{3.36}$$

The term proportional to m_0^2 is the anomaly contribution for the η' meson mass, and the potentials $\mathcal{U}$ and $\mathcal{U}'$ describes the lattice artifacts of taste symmetry

breaking : $SU(4)_T$ to $SO(4)_T$. They are

$$\begin{aligned} -\mathcal{U} = & \quad C_1 \mathrm{Tr}(\xi_5^{(n)} \Sigma \xi_5^{(n)} \Sigma^\dagger) \\ & + \quad C_6 \sum_{\mu<\nu} \mathrm{Tr}(\xi_{\mu\nu}^{(n)} \Sigma \xi_{\nu\mu}^{(n)} \Sigma^\dagger) \\ & + \quad C_3 \tfrac{1}{2} \sum_\nu [\mathrm{Tr}(\xi_\nu^{(n)} \Sigma \xi_\nu^{(n)} \Sigma) + h.c.] \\ & + \quad C_4 \tfrac{1}{2} \sum_\nu [\mathrm{Tr}(\xi_{\nu 5}^{(n)} \Sigma \xi_{5\nu}^{(n)} \Sigma) + h.c.] \,, \end{aligned} \tag{3.37}$$

$$\begin{aligned} -\mathcal{U}' = & \quad C_{2V} \tfrac{1}{4} \sum_\nu [\mathrm{Tr}(\xi_\nu^{(n)} \Sigma) \mathrm{Tr}(\xi_\nu^{(n)} \Sigma) + h.c.] \\ & + \quad C_{2A} \tfrac{1}{4} \sum_\nu [\mathrm{Tr}(\xi_{\nu 5}^{(n)} \Sigma) \mathrm{Tr}(\xi_{5\nu}^{(n)} \Sigma) + h.c.] \\ & + \quad C_{5V} \tfrac{1}{2} \sum_\nu [\mathrm{Tr}(\xi_\nu^{(n)} \Sigma) \mathrm{Tr}(\xi_\nu^{(n)} \Sigma^\dagger)] \\ & + \quad C_{5A} \tfrac{1}{2} \sum_\nu [\mathrm{Tr}(\xi_{\nu 5}^{(n)} \Sigma) \mathrm{Tr}(\xi_{5\nu}^{(n)} \Sigma^\dagger)] \,. \end{aligned} \tag{3.38}$$

where $T^{a(n)} = T^{a(3)} \equiv I_3 \otimes T^a$ in a theory with three flavors and four tastes for flavor, which we call the 4+4+4 theory. For details, see Refs. [90, 91, 92]

3.2.2. Propagators

The propagators for the PGB fields are obtained expanding the LO Lagrangian through $\mathcal{O}(\phi^2)$ [91]:

$$\langle \phi_{ij}^a \phi_{kl}^b \rangle = \delta^{ab} \left(\delta_{il}\delta_{jk} \frac{1}{q^2 + \frac{1}{2}(I_a + J_a)} + \delta_{ij}\delta_{kl} D_{il}^a \right) \tag{3.39}$$

where $i, j, k, l \in \{u, d, s\}$ are flavor indices, and a, b are taste indices in the adjoint irrep. Here D_{il}^a is a disconnected propagator defined by

$$D_{il}^a = -\frac{\delta_a}{(q^2 + I_a)(q^2 + L_a)} \times \frac{(q^2 + U_a)(q^2 + D_a)(q^2 + S_a)}{(q^2 + \pi_a^0)(q^2 + \eta_a)(q^2 + \eta_a')}, \tag{3.40}$$

where

$$\delta_I = 4m_0^2/3, \quad \delta_5 = 0, \quad \delta_\mu = a^2 \delta_V', \quad \delta_{\mu 5} = a^2 \delta_A', \quad \delta_{\mu\nu} = 0, \tag{3.41}$$

where $\delta_{V,A}'$ are the hairpin couplings defined below. The names of mesons are the squared tree-level masses of the meson. From the LO Lagrangian,

Eq. (3.35), the tree-level meson masses for $X \in \{I, J, L, U, D, S\}$ are

$$X_a \equiv m_{X_a}^2 = 2\mu m_x + a^2 \Delta_a, \tag{3.42}$$

where m_x is the mass of the quark of flavor x and Δ_a is the taste splitting defined below. For $X \in \{\pi^0, \eta, \eta'\}$, the tree-level meson masses are the eigenvalues of the matrix,

$$\begin{pmatrix} U_a + \delta_a & \delta_a & \delta_a \\ \delta_a & D_a + \delta_a & \delta_a \\ \delta_a & \delta_a & S_a + \delta_a \end{pmatrix}. \tag{3.43}$$

In the case of flavor-charged mesons (ϕ_{xy} for $x \neq y$), the tree-level meson masses are

$$P_t \equiv \frac{1}{2}(X_a + Y_a) = \mu(m_x + m_y) + a^2 \Delta_a, \tag{3.44}$$

where $X, Y \in \{U, D, S\}$ and $x, y \in \{u, d, s\}$. Expanding the LO Lagrangian, one finds the hairpin couplings $\delta'_{V,A}$ and taste splittings Δ_a [91],

$$\delta'_V = \frac{16}{f^2}(C_{2V} - C_{5V}), \tag{3.45}$$

$$\delta'_A = \frac{16}{f^2}(C_{2A} - C_{5A}), \tag{3.46}$$

and

$$\Delta_I = \frac{16}{f^2}(4C_3 + 4C_4), \tag{3.47}$$

$$\Delta_P = 0, \tag{3.48}$$

$$\Delta_V = \frac{16}{f^2}(C_1 + 3C_6 + C_3 + 3C_4), \tag{3.49}$$

$$\Delta_A = \frac{16}{f^2}(C_1 + 3C_6 + 3C_3 + C_4), \tag{3.50}$$

$$\Delta_T = \frac{16}{f^2}(4C_6 + 2C_3 + 2C_4). \tag{3.51}$$

4. Decay Constants in Staggered Chiral Perturbation Theory

The hadronic matrix elements in the leptonic decays of the pion and kaon are parameterized by the decay constants, f_π and f_K. From the ratio f_K/f_π and the semileptonic decay $K \to \pi\ell\nu$, one can obtain the $|V_{us}|$, and it can be used to test the unitarity of the CKM matrix [28, 50]. f_π in the lattice also can be used to determine the lattice spacing [51].

Lattice calculations of the decay constants have been limited to the taste Goldstone sector associated with the exact chiral symmetry of the staggered action. In Ref. [49], Aubin and Bernard calculated the decay constants of the taste Goldstone pions and kaons through next-to-leading order (NLO) in staggered chiral perturbation theory. Here we extend their calculation to the taste non-Goldstone pions and kaons; we find that the general results are simply related to those in the taste Goldstone case [93].

In Sec. 4.1 we recall the staggered chiral Lagrangian and the tree-level propagators. In Sec. 4.2 we consider the definition of the decay constants, recall the various contributions through NLO in staggered chiral perturbation theory, and write down the general results in the 4+4+4 theory. Sec. 4.3 contains the results for specific cases of interest in the 1+1+1 theory, and we conclude in Sec. 4.4.

4.1. Chiral Lagrangian that contribute to the decay constants at NLO

In this section, we specify the chiral Lagrangian terms that contribute to the decay constant at NLO. We use the standard power counting Eq. (3.28). The order of Lagrangian operators can be parameterized in terms of n_{p^2}, n_m and n_{a^2}, which are the number of derivative pairs, powers in quark mass and powers in the squared lattice spacing, respectively. At leading order (LO), the Lagrangian operators fall into three classes: $(n_{p^2}, n_m, n_{a^2}) = (1,0,0)$, $(0,1,0)$ and $(0,0,1)$, and we have the LO chiral Lagrangian, Eq. (3.35).

At NLO, there are six classes that satisfies $n_{p^2} + n_m + n_{a^2} = 2$, but only the two combinations of operators contribute to the decay constants: $(n_{p^2}, n_m, n_{a^2}) =$, $(1,1,0)$ and $(1,0,1)$. The former composes the Gasser-Leutwyler terms [73],

$$\mathcal{L}_{\rm GL} = L_4 \mathrm{Tr}(\partial_\mu \Sigma^\dagger \partial_\mu \Sigma)\mathrm{Tr}(\chi^\dagger \Sigma + \chi \Sigma^\dagger) + L_5 \mathrm{Tr}(\partial_\mu \Sigma^\dagger \partial_\mu \Sigma(\chi^\dagger \Sigma + \Sigma^\dagger \chi))\,, \quad (4.1)$$

(n_{p^2}, n_m, n_{a^2})	Form	Comment
(2,0,0)	$\mathcal{O}(p^4)$	Do not contribute to f_π at NLO ($(\partial_\mu\phi)^4$)
(0,2,0)	$\mathcal{O}(m^2)$	Do not contribute to f_π (Constant)
(0,1,1)	$\mathcal{O}(a^2 m)$	Do not contribute to f_π (No derivatives)
(0,0,2)	$\mathcal{O}(a^4)$	Do not contribute to f_π (Constant)
(1,1,0)	$\mathcal{O}(mp^2)$	Gasser-Leutwyler terms
(1,0,1)	$\mathcal{O}(a^2p^2)$	Sharpe and Van de Water terms

Table 4.1: Classes at NLO Lagrangian that satisfies $n_{p^2} + n_m + n_{a^2} = 2$. Only the operators in the last two classes contribute to the decay constant at NLO.

where $\chi = 2\mu M$, and the later composes the $\mathcal{O}(p^2a^2)$ terms enumerated by Sharpe and Van de Water [92]. The contributions from those NLO terms to the decay constants will be discussed in Sec. 4.2.3. In Table 4.1, we list the six classes of NLO Lagrangian.

4.2. Decay constants of flavor-charged pseudo-goldstone bosons

For a PGB with taste t, P_t^+, the decay constant $f_{P_t^+}$ is defined by the matrix element of the axial current, $A_{\mu,t}^{P^+}$, between the single-particle state and the vacuum:

$$\langle 0|A_{\mu,t}^{P^+}|P_t^+(p)\rangle = -if_{P_t^+}p_\mu. \tag{4.2}$$

Let us derive the axial current. The full chiral symmetry transforms the Σ field by

$$\Sigma \to L\Sigma R^\dagger, \tag{4.3}$$

where $L, R \in SU(12)_{L,R}$. For a fixed flavor, P^+, we can consider the following symmetry transformation:

$$\Sigma \to L_{P^+}\Sigma R_{P^+}^\dagger, \tag{4.4}$$

where L_{P^+} and R_{P^+} are defined by

$$L_{P^+} \equiv \exp\left[-i\frac{1}{2}\theta_t^L \mathcal{P}^{P^+} T^{t(3)}\right] \tag{4.5}$$

$$R_{P^+} \equiv \exp\left[-i\frac{1}{2}\theta_t^R \mathcal{P}^{P^+} T^{t(3)}\right], \tag{4.6}$$

$T^{t(3)} \equiv I_3 \otimes T^t$, I_3 is the identity matrix in flavor space, and $\mathcal{P}^{P^+}$ is a projection operator that chooses the P^+ from the Σ field. For example, for π^+ it is $\mathcal{P}^{\pi^+}_{ij} = \delta_{i1}\delta_{j2}$. In general, $\mathcal{P}^{P^+}_{ij} = \delta_{ix}\delta_{jy}$, where x and y are the light quarks in P^+. Note that $\mathcal{P}^{P^+}$ and $T^{t(3)}$ commute with each other.

To calculate the left current, $A^{P^+_t,L}_{\mu,t}$, we set $\theta^R_t = 0$ and $\theta^L_t = \theta^L_t(x)$ in Eq. (4.4), which yields

$$\Sigma \to L_{P^+}\Sigma R^\dagger_{P^+} = \left[1 - i\frac{1}{2}\theta^L_t \mathcal{P}^{P^+} T^{t(3)}\right]\Sigma \tag{4.7}$$

$$\Sigma^\dagger \to R_{P^+}\Sigma^\dagger L^\dagger_{P^+} = \Sigma^\dagger\left[1 + i\frac{1}{2}\theta^L_t \mathcal{P}^{P^+} T^{t(3)}\right]. \tag{4.8}$$

Using above infinitesimal transformation and Eq. (A.9), one can calculate the left current from the leading order Lagrangian, Eq. (3.35) as follows:

$$A^{P^+,L}_{\mu,t} = i\frac{f^2}{8}\mathrm{Tr}\left[T^{t(3)}\mathcal{P}^{P^+}\partial_\mu\Sigma\Sigma^\dagger\right]. \tag{4.9}$$

Similarly, the right current is

$$A^{P^+,R}_{\mu,t} = -i\frac{f^2}{8}\mathrm{Tr}\left[T^{t(3)}\mathcal{P}^{P^+}\Sigma^\dagger\partial_\mu\Sigma\right]. \tag{4.10}$$

Subtracting the left current from the right current, the axial current at leading order Lagrangian is written as

$$\begin{aligned} A^{P^+}_{\mu,t} &= A^{P^+,R}_{\mu,t} - A^{P^+,L}_{\mu,t} \\ &= -i\frac{f^2}{8}\mathrm{Tr}\left[T^{t(3)}\mathcal{P}^{P^+}(\partial_\mu\Sigma\Sigma^\dagger + \Sigma^\dagger\partial_\mu\Sigma)\right]. \end{aligned} \tag{4.11}$$

Expanding Σ gives

$$\partial_\mu\Sigma\Sigma^\dagger + \Sigma^\dagger\partial_\mu\Sigma = \frac{2i}{f}\partial_\mu\phi - \frac{i}{3f^3}\left(\partial_\mu\phi\phi^2 - 2\phi\partial_\mu\phi\phi + \phi^2\partial_\mu\phi\right) + \cdots. \tag{4.12}$$

The $\mathcal{O}(\phi)$ term of the axial current gives the LO term of the decay constants, f, and NLO corrections from the wavefunction renormalization, $\delta f^Z_{P^+_t}$. The $\mathcal{O}(\phi^3)$ term of the axial current gives one-loop corrections at NLO, $\delta f^{\text{current}}_{P^+_t}$. Fig. 4.1(a) and Fig. 4.1(b) show these two corrections to the decay constant. In addition, there is an analytic contribution to the decay constants, $\delta f^{\text{anal}}_{P^+_t}$, from the NLO current. Combining these three types of corrections, we write the decay constants:

$$f_{P^+_t} = f\left[1 + \frac{1}{16\pi^2 f^2}\left(\delta f^Z_{P^+_t} + \delta f^{\text{current}}_{P^+_t}\right) + \delta f^{\text{anal}}_{P^+_t}\right]. \tag{4.13}$$

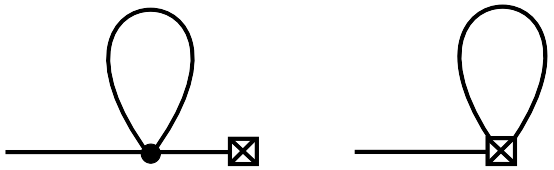


Figure 4.1: Diagrams contributing to the decay constants. (a) is the wavefunction renormalization correction and (b) is the current correction.

In this section we outline the calculation of $\delta f_{P_t^+} \equiv \delta f^Z_{P_t^+} + \delta f^{\text{current}}_{P_t^+}$ and the analytic contributions, $\delta f^{\text{anal}}_{P_t^+}$, and present results for the 4+4+4 theory.

4.2.1. Wavefunction renormalization correction

Up to $\mathcal{O}(\phi)$, we can write the axial current, Eq. (4.11), as follows:

$$A^{P^+,\phi}_{\mu,t} = f\left(\mathcal{P}^{P^+}_{ij}\partial_\mu \phi^t_{ji}\right), \tag{4.14}$$

where we used $\tau_{ta} = 4\delta_{ta}$. Here $\tau_{abcd\cdots}$ is defined by

$$\tau_{abcd\cdots} \equiv \text{Tr}(T_a T_b T_c T_d \cdots). \tag{4.15}$$

The contributions of the $\mathcal{O}(\phi)$ current to the decay constants are defined by the matrix element

$$\langle 0|A^{P^+,\phi}_{\mu,t}|P^+_t(p)\rangle = f(-ip_\mu)\langle 0|\phi^t_{yx}|P^+_t(p)\rangle \tag{4.16}$$

$$= f(-ip_\mu)\sqrt{Z_{P_t^+}}, \tag{4.17}$$

where we used $\mathcal{P}^{P^+}_{ij} = \delta_{ix}\delta_{jy}$, and here, $Z_{P_t^+} = 1 + \delta Z_{P_t^+}$ is the wavefunction renormalization constant of the ϕ field. The wavefunction renormalization corrections to the decay constants at NLO is

$$\delta f^Z_{P_t^+} \equiv \frac{16\pi^2 f^2}{2}\delta Z_{P_t^+} = -\frac{16\pi^2 f^2}{2}\frac{d\Sigma(p^2)}{dp^2}, \tag{4.18}$$

where $\Sigma(p^2)$ is the self-energy of P_t^+. Using the self-energy from Ref. [94], we find the one-loop corrections

$$\delta f^Z_{P_t^+} = \frac{1}{24}\sum_a \left[\sum_Q l(Q_a) + 16\pi^2 \int \frac{d^4q}{(2\pi)^4} \left(D^a_{xx} + D^a_{yy} - 2\theta^{at} D^a_{xy} \right) \right], \quad (4.19)$$

where Q runs over six flavor combinations, xi and yi for $i \in \{u, d, s\}$, a runs over the 16 PGB tastes in the **15** and **1** of $SU(4)_T$ and Q_a is the squared tree-level meson mass with flavor Q and taste a. Here $l(Q_a)$ is the chiral logarithm,

$$l(X) \equiv X\Big(\ln X/\Lambda^2 + \delta_1(\sqrt{X}L)\Big) \quad (4.20)$$

for any squared meson mass X and $\delta_1(\sqrt{X}L)$ is is the finite-volume correction [95],

$$\delta_1(\sqrt{X}L) \equiv \frac{4}{\sqrt{X}L} \sum_{\vec{n}\neq\vec{0}} \frac{K_1(|\vec{n}|\sqrt{X}L)}{|\vec{n}|}, \quad (4.21)$$

where K_1 is the Bessel function. Note that $\delta_1(\sqrt{X}L) \to 0$ in infinite volume. The self-energy also contains NLO analytic corrections to the decay constants. It will be discussed in Sec. 4.2.3.

The renormalization scale parameter Λ in Eq. (4.20) comes from the dimensional regularization of the divergent integral. Note that the NLO loop integrals are suppressed by a factor of $1/(4\pi f)^2$, and the NLO operators in the Lagrangian has more powers of ∂_μ^2 or μM that should be suppressed by the scale parameter $1/\Lambda^2$. Thus the natural (not necessary) choice of the scale is $\Lambda \approx 4\pi f \sim 1\text{GeV}$. The Λ dependence of the chiral logarithm is observed by the low energy constants in the analytic terms, L_4, L_5 and $\mathscr{F}$, in Eq. (4.43).

4.2.2. Current correction

Collecting all the $\mathcal{O}(\phi^3)$ terms from the axial current, we find

$$A^{P^+,\phi^3}_{\mu,t} = -\frac{1}{24f}\tau_{tabc}\Big(\mathcal{P}_{ij}\partial_\mu\phi^a_{jk}\phi^b_{kl}\phi^c_{li} - 2\mathcal{P}_{ij}\phi^a_{jk}\partial_\mu\phi^b_{kl}\phi^c_{li} + \mathcal{P}_{ij}\phi^a_{jk}\phi^b_{kl}\partial_\mu\phi^c_{li}\Big). \quad (4.22)$$

In the calculation of the matrix element defined in Eq. (4.2), each term of Eq. (4.22) contributes one contraction because the derivatively coupled field in the current must contract with the external field for nonzero results. After con-

tracting, each term becomes

$$\tau_{tabc}\mathcal{P}_{ij}\partial_\mu\phi^a_{jk}\overbrace{\phi^b_{kl}\phi^c_{li}} \quad \to \quad -ip_\mu\delta^{ta,bc}_{ix,jy,kx}K^b_{kl,li}, \tag{4.23}$$

$$-2\tau_{tabc}\mathcal{P}_{ij}\overbrace{\phi^a_{jk}\partial_\mu\phi^b_{kl}}\phi^c_{li} \quad \to \quad 2ip_\mu\delta^{tb,ac}_{ix,jy,ky,lx}K^a_{jk,li}, \tag{4.24}$$

$$\tau_{tabc}\mathcal{P}_{ij}\overbrace{\phi^a_{jk}\phi^b_{kl}}\partial_\mu\phi^c_{li} \quad \to \quad -ip_\mu\delta^{tc,ab}_{ix,jy,ly}K^a_{jk,kl}, \tag{4.25}$$

where

$$K^a_{ij,kl} \equiv \int \frac{d^4q}{(2\pi)^4}\langle\phi^a_{ij}\phi^a_{kl}\rangle \qquad \text{(no sum)}. \tag{4.26}$$

Collecting the three contributions from Eq. (4.22), we find

$$i\frac{p_\mu}{6f}\sum_a\left[\sum_j\left(K^a_{xj,jx}+K^a_{yj,jy}\right)-2\theta^{at}K^a_{xx,yy}\right], \tag{4.27}$$

where j runs over $\{u,d,s\}$ and $\theta^{ab}\equiv\frac{1}{4}\tau_{abab}$.

The integral in the $K^a_{ij,kl}$ can be performed using the propagators given in Eq. (3.39). For $x\neq l$,

$$K^a_{xl,lx}=\int\frac{d^4q}{(2\pi)^4}\frac{1}{q^2+\frac{1}{2}(X_a+L_a)}=\frac{1}{16\pi^2}l\left(\frac{1}{2}X_a+\frac{1}{2}L_a\right), \tag{4.28}$$

where we use dimensional regularization to obtain the chiral logarithm $l(X)$ in the last equality. For $x=l$,

$$K^a_{xl,lx}=\frac{1}{16\pi^2}l\left(\frac{1}{2}X_a+\frac{1}{2}L_a\right)+\int\frac{d^4q}{(2\pi)^4}D^a_{xx}. \tag{4.29}$$

Combining these two results, we find

$$\sum_j\left(K^a_{xj,jx}+K^a_{yj,jy}\right)=\sum_Q\frac{1}{16\pi^2}l(Q_a)+\int\frac{d^4q}{(2\pi)^4}\left(D^a_{xx}+D^a_{yy}\right), \tag{4.30}$$

where Q runs over six flavor combinations, xi and yi for $i\in\{u,d,s\}$. Since only the flavor-charged ($x\neq y$) cases are the interest here, $K^a_{xx,yy}$ is evaluated as

$$K^a_{xx,yy}=\int\frac{d^4q}{(2\pi)^4}D^a_{xy}. \tag{4.31}$$

Using these results, we find the one-loop current correction to the pion decay

constants:

$$\delta f_{P_t^+}^{\text{current}} \equiv -\frac{1}{6}\sum_a \left[\sum_Q l(Q_a) + 16\pi^2 \int \frac{d^4q}{(2\pi)^4} \left(D^a_{xx} + D^a_{yy} - 2\theta^{at} D^a_{xy} \right) \right]. \quad (4.32)$$

Note that $\delta f_{P_t^+}^{\text{current}}$ is proportional to the wavefunction renormalization correction, $\delta f_{P_t^+}^{Z}$.

$$\delta f_{P_t^+}^{\text{current}} = -4\delta f_{P_t^+}^{Z} \quad (4.33)$$

This was shown in the taste Goldstone case in Ref. [49].

4.2.3. Next-to-leading order analytic contributions

Now we consider the analytic contributions to the decay constants. The analytic contributions come from the terms of NLO Lagrangian given in Sec. 4.1. The contributions of $\mathcal{O}(p^2a^2)$ terms in the Lagrangian to the decay constants are $fa^2\mathcal{F}_t$ with low-energy constants (LECs) $\mathcal{F}_t$.

For example, one of the $\mathcal{O}(p^2a^2)$ operators is

$$a^2\text{Tr}(\partial_\mu\Sigma^\dagger\partial_\mu\Sigma\Sigma^\dagger T_a\Sigma T_a) + (\Sigma \leftrightarrow \Sigma^\dagger) \quad (4.34)$$

where $(\Sigma \leftrightarrow \Sigma^\dagger)$ is the parity conjugate term. The contribution from the operator to the decay constants at NLO arises in two ways: tree-level wavefunction renormalization correction and tree-level current correction. The wavefunction renormalization correction is obtained by expanding the operator through $\mathcal{O}(\phi^2)$,

$$\tau_{ab}\partial_\mu\phi^a_{ij}\partial_\mu\phi^b_{ji}. \quad (4.35)$$

The corresponding self-energy is

$$\Sigma(p^2) \sim a^2p^2, \quad (4.36)$$

and the corresponding wavefunction renormalization correction is

$$\delta f^Z_{a^2p^2} = -\frac{1}{2}\frac{d\Sigma(p^2)}{dp^2} \sim a^2 \times \text{constant} \quad (4.37)$$

The current correction is obtained by calculating the axial current,

$$A^{P^+}_{\mu,t} \sim a^2 \partial_\mu \phi^t_{yx}, \tag{4.38}$$

$$\langle 0|A^{P^+}_{\mu,t}|P^+_t(p)\rangle \sim a^2 p_\mu, \tag{4.39}$$

$$\delta f^{\text{current}}_{a^2p^2} \sim a^2 \times \text{constant}. \tag{4.40}$$

The other $\mathcal{O}(p^2a^2)$ terms have the same structure (two derivatives), and they give the same form of NLO correction to the decay constants. Since there are splittings between tastes, we write the correction as $\mathcal{F}_t$.

$\mathcal{F}_t$ are degenerate within the irreps of the lattice symmetry group. As mentioned in Ref. [92], there is no relations between the SO(4)-violations in the pion masses and the axial current decay constants due to the contributions from the $\mathcal{O}(p^2a^2)$ source operators. At NLO axial currents, there are extra terms in addition to the axial currents derived from the NLO Lagrangian. They are incorporated by adding additional source terms in the chiral Lagrangian [92]. Since they affect only the axial currents, the pattern of SO(4)-violations in the axial currents is different from that in the pion masses.

The Gasser-Leutwyler terms given in Eq. (4.1) contribute to the decay constants through wavefunction renormalization and the current. The self-energy from the Gasser-Leutwyler terms calculated in Ref. [94] gives the wavefunction renormalization correction is

$$\delta f^{Z,\text{GL}}_{P^+_t} = -\frac{64}{f^2} L_4 \mu (m_u + m_d + m_s) - \frac{8}{f^2} L_5 \mu (m_x + m_y), \tag{4.41}$$

while a direct calculation of NLO corrections from the Gasser-Leutwyler terms gives the current correction,

$$\delta f^{\text{current,GL}}_{P^+_t} = \frac{128}{f^2} L_4 \mu (m_u + m_d + m_s) + \frac{16}{f^2} L_5 \mu (m_x + m_y). \tag{4.42}$$

Collecting these three corrections, we find the analytic contribution to the decay constants:

$$\delta f^{\text{anal}}_{P^+_t} = \frac{64}{f^2} L_4 \mu (m_u + m_d + m_s) + \frac{8}{f^2} L_5 \mu (m_x + m_y) + a^2 \mathcal{F}_t. \tag{4.43}$$

4.3. Results

To formulate the full QCD and (partially) quenched results in rooted staggered chiral perturbation theory (rSChPT), we employ the replica method [96, 97, 98]. Rooting introduces a factor of 1/4 in the chiral logarithms in Eqs. (4.19) and (4.32) and in the L_4 term in Eq. (4.43). The rooting also replace the tree-level meson masses for π^0, η, η' defined in Eq. (3.43) by the eigen-

$B\backslash F$	V	A	T	P	I
V	-2	2	0	-4	4
A	2	-2	0	-4	4
T	0	0	-2	6	6
P	-1	-1	1	1	1
I	1	1	1	1	1

Table 4.2: The coefficient Θ^{BF} defined in Eq. (4.50) is in row B and column F.

values of

$$\begin{pmatrix} U_a+\delta_a/4 & \delta_a/4 & \delta_a/4 \\ \delta_a/4 & D_a+\delta_a/4 & \delta_a/4 \\ \delta_a/4 & \delta_a/4 & S_a+\delta_a/4 \end{pmatrix}. \tag{4.44}$$

The results in rSChPT are

$$\delta f_{P_F^+} = \delta f_{P_F^+}^{\text{con}} + \delta f_{P_F^+}^{\text{disc}}, \tag{4.45}$$

$$\delta f_{P_t^+}^{\text{anal}} = \frac{16}{f^2} L_4 \mu (m_u + m_d + m_s) + \frac{8}{f^2} L_5 \mu (m_x + m_y) + a^2 \mathcal{F}_t. \tag{4.46}$$

where

$$\delta f_{P_F^+}^{\text{con}} \equiv -\frac{1}{32} \sum_{Q,B} g_B \; l(Q_B), \tag{4.47}$$

$$\begin{aligned} \delta f_{P_F^+}^{\text{disc}} \equiv -2\pi^2 \int \frac{d^4q}{(2\pi)^4} \Big(& D_{xx}^I + D_{yy}^I - 2D_{xy}^I \\ & + 4D_{xx}^V + 4D_{yy}^V - 2\Theta^{VF} D_{xy}^V \\ & + 4D_{xx}^A + 4D_{yy}^A - 2\Theta^{AF} D_{xy}^A \Big). \end{aligned} \tag{4.48}$$

Here we summed over a within each irrep for Eqs. (4.19) and (4.32), B and F represent the $SO(4)_T$ irreps,

$$B, F \in \{I, V, T, A, P\}, \tag{4.49}$$

$t \in F$ and

$$\Theta^{BF} \equiv \sum_{a \in B} \theta^{at}, \quad g_B \equiv \sum_{a \in B} 1. \tag{4.50}$$

The coefficients Θ^{BF} are given in Table 4.2.

In this section, we give the one-loop contributions to the decay constants in various cases. In Sec. 4.3.1.1 and Sec. 4.3.1.2, fully dynamical and partially

quenched results for the 1+1+1 and 2+1 flavor cases in SU(3) chiral perturbation theory are given. The analogous results in SU(2) chiral perturbation theory are presented in Sec. 4.3.2. In Sec. 4.3.1.3, the results in the quenched case are given. In this section, we use the same notation as in Ref. [94].

4.3.1. SU(3) chiral perturbation theory

4.3.1.1. Fully dynamical case

In Eq. (4.47) Q runs over six flavor combinations, xi and yi for $i \in \{u, d, s\}$. Setting $xy = ud,\ us,\ ds$ gives the results for the π^+, K^+, and K^0 in the fully dynamical 1+1+1 flavor case:

$$\delta f^{\rm con}_{\pi^+_F} = -\frac{1}{32}\sum_B g_B\Big(l(U_B) + 2l(\pi^+_B) + l(K^+_B) + l(D_B) + l(K^0_B)\Big), \tag{4.51}$$

$$\delta f^{\rm con}_{K^+_F} = -\frac{1}{32}\sum_B g_B\Big(l(U_B) + l(\pi^+_B) + 2l(K^+_B) + l(K^0_B) + l(S_B)\Big). \tag{4.52}$$

$$\delta f^{\rm con}_{K^0_F} = -\frac{1}{32}\sum_B g_B\Big(l(\pi^+_B) + l(D_B) + 2l(K^0_B) + l(K^+_B) + l(S_B)\Big). \tag{4.53}$$

The integrals in the disconnected parts, Eq.(4.48) can be performed as explained in Ref. [91]. Expanding the integrands (D^a_{ij}) in terms of the simple poles multiplied by their residues, the integrals becomes chiral logarithms multiplied by residues. For example,

$$D^I_{uu} = -\frac{\delta_I(q^2 + D_I)(q^2 + S_I)}{(q^2 + U_I)(q^2 + \pi^0_I)(q^2 + \eta_I)(q^2 + \eta'_I)} \tag{4.54}$$

$$= -\delta_I \sum_X \frac{1}{q^2 + X_I} R^{DS}_{U\pi^0\eta\eta'}(X_I), \tag{4.55}$$

where X runs over $\{U, \pi^0, \eta, \eta'\}$, and

$$R^{DS}_{U\pi^0\eta\eta'}(X_I) = \frac{(D_I - X_I)(S_I - X_I)}{(\pi^0_I - X_I)(\eta_I - X_I)(\eta'_I - X_I)}, \tag{4.56}$$

Performing the integral,

$$\int \frac{d^4q}{(2\pi)^4} D^I_{uu} = -\delta_I \frac{1}{16\pi^2} \sum_X R^{DS}_{U\pi^0\eta\eta'}(X_I) l(X_I) \tag{4.57}$$

$$= -\frac{4}{3}\frac{1}{16\pi^2} \sum_X R^{DS}_{U\pi^0\eta}(X_I) l(X_I), \tag{4.58}$$

where we take $m_0 \to \infty$ while $\delta_I = \frac{4}{3} m_0^2$ and $m_{\eta'}^2 = m_0^2$ in the last equality to decouple the η' meson.

After performing the integrals and decoupling the η'_I by taking $m_0^2 \to \infty$ [99], we find

$$\begin{aligned}\delta f^{\text{disc}}_{\pi^+_F} = \sum_X \Bigg[& \frac{1}{6}\Big\{ R^{DS}_{U\pi^0\eta}(X_I) l(X_I) + R^{US}_{D\pi^0\eta}(X_I) l(X_I) - 2R^{S}_{\pi^0\eta}(X_I) l(X_I) \Big\} \\ & + \frac{1}{4} a^2 \delta'_V \Big\{ 2R^{DS}_{U\pi^0\eta\eta'}(X_V) l(X_V) + 2R^{US}_{D\pi^0\eta\eta'}(X_V) l(X_V) \\ & \qquad - \Theta^{VF} R^{S}_{\pi^0\eta\eta'}(X_V) l(X_V) \Big\} \\ & + (V \to A) \Bigg],\end{aligned} \tag{4.59}$$

$$\begin{aligned}\delta f^{\text{disc}}_{K^+_F} = \sum_X \Bigg[& \frac{1}{6}\Big\{ R^{DS}_{U\pi^0\eta}(X_I) l(X_I) + R^{UD}_{S\pi^0\eta}(X_I) l(X_I) - 2R^{D}_{\pi^0\eta}(X_I) l(X_I) \Big\} \\ & + \frac{1}{4} a^2 \delta'_V \Big\{ 2R^{DS}_{U\pi^0\eta\eta'}(X_V) l(X_V) + 2R^{UD}_{S\pi^0\eta\eta'}(X_V) l(X_V) \\ & \qquad - \Theta^{VF} R^{D}_{\pi^0\eta\eta'}(X_V) l(X_V) \Big\} \\ & + (V \to A) \Bigg],\end{aligned} \tag{4.60}$$

$$\begin{aligned}\delta f^{\text{disc}}_{K^0_F} = \sum_X \Bigg[& \frac{1}{6}\Big\{ R^{US}_{D\pi^0\eta}(X_I) l(X_I) + R^{UD}_{S\pi^0\eta}(X_I) l(X_I) - 2R^{U}_{\pi^0\eta}(X_I) l(X_I) \Big\} \\ & + \frac{1}{4} a^2 \delta'_V \Big\{ 2R^{US}_{D\pi^0\eta\eta'}(X_V) l(X_V) + 2R^{UD}_{S\pi^0\eta\eta'}(X_V) l(X_V) \\ & \qquad - \Theta^{VF} R^{U}_{\pi^0\eta\eta'}(X_V) l(X_V) \Big\} \\ & + (V \to A) \Bigg].\end{aligned} \tag{4.61}$$

In the sum, X runs over the subscripts of the residue, R, where the residues are defined by

$$R^{A_1 A_2 \cdots A_k}_{B_1 B_2 \cdots B_n}(X_F) \equiv \frac{\prod_{A_j}(A_{jF} - X_F)}{\prod_{B_i \neq X}(B_{iF} - X_F)}, \tag{4.62}$$

where $X \in \{B_1, B_2, \ldots, B_n\}$ and $F \in \{V, A, I\}$ is the $SO(4)_T$ irrep.

The results in the 2+1 flavor case are easily obtained by setting $xy = ud,\ us$ and $m_u = m_d$. Eq. (4.47) gives connected contributions for the π and K:

$$\delta f^{\rm con}_{\pi_F} = -\frac{1}{16}\sum_B g_B \Big(2l(\pi_B) + l(K_B) \Big), \tag{4.63}$$

$$\delta f^{\rm con}_{K_F} = -\frac{1}{32}\sum_B g_B \Big(2l(\pi_B) + 3l(K_B) + l(S_B) \Big). \tag{4.64}$$

Setting $xy = ud,\ us$ and $m_u = m_d$ in Eq. (4.48) gives

$$\begin{aligned} \delta f^{\rm disc}_{\pi_F} = {} & \frac{1}{4} a^2 \delta'_V (4 - \Theta^{VF}) \sum_X R^S_{\pi\eta\eta'}(X_V) l(X_V) \\ & + (V \to A), \end{aligned} \tag{4.65}$$

and

$$\begin{aligned} \delta f^{\rm disc}_{K_F} = {} & \frac{1}{6}\sum_X \Big\{ R^S_{\pi\eta}(X_I) l(X_I) + R^\pi_{S\eta}(X_I) l(X_I) \Big\} - 2l(\eta_I) \\ & + \frac{1}{4} a^2 \delta'_V \sum_X \Big\{ 2R^S_{\pi\eta\eta'}(X_V) l(X_V) + 2R^\pi_{S\eta\eta'}(X_V) l(X_V) \\ & \qquad - \Theta^{VF} R_{\eta\eta'}(X_V) l(X_V) \Big\} \\ & + (V \to A), \end{aligned} \tag{4.66}$$

where $R_{B_1 B_2}(X_F)$ is defined by

$$R_{B_1 B_2}(X_F) = \begin{cases} \frac{1}{B_2 - B_1} & (X_F = B_1) \\ \frac{1}{B_1 - B_2} & (X_F = B_2) \end{cases} . \tag{4.67}$$

Using the tree-level masses of the taste singlet channel, one finds

$$R^S_{\pi\eta}(\pi_I) = \frac{3}{2}, \quad R^S_{\pi\eta}(\eta_I) = -\frac{1}{2}, \tag{4.68}$$

$$R^\pi_{S\eta}(S_I) = 3, \quad R^\pi_{S\eta}(\eta_I) = -2. \tag{4.69}$$

They simplify the results, Eqs. (4.65) and (4.66):

$$\begin{aligned}\delta f^{\rm disc}_{\pi_F} =& \frac{1}{4}a^2\delta'_V(4-\Theta^{VF})\Bigg[\frac{S_V-\pi_V}{(\eta_V-\pi_V)(\eta'_V-\pi_V)}l(\pi_V)\\ &+\frac{S_V-\eta_V}{(\pi_V-\eta_V)(\eta'_V-\eta_V)}l(\eta_V)+\frac{S_V-\eta'_V}{(\pi_V-\eta'_V)(\eta_V-\eta'_V)}l(\eta'_V)\Bigg]\\ &+(V\to A),\end{aligned} \tag{4.70}$$

$$\begin{aligned}\delta f^{\rm disc}_{K_F} =& \frac{1}{12}\Big[3l(\pi_I)-5l(\eta_I)+6l(S_I)-4l(\eta_I)\Big]\\ &+\frac{1}{2}a^2\delta'_V\Bigg[\frac{S_V-\pi_V}{(\eta_V-\pi_V)(\eta'_V-\pi_V)}l(\pi_V)\\ &+\frac{(\pi_V-\eta_V)^2+(S_V-\eta_V)^2}{(\pi_V-\eta_V)(\eta'_V-\eta_V)(S_V-\eta_V)}l(\eta_V)\\ &+\frac{(\pi_V-\eta'_V)^2+(S_V-\eta'_V)^2}{(\pi_V-\eta'_V)(\eta_V-\eta'_V)(S_V-\eta'_V)}l(\eta'_V)\\ &+\frac{\pi_V-S_V}{(\eta_V-S_V)(\eta'_V-S_V)}l(S_V)-\frac{1}{2}\Theta^{VF}\frac{1}{\eta_V-\eta'_V}\Big\{l(\eta'_V)-l(\eta_V)\Big\}\Bigg]\\ &+(V\to A).\end{aligned} \tag{4.71}$$

4.3.1.2. Partially quenched case

In the partially quenched case, the valence quark masses, m_x and m_y are not degenerate with the sea quark masses, m_u, m_d and m_s. The connected contributions to the decay constants in the partially quenched 1+1+1 flavor case are

$$\delta f^{\rm con}_{P^+_F} = -\frac{1}{32}\sum_{Q,B} g_B\ l(Q_B). \tag{4.72}$$

The disconnected parts in Eq. (4.48) can be done using the residue method used in the fully dynamical cases. For example,

$$D^a_{xy} = -\frac{\delta_a(q^2+U_a)(q^2+D_a)(q^2+S_a)}{(q^2+X_a)(q^2+Y_a)(q^2+\pi^0_a)(q^2+\eta_a)(q^2+\eta'_a)} \tag{4.73}$$

$$= -\delta_a\sum_Z \frac{1}{q^2+Z_a}R^{UDS}_{XY\pi^0\eta\eta'}(Z_a), \tag{4.74}$$

where Z runs over $\{X, Y, \pi^0, \eta, \eta'\}$. The integrals give chiral logarithms. In the case of D^a_{xx} or D^a_{yy}, there are double poles, which need special treatments. For

D^a_{xx},

$$D^a_{xx} = -\frac{\delta_a(q^2+U_a)(q^2+D_a)(q^2+S_a)}{(q^2+X_a)^2(q^2+\pi^0_a)(q^2+\eta_a)(q^2+\eta'_a)} \tag{4.75}$$

$$= \frac{\partial}{\partial X}\left[\frac{\delta_a(q^2+U_a)(q^2+D_a)(q^2+S_a)}{(q^2+X_a)(q^2+\pi^0_a)(q^2+\eta_a)(q^2+\eta'_a)}\right] \tag{4.76}$$

$$= \frac{\partial}{\partial X}\left[\delta_a \sum_Z \frac{1}{q^2+Z_a} R^{UDS}_{X\pi^0\eta\eta'}(Z_a)\right] \tag{4.77}$$

$$= -\delta_a \frac{1}{(q^2+X_a)^2} R^{UDS}_{X\pi^0\eta\eta'}(X_a) - \delta_a \sum_Z \frac{1}{q^2+Z_a} D^{UDS}_{X\pi^0\eta\eta'}(Z_a) \tag{4.78}$$

where Z runs over $\{X, \pi^0, \eta, \eta'\}$ and

$$D^{UDS}_{X\pi^0\eta\eta'}(X_a)(Z_a) = -\frac{\partial}{\partial X} R^{UDS}_{X\pi^0\eta\eta'}(Z_a). \tag{4.79}$$

For the integrals, we define a new chiral logarithm $\tilde{l}(X)$:

$$\frac{1}{16\pi^2} l(X) = \int \frac{d^4q}{(2\pi)^4} \frac{1}{q^2+X} \tag{4.80}$$

$$\frac{1}{16\pi^2} \tilde{l}(X) \equiv \int \frac{d^4q}{(2\pi)^4} \frac{1}{(q^2+X)^2} = -\frac{\partial}{\partial X}\left[\frac{1}{16\pi^2} l(X)\right]. \tag{4.81}$$

Considering finite volume effect,

$$\tilde{l}(X) \equiv -\Big(\ln X/\Lambda^2 + 1\Big) + \delta_3(\sqrt{X}L), \tag{4.82}$$

where $\delta_3(\sqrt{X}L)$ is the finite volume correction

$$\delta_3(\sqrt{X}L) = 2\sum_{\vec{n}\neq 0} K_0(|\vec{n}|\sqrt{X}L) \tag{4.83}$$

where K_0 is the Bessel function [95].

Performing the integrals in Eq. (4.48) keeping all quark masses distinct gives

the disconnected contributions for the partially quenched 1+1+1 flavor case:

$$\begin{aligned}\delta f^{\rm disc}_{P_F^+,m_x\neq m_y} = \sum_Z \Bigg[& \frac{1}{6}\Big\{ D^{UDS}_{X\pi^0\eta,X}(Z_I) l(Z_I) \\ & + D^{UDS}_{Y\pi^0\eta,Y}(Z_I) l(Z_I) - 2R^{UDS}_{XY\pi^0\eta}(Z_I) l(Z_I) \Big\} \\ & + \frac{1}{4} a^2 \delta'_V \Big\{ 2D^{UDS}_{X\pi^0\eta\eta',X}(Z_V) l(Z_V) + 2D^{UDS}_{Y\pi^0\eta\eta',Y}(Z_V) l(Z_V) \\ & - \Theta^{VF} R^{UDS}_{XY\pi^0\eta\eta'}(Z_V) l(Z_V) \Big\} + (V \to A) \Bigg] \\ & + \frac{1}{6}\Big\{ R^{UDS}_{X\pi^0\eta}(X_I)\tilde{l}(X_I) + R^{UDS}_{Y\pi^0\eta}(Y_I)\tilde{l}(Y_I) \Big\} \\ & + \frac{1}{2} a^2 \delta'_V \Big\{ R^{UDS}_{X\pi^0\eta\eta'}(X_V)\tilde{l}(X_V) + R^{UDS}_{Y\pi^0\eta\eta'}(Y_V)\tilde{l}(Y_V) \Big\} + (V \to A), \end{aligned} \tag{4.84}$$

where

$$D^{A_1 A_2 \cdots A_k}_{B_1 B_2 \cdots B_n, B_i}(X_F) \equiv -\frac{\partial}{\partial B_{iF}} R^{A_1 A_2 \cdots A_k}_{B_1 B_2 \cdots B_n}(X_F). \tag{4.85}$$

For $m_x = m_y$, we find

$$\begin{aligned}\delta f^{\rm disc}_{P_F^+,m_x=m_y} = & \frac{1}{4} a^2 \delta'_V (4 - \Theta^{VF}) \Bigg[R^{UDS}_{X\pi^0\eta\eta'}(X_V)\tilde{l}(X_V) \\ & + \sum_Z D^{UDS}_{X\pi^0\eta\eta',X}(Z_V) l(Z_V) \Bigg] \\ & + (V \to A). \end{aligned} \tag{4.86}$$

The connected contributions in the 2+1 flavor case are obtained by setting $m_u = m_d$ in Eq. (4.72). To obtain the disconnected contributions, we perform

the integrals in Eq. (4.48) setting $m_u = m_d$. For $m_x \neq m_y$, we find

$$\begin{aligned}
\delta f^{\text{disc}}_{P_F^+, m_x \neq m_y} = \sum_Z \Bigg[& \frac{1}{6}\Big\{ D^{\pi S}_{X\eta,X}(Z_I) l(Z_I) \\
& + D^{\pi S}_{Y\eta,Y}(Z_I) l(Z_I) - 2R^{\pi S}_{XY\eta}(Z_I) l(Z_I) \Big\} \\
& + \frac{1}{4} a^2 \delta'_V \Big\{ 2D^{\pi S}_{X\eta\eta',X}(Z_V) l(Z_V) \\
& + 2D^{\pi S}_{Y\eta\eta',Y}(Z_V) l(Z_V) - \Theta^{VF} R^{\pi S}_{XY\eta\eta'}(Z_V) l(Z_V) \Big\} + (V \to A) \Bigg] \\
& + \frac{1}{6}\Big\{ R^{\pi S}_{X\eta}(X_I) \tilde{l}(X_I) + R^{\pi S}_{Y\eta}(Y_I) \tilde{l}(Y_I) \Big\} \\
& + \frac{1}{2} a^2 \delta'_V \Big\{ R^{\pi S}_{X\eta\eta'}(X_V) \tilde{l}(X_V) + R^{\pi S}_{Y\eta\eta'}(Y_V) \tilde{l}(Y_V) \Big\} \\
& + (V \to A).
\end{aligned} \tag{4.87}$$

For $m_x = m_y$, we find

$$\begin{aligned}
\delta f^{\text{disc}}_{P_F^+, m_x = m_y} = & \frac{1}{4} a^2 \delta_V (4 - \Theta^{VF}) \left[R^{\pi S}_{X\eta\eta'}(X_V) \tilde{l}(X_V) + \sum_Z D^{\pi S}_{X\eta\eta',X}(Z_V) l(Z_V) \right] \\
& + (V \to A).
\end{aligned} \tag{4.88}$$

4.3.1.3. Quenched case

In the quenched case, there is no connected contribution, Eq. (4.47). Furthermore, there is no sea quarks in the denominator of D_{il}. As a results, the m_0^2 cannot be taken to infinity and it cannot be decoupled. Hence, additional η'_I dependent term is required, and it can be achieved by replacing m_0^2 by $m_0^2 + \alpha q^2$, where α is an additional parameter [100].

As explained in Refs.[91, 95, 100], quenching the sea quarks in the disconnected part can be done by replacing the disconnected propagator with

$$D^{a,\text{quench}}_{il} = -\frac{\delta^{\text{quench}}_a}{(q^2 + I_a)(q^2 + L_a)}, \tag{4.89}$$

where

$$\delta^{\text{quench}}_a = \begin{cases} 4(m_0^2 + \alpha q^2)/3 & \text{if } a = I \\ \delta_a & \text{if } a \neq I. \end{cases} \tag{4.90}$$

Replacing D^a_{il} with the quenched disconnected propagator in Eq. (4.48) for

$m_x \neq m_y$ gives

$$\begin{aligned}\delta f^{\rm disc}_{P_F^+, m_x \neq m_y} =& \frac{\alpha}{6}\Big\{\frac{Y_I + X_I}{Y_I - X_I}(l(X_I) - l(Y_I)) - X_I\tilde{l}(X_I) - Y_I\tilde{l}(Y_I)\Big\} \\ &+ \frac{m_0^2}{6}\Big\{\tilde{l}(X_I) + \tilde{l}(Y_I) - 2\frac{l(X_I) - l(Y_I)}{Y_I - X_I}\Big\} \\ &+ \frac{1}{4}a^2\delta'_V\Big\{2\tilde{l}(X_V) + 2\tilde{l}(Y_V) - \Theta^{VF}\frac{l(X_V) - l(Y_V)}{Y_V - X_V}\Big\} \\ &+ (V \to A),\end{aligned} \tag{4.91}$$

and for $m_x = m_y$,

$$\delta f^{\rm disc}_{P_F^+, m_x = m_y} = \frac{1}{4}a^2\delta'_V(4 - \Theta^{VF})\tilde{l}(X_V) + (V \to A). \tag{4.92}$$

Quenching the sea quarks also affects the analytic terms. In the quenched version of Eq. (4.43), there is no L_4 term of Eq. (4.43), which is coming from the sea quarks.

4.3.2. SU(2) chiral perturbation theory

In some cases, using the SU(2) SChPT in data analysis gives better results than those in the SU(3) SChPT, mainly because the fitting function in the SU(2) SChPT has smaller number of fitting parameters than that in the SU(3) SChPT. In this subsection, we decouple m_s and m_0 by taking the limit

$$m_x, m_y, m_u, m_d \ll m_s \ll m_0, \tag{4.93}$$

to obtain the results in the SU(2) SChPT.

4.3.2.1. Fully dynamical case

From Eqs. (4.51), (4.52), and (4.53), we obtain the connected contributions for the fully dynamical 1+1+1 flavor case:

$$\delta f^{\rm con}_{\pi_F^+} = -\frac{1}{32}\sum_B g_B\Big(l(U_B) + 2l(\pi_B^+) + l(D_B)\Big), \tag{4.94}$$

$$\delta f^{\rm con}_{K_F^+} = -\frac{1}{32}\sum_B g_B\Big(l(U_B) + l(\pi_B^+)\Big), \tag{4.95}$$

$$\delta f^{\rm con}_{K_F^0} = -\frac{1}{32}\sum_B g_B\Big(l(D_B) + l(\pi_B^+)\Big). \tag{4.96}$$

For the disconnected contributions, we find from Eqs. (4.59), (4.60) and (4.61):

$$\begin{aligned}\delta f^{\text{disc}}_{\pi^+_F} &= \frac{1}{2}\big(l(U_I)+l(D_I)\big) - l(\pi^0_I) \\ &\quad + \frac{1}{4}a^2\delta'_V \sum_X \Big\{2R^D_{U\pi^0\eta}(X_V)l(X_V) + 2R^U_{D\pi^0\eta}(X_V)l(X_V)\Big\} \\ &\qquad + \frac{1}{4}a^2\delta'_V\Theta^{VF}\frac{l(\eta_V)-l(\pi^0_V)}{\eta_V-\pi^0_V} \\ &\quad + (V\to A),\end{aligned} \tag{4.97}$$

$$\begin{aligned}\delta f^{\text{disc}}_{K^+_F} =& \frac{1}{2}l(U_I) - \frac{1}{4}l(\pi^0_I) \\ &+ \frac{1}{2}a^2\delta'_V\sum_X R^D_{U\pi^0\eta}(X_V)l(X_V) + (V\to A),\end{aligned} \tag{4.98}$$

and

$$\begin{aligned}\delta f^{\text{disc}}_{K^0_F} =& \frac{1}{2}l(D_I) - \frac{1}{4}l(\pi^0_I) \\ &+ \frac{1}{2}a^2\delta'_V\sum_X R^U_{D\pi^0\eta}(X_V)l(X_V) + (V\to A).\end{aligned} \tag{4.99}$$

One can simplify the results using the tree-level mass relations given in Ref. [94]:

$$R^D_{U\pi^0\eta}(U_B) = \frac{4}{a^2\delta'_B}, \tag{4.100}$$

$$R^D_{U\pi^0\eta}(\pi^0_B) = -\frac{2}{a^2\delta'_B}(1+\sin\beta_B), \tag{4.101}$$

$$R^D_{U\pi^0\eta}(\eta_B) = -\frac{2}{a^2\delta'_B}(1-\sin\beta_B), \tag{4.102}$$

$$R^U_{D\pi^0\eta}(D_B) = \frac{4}{a^2\delta'_B}, \tag{4.103}$$

$$R^U_{D\pi^0\eta}(\pi^0_B) = -\frac{2}{a^2\delta'_B}(1-\sin\beta_B), \tag{4.104}$$

$$R^U_{D\pi^0\eta}(\eta_B) = -\frac{2}{a^2\delta'_B}(1+\sin\beta_B), \tag{4.105}$$

$$R_{\pi^0\eta}(\pi^0_B) = \frac{2}{a^2\delta'_B}\cos\beta_B, \tag{4.106}$$

$$R_{\pi^0\eta}(\eta_B) = -\frac{2}{a^2\delta'_B}\cos\beta_B, \tag{4.107}$$

$$R^D_{\pi^0\eta}(\pi^0_B) = \frac{1}{2}\left(1+\sin\beta_B-\cos\beta_B\right), \tag{4.108}$$

$$R^D_{\pi^0\eta}(\eta_B) = \frac{1}{2}\left(1-\sin\beta_B+\cos\beta_B\right), \tag{4.109}$$

for $B = V, A$, where

$$\sin\beta_B \equiv (\operatorname{sgn}\delta'_B)\frac{D_B-U_B}{\sqrt{(D_B-U_B)^2+\frac{1}{4}(a^2\delta'_B)^2}} \tag{4.110}$$

$$\cos\beta_B \equiv (\operatorname{sgn}\delta'_B)\frac{\frac{1}{2}a^2\delta'_B}{\sqrt{(D_B-U_B)^2+\frac{1}{4}(a^2\delta'_B)^2}}. \tag{4.111}$$

The connected contributions in the fully dynamical 2+1 flavor case are

$$\delta f^{\rm con}_{\pi_F} = -\frac{1}{8}\sum_B g_B l(\pi_B), \tag{4.112}$$

$$\delta f^{\rm con}_{K_F} = -\frac{1}{16}\sum_B g_B l(\pi_B). \tag{4.113}$$

For the disconnected contributions in the fully dynamical 2+1 flavor case, we

find

$$\delta f_{\pi_F}^{\rm disc} = \frac{1}{2}(4-\Theta^{VF})\Big\{l(\pi_V)-l(\eta_V)\Big\} + (V\to A), \tag{4.114}$$

$$\delta f_{K_F}^{\rm disc} = \frac{1}{4}l(\pi_I)+l(\pi_V)-l(\eta_V)+(V\to A). \tag{4.115}$$

4.3.2.2. Partially quenched case

Considering x and y to be light quarks, the connected contributions to the decay constants in the partially quenched 1+1+1 flavor case can be obtained by dropping terms corresponding to strange sea quark loops from Eq. (4.72). Eqs. (4.84), (4.86) and (4.48) give the disconnected contributions:

$$\begin{aligned}
\delta f^{\rm disc}_{P_F^+,m_x\neq m_y} = \sum_Z \Bigg[& \frac{1}{4}\Big\{ D^{UD}_{X\pi^0,X}(Z_I)l(Z_I) \\
& \qquad + D^{UD}_{Y\pi^0,Y}(Z_I)l(Z_I) - 2R^{UD}_{XY\pi^0}(Z_I)l(Z_I)\Big\} \\
& + \frac{1}{4}a^2\delta'_V\Big\{2D^{UD}_{X\pi^0\eta,X}(Z_V)l(Z_V) \\
& \quad + 2D^{UD}_{Y\pi^0\eta,Y}(Z_V)l(Z_V) - \Theta^{VF}R^{UD}_{XY\pi^0\eta}(Z_V)l(Z_V)\Big\} + (V\to A)\Bigg] \\
& + \frac{1}{4}\Big\{R^{UD}_{X\pi^0}(X_I)\tilde{l}(X_I) + R^{UD}_{Y\pi^0}(Y_I)\tilde{l}(Y_I)\Big\} \\
& + \frac{1}{2}a^2\delta'_V\Big\{R^{UD}_{X\pi^0\eta}(X_V)\tilde{l}(X_V) + R^{UD}_{Y\pi^0\eta}(Y_V)\tilde{l}(Y_V)\Big\} \\
& + (V\to A),
\end{aligned} \tag{4.116}$$

and

$$\begin{aligned}
\delta f^{\rm disc}_{P_F^+,m_x=m_y} = & \frac{1}{4}a^2\delta'_V(4-\Theta^{VF})\Bigg[R^{UD}_{X\pi^0\eta}(X_V)\tilde{l}(X_V) + \sum_Z D^{UD}_{X\pi^0\eta,X}(Z_V)l(Z_V)\Bigg] \\
& + (V\to A).
\end{aligned} \tag{4.117}$$

The connected contributions to the decay constants in the partially quenched 2+1 flavor case can be obtained by setting $m_u = m_d$ in the 1+1+1 flavor case, Eq. (4.72). Setting $m_u = m_d$, in Eqs. (4.84) and (4.86), we find the disconnected

contributions in the 2+1 flavor case:

$$\begin{aligned}\delta f^{\rm disc}_{P_F,m_x\neq m_y} = \sum_Z \Bigg[&-\frac{1}{2}R^\pi_{XY}(Z_I)l(Z_I) \\ &+\frac{1}{4}a^2\delta'_V\Big\{2D^\pi_{X\eta,X}(Z_V)l(Z_V) \\ &\qquad +2D^\pi_{Y\eta,Y}(Z_V)l(Z_V)-\Theta^{VF}R^\pi_{XY\eta}(Z_V)l(Z_V)\Big\}+(V\to A)\Bigg] \\ &+\frac{1}{4}\Big\{l(X_I)+(\pi_I-X_I)\tilde{l}(X_I)+l(Y_I)+(\pi_I-Y_I)\tilde{l}(Y_I)\Big\} \\ &+\frac{1}{2}a^2\delta'_V\Big\{R^\pi_{X\eta}(X_V)\tilde{l}(X_V)+R^\pi_{Y\eta}(Y_V)\tilde{l}(Y_V)\Big\} \\ &+(V\to A),\end{aligned} \tag{4.118}$$

and

$$\begin{aligned}\delta f^{\rm disc}_{P_F,m_x=m_y} =&\frac{1}{4}a^2\delta'_V(4-\Theta^{VF})\Bigg[R^\pi_{X\eta}(X_V)\tilde{l}(X_V)+\sum_Z D^\pi_{X\eta,X}(Z_V)l(Z_V)\Bigg] \\ &+(V\to A).\end{aligned} \tag{4.119}$$

Considering x to be a light quark and y to be a heavy quark, the connected contributions to the decay constants is the same as those in the case of x and y are the light quarks. The strange quark and y can be decoupled by taking the limit $m_s, m_y \gg m_u, m_d, m_x$. From Eq. (4.84), we find the disconnected contribution for the partially quenched 1+1+1 flavor case,

$$\begin{aligned}\delta f^{\rm disc}_{P^+_F} =&\frac{1}{4}\sum_Z\Bigg[D^{UD}_{X\pi^0,X}(Z_I)l(Z_I)+2a^2\delta'_V D^{UD}_{X\pi^0\eta,X}(Z_V)l(Z_V)+(V\to A)\Bigg] \\ &+\frac{1}{4}R^{UD}_{X\pi^0}(X_I)\tilde{l}(X_I)+\frac{1}{2}a^2\delta'_V R^{UD}_{X\pi^0\eta}(X_V)\tilde{l}(X_V)+(V\to A).\end{aligned} \tag{4.120}$$

For the 2+1 flavor case, we find

$$\begin{aligned}\delta f^{\rm disc}_{P_F} =& \sum_Z\Bigg[\frac{1}{2}a^2\delta'_V D^\pi_{X\eta,X}(Z_V)l(Z_V)+(V\to A)\Bigg] \\ &+\frac{1}{4}\Big\{l(X_I)+(\pi_I-X_I)\tilde{l}(X_I)\Big\}+\frac{1}{2}a^2\delta'_V R^\pi_{X\eta}(X_V)\tilde{l}(X_V)+(V\to A).\end{aligned} \tag{4.121}$$

4.4. Conclusion

The general results for the decay constants are given compactly by Eq. (4.13) with Eqs. (4.45) through (4.46); they reduce to those of Ref. [49] in the taste Goldstone sector. The only new low-energy couplings are those parametrizing the analytic corrections proportional to a^2; the SO(4)-violating corrections are independent of those in the masses due to contributions from mesonic source operators [92]. The loop corrections respect $SO(4)_T$, which is broken only by the analytic terms $\propto a^2$.

Results for special cases of interest can be obtained by expanding the disconnected pieces of the propagators in Eq. (4.48). For the fully dynamical case with three non-degenerate quarks, the loop corrections in the SU(3) chiral theory are in Eqs. (4.51)-(4.61). Results in the isospin limit are in Eqs. (4.63)-(4.71). For the partially quenched case with three non-degenerate sea quarks, loop corrections in the SU(3) chiral theory are in Eqs. (4.72)-(4.86). Results in the isospin limit are in Eqs. (4.87)-(4.88). For the quenched case the results are in Eqs. (4.91)-(4.92). Results in SU(2) chiral perturbation theory are in Eqs. (4.94)-(4.115) and Eqs. (4.116)-(4.121). These results can be used to improve determinations of the decay constants, quark masses and the Gasser-Leutwyler LECs.

Part II

Kaon Mixing Matrix Elements from BSM Operators

5. Introduction to the Kaon Mixing Matrix Elements from BSM Operators

5.1. Kaon mixing matrix elements from the Standard Model

Flavor eigenstates or strong interaction eigenstate of the netral kaons are

$$K^0 = (\bar{s}d), \qquad \bar{K}^0 = (s\bar{d}). \tag{5.1}$$

They are antiparticles to each other and they have opposite strangeness: $S = 1$ for K^0 and $S = -1$ for $\bar{K}^0$. If there is no mixing between K^0 and $\bar{K}^0$, they whould have the same mass. Via weak interaction, however, strangeness is not conserved and K^0 and $\bar{K}^0$ can be mixed. Although they are eigenstates in strong interaction, Since CP operation (charge conjugation and parity) changes them to each other, we can write the CP eigenstates as follows:

$$K_\pm = \frac{1}{2}\big(K^0 \pm \bar{K}^0\big)/\sqrt{2}, \tag{5.2}$$

$$CP|K_\pm\rangle = \pm|K_\pm\rangle. \tag{5.3}$$

If the CP is an exact symmetry, the CP eigenstates $K_\pm$ are should be the Hamiltonian eigenstates. In nature, however, the weak interaction breaks CP symmetry and the Hamiltonian eigenstates are mixture of $K_\pm$:

$$K_S = \frac{K_+ + \bar{\epsilon}K_-}{\sqrt{1+|\bar{\epsilon}|^2}}, \qquad K_L = \frac{K_- + \bar{\epsilon}K_+}{\sqrt{1+|\bar{\epsilon}|^2}}, \tag{5.4}$$

where $\bar{\epsilon} \simeq \mathcal{O}(10^{-3})$. The subscripts S and L stand for "short" and "long", which corresponds to their lifetime: $\tau_S \sim 10^{-10}$s versus $\tau_L \sim 5 \times 10^{-8}$s.

Their hadronic decay modes are two pion states and three pion states. Two pion ($\pi^0\pi^0$, $\pi^+\pi^-$) $l = 0$ states are CP even while three pion states ($\pi^+\pi^-\pi^0$, $\pi^0\pi^0\pi^0$) are CP odd. Hence, if CP is exact symmetry($\bar{\epsilon} = 0$), K_S decays only to two pion states and K_L decays only to three pion states. This is the major reason of the lifetime difference in K_L and K_S. Since two pion final states has much larger phase space than that of three pion states, K_S decays much faster than K_S.

In real world, CP is not a exact symmetry and K_L can decay to two pion states. The $K_L \to \pi\pi$ can happen in two ways. First, the CP odd component

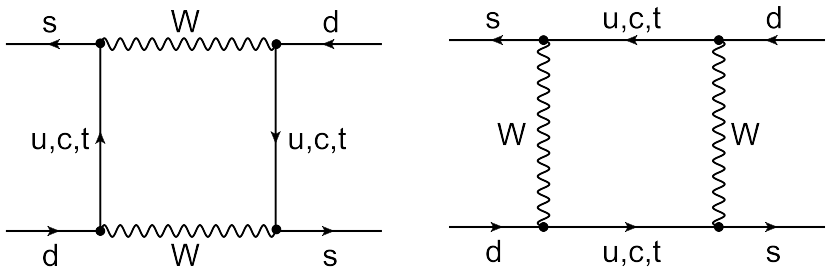


Figure 5.1: Box diagrams that are contributing to $K^0 - \bar{K}^0$ mixing in the Standard Model.

(K_-) of K_L can decay to the two pion states. This shows CP violation explicitly and it is called *direct CP violation*. The size of direct CP violation is parameterized by ϵ' defined as

$$\epsilon' = \frac{1}{\sqrt{2}} \left(\frac{A[K_L \to (\pi\pi)_2]}{A[K_S \to (\pi\pi)_2]} - \epsilon \frac{A[K_S \to (\pi\pi)_2]}{A[K_S \to (\pi\pi)_0]} \right) \tag{5.5}$$

where subscript 0 and 2 indicates the isospin of the final two pion states, A means the amplitude and ϵ is the size of indirect CP violation defined below. Another way of $K_L \to \pi\pi$ is the decay of CP even component $(\bar{\epsilon} K_+)$ in K_L to two pion states. Since both of them are CP even, it dose not seem to breaking CP symmetry. However, the existence of CP even component in K_L indicates the CP violation. Hence, it is called *indirect CP violation*. The size of indirect CP violation is parameterized by ϵ defined by

$$\epsilon = \frac{A[K_L \to (\pi\pi)_0]}{A[K_S \to (\pi\pi)_0]} \,. \tag{5.6}$$

K_L can have small CP even component through $K^0 - \bar{K}^0$ mixing. The $K^0 - \bar{K}^0$ mixing arises from the $\Delta S = 2$, $s\bar{d} \to \bar{s}d$ flavor changing neutral current(FCNC). It is responsible for indirect CP violation and the mass difference between K_L and K_S, ΔM_K. The $K^0 - \bar{K}^0$ mixing is dominated by the box digrams given in Fig. 5.1. At low energy, the W-boson of the box diagram can be integrated out and it can be replaced by a local, four-quark operator. The resulting effective Hamiltonian is given by [28]

$$H_{\text{eff}}^{\Delta S=2} = \frac{G_F^2 M_W^2}{16\pi^2} F^0 Q_1 + \text{h.c.}\,, \tag{5.7}$$

where G_F is the Fermi coupling, M_W is the W-boson mass, F^0 is a function

that contains the kinematic factors and CKM matrix elements:

$$F^0 = \lambda_c^2 S_0(x_c) + \lambda_t^2 S_0(x_t) + 2\lambda_c\lambda_t S_0(x_c, x_t)\,, \tag{5.8}$$

where $\lambda_a = V_{as}^* V_{ad}$, $a = c, t$, which is a flavor index, $x_a = m_a^2/M_W^2$ and $S_0(x_c)$, $S_0(x_t)$ and $S_0(x_c, x_t)$ are the Inami-Lim functions [101]. Q_1 is the four-quark operator defined by

$$Q_1 = [\bar{s}\gamma_\mu(1-\gamma_5)d][\bar{s}\gamma_\mu(1-\gamma_5)d]\,. \tag{5.9}$$

From this effective hamiltonian, the indirect CP violation, ϵ is predicted as

$$\epsilon = \exp(i\phi_\epsilon)\ \sqrt{2}\sin(\phi_\epsilon)\ C_\epsilon\ \mathrm{Im}\lambda_t\ X\ \widehat{B}_K + \xi\,, \tag{5.10}$$

where

$$\begin{aligned} X &= \mathrm{Re}\lambda_c[\eta_1 S_0(x_c) - \eta_3 S_3(x_c, x_t)] - \mathrm{Re}\lambda_t\eta_2 S_0(x_t) \\ C_\epsilon &= \frac{G_F^2 F_K^2 m_K M_W^2}{6\sqrt{2}\pi^2\Delta M_K} \\ \xi &= \exp(i\phi_\epsilon)\sin(\phi_\epsilon)\frac{\mathrm{Im}A_0}{\mathrm{Re}A_0}\,. \end{aligned} \tag{5.11}$$

Here $\widehat{B}_K$ is a RG(renormalization group) invariant form of the kaon bag parameter B_K, defined by

$$\widehat{B}_K = C(\mu)B_K(\mu) \tag{5.12}$$

$$B_K(\mu) = \frac{\sum_\nu \langle \bar{K}_0|[\bar{s}\gamma_\nu(1-\gamma_5)d][\bar{s}\gamma_\nu(1-\gamma_5)d]|K_0\rangle}{\frac{8}{3}\langle \bar{K}_0|\bar{s}\gamma_0\gamma_5 d|0\rangle\langle 0|\bar{s}\gamma_0\gamma_5 d|K_0\rangle}\,, \tag{5.13}$$

where $C(\mu)$ is the Wilson coefficient that makes $\widehat{B}_K$ RG invariant and μ is the renormalization scale. The $\widehat{B}_K$ contains all the non-perturbative QCD contributions for the ϵ and it can be calculated from the lattice simulations [46, 102].

5.2. Kaon mixing matrix elements from beyond the Standard Model

In the Standard Model, only the left handed quarks couple to the W-bosons, hence only the four-quark operator of the "left-left" Dirac structure form contribute to the $K^0 - \bar{K}^0$ mixing, given in Eq. (5.9). Considering beyond the Standard Model (BSM) physics, however, there are heavy particles that couple with the both left and right handed quarks.

For example, let us consider a supersymmetric (SUSY) model [103]. They give the gluino-mediated contribution to $\Delta S = 2$ transitions, shown in Fig. 5.2.

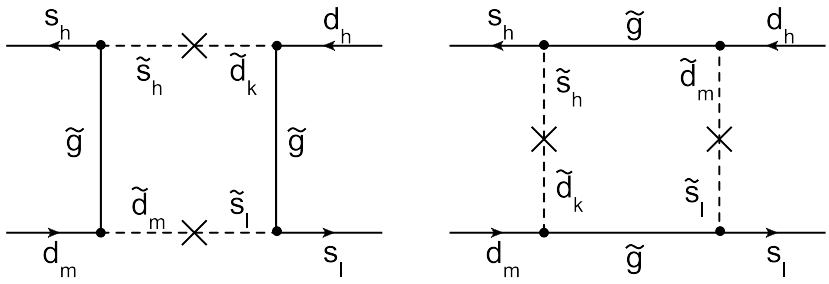


Figure 5.2: Box diagrams that are contributing to $K^0 - \bar{K}^0$ mixing in the SUSY Model. Here $h, k, l, m = \{L, R\}$.

In the process, both the left and right handed quarks contribute to the $K^0 - \bar{K}^0$ exchanging gnuinos and squarks.

Integrating out the heavy particles in the BSM, the resulting $\Delta S = 2$ four-fermion operators may have other than the "left-left" Dirac structure form. We adopt the operator basis used in perturbative calculations of anomalous dimensions [104] for the $\Delta S = 2$ four-fermion operators.

$$Q_2 = [\bar{s}^a(1-\gamma_5)d^a][\bar{s}^b(1-\gamma_5)d^b], \tag{5.14}$$
$$Q_3 = [\bar{s}^a\sigma_{\mu\nu}(1-\gamma_5)d^a][\bar{s}^b\sigma_{\mu\nu}(1-\gamma_5)d^b], \tag{5.15}$$
$$Q_4 = [\bar{s}^a(1-\gamma_5)d^a][\bar{s}^b(1+\gamma_5)d^b], \tag{5.16}$$
$$Q_5 = [\bar{s}^a\gamma_\mu(1-\gamma_5)d^a][\bar{s}^b\gamma_\mu(1+\gamma_5)d^b], \tag{5.17}$$

where $\sigma_{\mu\nu} = [\gamma_\mu, \gamma_\nu]/2$ and a, b are color indices. In the basis, the two-loop anomalous dimensions are known [104], which are used for the one-loop perturbative operator matching and the RG running. Note that this basis of operators differs slightly from the "SUSY basis" [105] used in other lattice calculations [106, 107]:

$$O_2 = \bar{s}^a(1-\gamma_5)d^a\bar{s}^b(1-\gamma_5)d^b, \tag{5.18}$$
$$O_3 = \bar{s}^a(1-\gamma_5)d^b\bar{s}^b(1-\gamma_5)d^a, \tag{5.19}$$
$$O_4 = \bar{s}^a(1-\gamma_5)d^a\bar{s}^b(1+\gamma_5)d^b, \tag{5.20}$$
$$O_5 = \bar{s}^a(1-\gamma_5)d^b\bar{s}^b(1+\gamma_5)d^a. \tag{5.21}$$

Including the BSM operators, the effective Hamiltonian is written as

$$H_{\rm eff}^{\Delta S=2} = \sum_{i=1}^{5} C_i Q_i \,, \tag{5.22}$$

where the coefficients C_i are determined once a model is specified. The BSM physics contributions to the kaon mixing matrix element are determined as the hadronic matrix elements of Q_{2-5} are calculated. Since the mixing of neutral kaon is well known in the experiments, calculating the hadronic matrix elements of Q_{2-5} can give strong constraints to the BSM physics.

As the B_K, it is convenient to calculate the B-parameters rather than the bare matrix elements $\langle \overline{K}_0 | Q_i | K_0 \rangle$. The B-parameters are define as

$$B_j = \frac{\langle \overline{K}_0 | Q_j | K_0 \rangle}{N_j \langle \overline{K}_0 | \bar{s}\gamma_5 d | 0 \rangle \langle 0 | \bar{s}\gamma_5 d | K_0 \rangle} \qquad \text{for } j = 2, 3, 4, 5, \tag{5.23}$$

$$(N_2,\ N_3,\ N_4,\ N_5) = (5/3,\ 4,\ -2,\ 4/3)\,. \tag{5.24}$$

Using the formula,

$$\langle \overline{K}_0 | \bar{s}\gamma_5 d | 0 \rangle \langle 0 | \bar{s}\gamma_5 d | K_0 \rangle = -\left(\frac{f_K M_K^2}{m_d + m_s} \right)^2, \tag{5.25}$$

it is possible to calculate the bare matrix elements $\langle \overline{K}_0 | Q_i | K_0 \rangle$ from the B-parameters.

6. Numerical Study of Kaon Mixing Matrix Elements from BSM Operators

6.1. Computation of BSM B-parameters

We use MILC ensembles generated with $N_f = 2 + 1$ asqtad staggered sea quarks, listed in Table 6.1. The number of ensembles and the mass of light sea quarks (am_l) and strange sea quarks(am_s) are given in the table. Details of generation and properties of lattices are given in Ref. [6]. To convert our data to physical unit, we use r_1/a given in Ref. [6]. For the r_1, we use $r_1 = 0.3117(6)(^{+12}_{-31})$ fm, which is obtained by the f_π analysis [88].

For the valence quarks, we use staggered quarks with HYP-smeared [63] gauge links. The parameters for the HYP smearing are chosen so that remove $\mathcal{O}(a^2)$ taste-symmetry breaking at tree level [108]. We use 10 different quark masses for the valence quarks (m_x, m_y):

$$m_{x,y} = m_s \times \frac{n}{10} \qquad \text{with } n = 1, 2, 3, \cdots, 10\,. \tag{6.1}$$

As explained in Ref. [46], the valence strange quark mass m_y is tuned so that the mass of $\bar{y}y$ meson gives 0.6858 GeV, which is the continuum $\bar{s}s$ meson mass mass [109], and m_x is tuned so that the mass of taste-Goldstone $\bar{x}y$ meson gives the physical K_0 mass. In Table 6.2, we show the valence quark masses (m_x and m_y) for the simulation and physical down and strange valence-quark masses ($m_d^{\rm phys}$ and $m_s^{\rm phys}$).

a (fm)	am_l/am_s	size	ens $\times$ meas	ID
0.12	0.01/0.05	$20^3 \times 64$	671×9	C3
0.09	0.0062/0.031	$28^3 \times 96$	995×9	F1
0.09	0.0093/0.031	$28^3 \times 96$	949×9	F3
0.06	0.0036/0.018	$48^3 \times 144$	749×9	S1
0.045	0.0028/0.014	$64^3 \times 192$	747×1	U1

Table 6.1: MILC lattices used for the numerical study. Here "ens" represents the number of gauge configurations, "meas" is the number of measurements per configuration, and ID will be used later to identify the corresponding MILC lattice.

Ensemble	$am_d^{\rm phys}$	$am_s^{\rm phys}$	am_x and am_y
C3	0.00213(2)	0.05204(5)	$0.005 \times n$
F1	0.00146(2)	0.03542(5)	$0.003 \times n$
S1	0.00104(1)	0.02372(3)	$0.0018 \times n$
U1	0.00076(1)	0.01693(3)	$0.0014 \times n$

Table 6.2: Valence quark masses (m_x and m_y) for the simulation and physical down and strange valence-quark masses ($m_d^{\rm phys}$ and $m_s^{\rm phys}$) in lattice units. Here $n = 1, 2, 3, \ldots, 10$.

The method of numerical calculation for the four-quark operator matrix elements is the same as that of B_K, explained elaborately in Ref. [46]. In terms of lattice operators, the B-parameters are

$$B_i(t) = \frac{2\langle \overline{K}^0_{P1} | z_{ij} \mathcal{O}_j^{\rm Lat}(t) | K^0_{P2} \rangle}{N_i \langle \overline{K}^0_{P1} | z_P \mathcal{O}_P^{\rm Lat}(t) | 0 \rangle \langle 0 | z_P \mathcal{O}_P^{\rm Lat}(t) | K^0_{P2} \rangle}, \tag{6.2}$$

where $\mathcal{O}_P^{\rm Lat}$ and $Q_j^{\rm Lat}$ are the lattice bilinear operator and the lattice four-fermion operators defined by

$$\mathcal{O}_P^{\rm Lat} = [P \times P], \tag{6.3}$$

and

$$\mathcal{O}_{S1}^{\rm Lat} = [S \times P][S \times P]_I \,, \tag{6.4}$$
$$\mathcal{O}_{S2}^{\rm Lat} = [S \times P][S \times P]_{II} \,, \tag{6.5}$$
$$\mathcal{O}_{P1}^{\rm Lat} = [P \times P][P \times P]_I \,, \tag{6.6}$$
$$\mathcal{O}_{P2}^{\rm Lat} = [P \times P][P \times P]_{II} \,, \tag{6.7}$$
$$\mathcal{O}_{V1}^{\rm Lat} = [V_\mu \times P][V_\mu \times P]_I \,, \tag{6.8}$$
$$\mathcal{O}_{V2}^{\rm Lat} = [V_\mu \times P][V_\mu \times P]_{II} \,, \tag{6.9}$$
$$\mathcal{O}_{A1}^{\rm Lat} = [A_\mu \times P][A_\mu \times P]_I \,, \tag{6.10}$$
$$\mathcal{O}_{A2}^{\rm Lat} = [A_\mu \times P][A_\mu \times P]_{II} \,, \tag{6.11}$$
$$\mathcal{O}_{T1}^{\rm Lat} = \sum_{\mu<\nu} [T_{\mu\nu} \times P][T_{\mu\nu} \times P]_I \,, \tag{6.12}$$
$$\mathcal{O}_{T2}^{\rm Lat} = \sum_{\mu<\nu} [T_{\mu\nu} \times P][T_{\mu\nu} \times P]_{II} \,. \tag{6.13}$$

Here the subscript I and II represents one color trace operators and two color trace operators, respectively [110], and $[S \times T]$ is the spin and taste structure of

ensemble	Δt	$\Delta t/T$	t_L	t_R	t_L (fm)
C3	26	0.41	10	15	1.19
F1	40	0.42	14	25	1.18
S1	60	0.42	22	37	1.29
U1	80	0.42	26	53	1.14

Table 6.3: Choices for the wall-source separation, Δt, and its ratio to the temporal length of the lattices, T, as well as the parameters determining the fitting range.

the bilinear and four-fermion operators (see Eq. (3.23)). z_P and z_{ij} are matching factors for the bilinear and four-fermion operators. The one-loop matching factor for the pseudosclar bilinear is

$$z_P = 1 + \frac{\alpha}{4\pi}\left(8\log(\mu a) + \frac{10}{3} - 1.57938\right). \tag{6.14}$$

Operator matching for the four-fermion operators at one-loop level is defined by

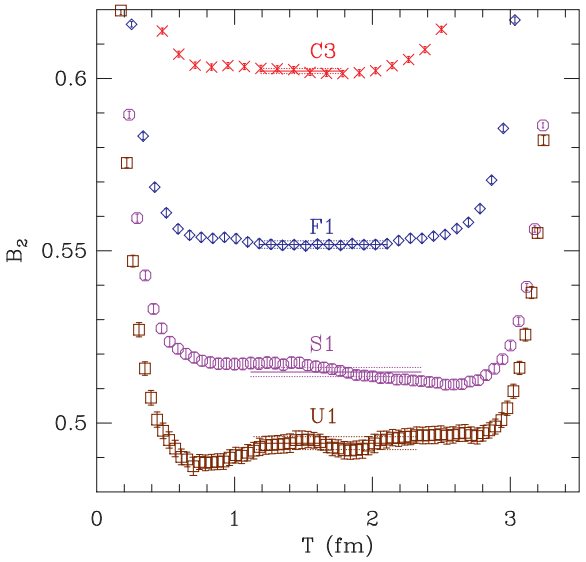
$$\mathcal{O}_i^{\mathrm{Cont}'} = \sum_{j\in(A)} z_{ij}\mathcal{O}_j^{\mathrm{Lat}}, \tag{6.15}$$

with $i = 2, 3, 4, 5$, and

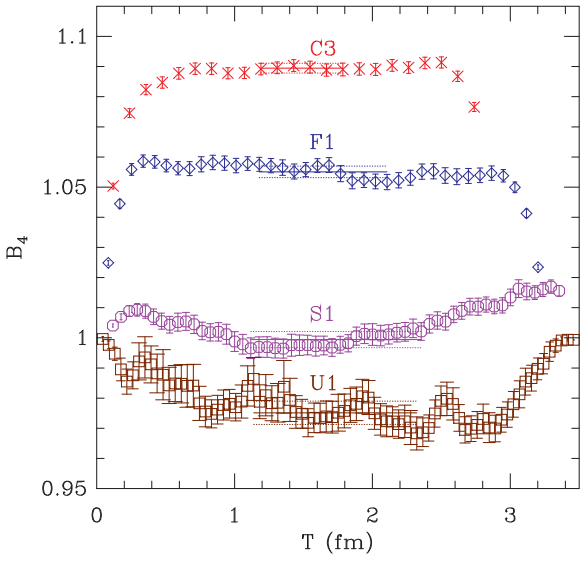
$$z_{ij} = b_{ij} + \frac{g^2}{(4\pi)^2}\Big(- \gamma_{ij}\log(\mu a) + d_{ij}^{\mathrm{Cont}} - d_{ij}^{\mathrm{Lat}} - C_F I_{MF} T_{ij}\Big). \tag{6.16}$$

To produce kaon and anti-kaon, we place U(1)-noise wall-sources at timeslice t_1 and $t_2 > t_1$. They produce kaon and anti-kaon with taste-P and having zero momentum. The four-quark operators are placed at t between t_1 and t_2 (i.e. $t_1 < t < t_2$) and contracted with the kaon and anti-kaon. The four-quark operators are expected to be independent of t when t is far enough from the sources at t_1 and t_2 so that the contamination from the excited states are negligible. Hence we find suitable range of t, which shows a plateau, and fit with constant in the region. In Fig. 6.1, we show the sample plots of the $B_2(\mu = 1/a)$ and $B_4(\mu = 1/a)$ as a function of $T = t - t_1$. The constant fit is performed ignoring the correlation between timeslices (diagonal approximation for the covariance matrix) to avoid instability of the fit come from the small eigenvalues of the covariance matrix [53]. The fitting errors are estimated by the jackknife method.

To determine the fit range, we use two-point correlator from the wall-sources to the taste-P axial current. From the effective mass plots for the two-point correlator, we find t_L that removes the contamination of excitation states. Then,



(a) B_2



(b) B_4

Figure 6.1: $B_{2-5}(\mu = 1/a)$ as a function of $T = t - t_1$. (Red) crosses are the results on C3 ensemble, with $(am_x, am_y) = (0.005, 0.05)$; and (blue) diamonds are the results on F1 ensemble, with $(am_x, am_y) = (0.003, 0.03)$; and (purple) octagons are the results on S1 ensemble, with $(am_x, am_y) = (0.0018, 0.018)$; and (brown) squares are the results on U1 ensemble, with $(am_x, am_y) = (0.0014, 0.014)$.

we fit from $t = t_1 + t_L$ to $t = t_1 + t_R = t_2 - t_L - 1$. Our choice of t_L and t_R is given in Table. 6.3. Detailed description for this procedure is given in Ref. [46].

To reduce the statistical error, we measure multiple times on each gauge configuration. The source position t_1 is chosen randomly, and t_2 is determined by $t_2 = t_1 + \Delta t$. The number of measurements for each gauge configuration is given in Table 6.1.

We find that the autocorrelation gets larger as the lattice spacing gets finer. As one can see in Fig. 6.2, the autocorrelation effect is about 100% for the MILC superfine lattice (S1) while that is about 25% for the MILC fine lattice (F1). In order to remove the autocorrelation effect, we use bin size = 5 to do the data analysis.

As discussed in Ref. [46], the wrap-around contributions for the kaons are negligible. For more detailed description for the calculation of B-parameters, see Ref. [46, 111].

6.2. SU(2) fitting

Next step is the chiral extrapolation of the quark masses to the physical down and strange quark masses [112, 113] using the formula given in Refs. [102, 114]. Here we use the SU(2) fitting, which extrapolates m_x to the $m_d^{\rm phys}$ fixing m_y close to $m_s^{\rm phys}$, and then extrapolates m_y to the $m_s^{\rm phys}$. For the extrapolation of m_x to the $m_d^{\rm phys}$, we use SU(2) staggered chiral perturbation theory (SχPT), which requires $m_x \ll m_y$. Hence we take lightest four quark masses for the m_x (e.g. $m_x = \{0.005,\ 0.01,\ 0.015,\ 0.02\}$ on the coarse ensemble) and heaviest three quark masses for the m_y (e.g. $m_i = \{0.04,\ 0.045,\ 0.05\}$ on the coarse ensemble) from the 10 different quark masses, given in Eq. (6.1). This choice of quark masses guarantees $m_x/m_y \le 1/2$.

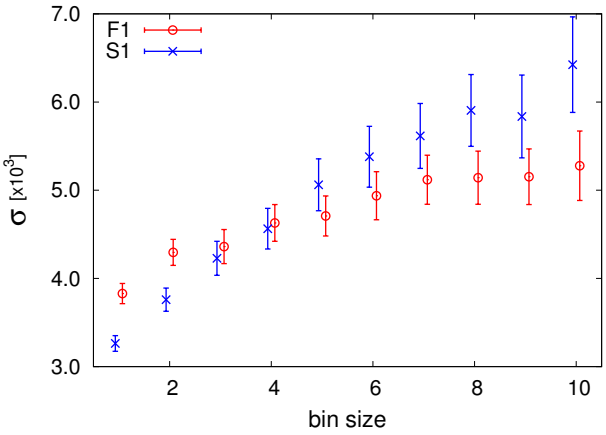
Let us consider the extrapolation of m_x. We call it the X-fit. For an ensemble and fixed m_y, the fitting functional form is simplified as follows:

$$B_j(\mathrm{NLO}) = c_1 F_0(j) + c_2 X \tag{6.17}$$

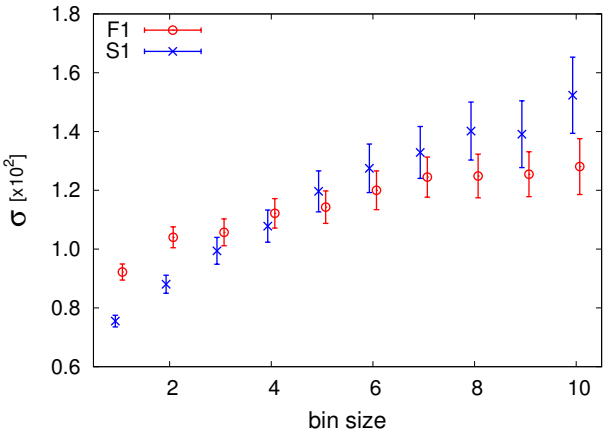
where $X \equiv \dfrac{X_P}{\Lambda^2}$, $\Lambda = 1\,\mathrm{GeV}$, and

$$F_0(j) = 1 \pm \frac{1}{32\pi^2 f^2}\left\{ \ell(X_I) + (L_I - X_I)\tilde{\ell}(X_I) - \frac{1}{16}\sum_B \ell(X_B) \right\}. \tag{6.18}$$

Here the upper sign applies for $j = 2, 3$, and the lower sign applies for $j = 4, 5$. B_K also uses the same fitting function with the upper sign. However, it turns out that the NLO fitting function is not accurate enough to describe the precise and highly correlated data. The fitting with $B_j(\mathrm{NLO})$ gives large χ^2 values.



(a) $[P \times P][P \times P]_I$



(b) $[P \times P][P \times P]_{II}$

Figure 6.2: Statistical error of raw data as a function of bin size. The operators are (a) $\mathcal{O}_{P1}^{\text{Lat}}$, and (b) $\mathcal{O}_{P2}^{\text{Lat}}$, respectively, at $t = 20$ (F1) and $t = 30$ (S1). (Red) circles are the results on F1 ensemble, with $(am_x, am_y) = (0.003, 0.03)$; and (blue) crosses are the results on S1 ensemble, with $(am_x, am_y) = (0.0018, 0.018)$.

ID	C3	F1	F3	S1	U1
c_1	0.5806(97)	0.4954(83)	0.4909(91)	0.4833(75)	0.4726(115)
c_2	−0.81(32)	0.27(31)	0.10(34)	−0.31(30)	−0.02(41)
c_3	0.88(29)	−0.23(26)	−0.24(29)	−0.03(22)	−0.24(34)
c_4	0.78(29)	−0.07(31)	0.21(34)	0.63(32)	0.30(42)
c_5	−0.55(20)	0.04(19)	−0.13(21)	−0.33(17)	−0.17(24)
c_6	−0.185(69)	0.013(56)	−0.038(62)	−0.088(46)	−0.050(70)
χ^2	1.11(82)	0.01(08)	0.08(24)	0.61(62)	0.21(60)
B_K	0.5755(53)	0.5070(42)	0.5010(46)	0.4873(36)	0.4804(62)

Table 6.4: Fit parameters for B_K X-fit. The valence strange quark mass is fixed at the heaviest value, i.e. $am_y = 0.05, 0.03, 0.018$ and 0.014 on coarse, fine, superfine and ultrafine lattices, respectively. Here χ^2 is the $\chi^2_{\rm aug}$/dof, and B_K are evaluated at $\mu = 1/a$.

Hence we added higher order terms to the fitting function:

$$B_j(\mathrm{NNNLO}) = c_1 F_0(j) + c_2 X + c_3 X^2 + c_4 X^2 \big(\ln(X)\big)^2 + c_5 X^2 \ln(X) + c_6 X^3. \tag{6.19}$$

The three terms $X^2, X^2\big(\ln(X)\big)^2$ and $X^2\ln(X)$ are the generic NNLO terms in the continuum chiral perturbation theory. We also included a single NNNLO term X^3. Since we have only four data points for the X-fit, we use Bayesian method to constrained the three fitting parameters c_{4-6} [115]. Our prior knowledge for the coefficients is that they are $\mathcal{O}(1)$ parameters. Hence, we constrained them to be $c_{4-6} = 0 \pm 1$, which means that the expectation value and standard deviation of the prior distribution for the c_{4-6} is 0 and 1, respectively. For the Bayesian fitting, the augmented χ^2 is defined by

$$\chi^2_{\rm aug} = \chi^2 + \chi^2_{\rm prior} \tag{6.20}$$

$$\chi^2_{\rm prior} = \sum_{i=4}^{6} \frac{(c_i - a_i)^2}{\sigma_i^2}\,, \tag{6.21}$$

where we set $a_i = 0$ and $\sigma_i = 1$. Then we minimize the $\chi^2_{\rm aug}$ to find the values for the fitting parameters, c_{1-6}. To estimate the systematic error of the Bayesian fitting in the X-fit, we doubled the prior width.

After fitting the data with the fitting function, Eq. (6.19), we extrapolate it to the physical point by setting $L_I = m^2_{\pi_0}$ and $m_x = m_d^{\rm phys}$. We also remove the taste symmetry breaking discretization error in sea and valence quarks ($\Delta_B = 0$) and set $X_B = 2m^2_{K_0,\rm phys} - m^2_{s\bar{s},\rm phys}$ [46]. In Figs. 6.3 and 6.4, we show sample

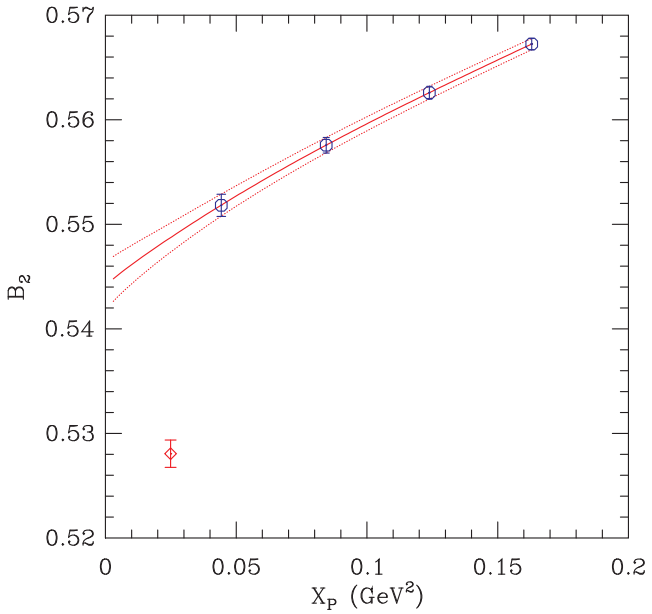


Figure 6.3: $B_2(\mu = 1/a)$ from the NNNLO Bayesian X-fit vs. X_P, on the fine ensemble F1, for $am_y = 0.03$. (Red) diamond is the extrapolation point. It is shifted from the fitting curve as the taste symmetry breaking discretization error is removed after the extrapolation.

ID	C3	F1	F3	S1	U1
c_1	0.5641(16)	0.5109(22)	0.5027(19)	0.4964(33)	0.4756(45)
c_2	0.51(05)	0.63(07)	0.79(06)	−0.24(12)	−0.19(16)
c_3	−0.57(04)	−0.69(06)	−0.85(05)	0.01(09)	0.00(13)
c_4	−0.30(04)	−0.41(06)	−0.52(06)	0.43(12)	0.37(16)
c_5	0.21(3)	0.25(4)	0.31(4)	−0.23(7)	−0.21(9)
c_6	0.071(10)	0.075(11)	0.095(11)	−0.061(17)	−0.062(27)
χ^2	0.14(04)	0.24(07)	0.37(08)	0.25(14)	0.19(17)
B_2	0.5801(10)	0.5281(13)	0.5226(12)	0.5005(17)	0.4800(24)

Table 6.5: Fit parameters for B_2 X-fit. Notation is the same as in Table 6.4

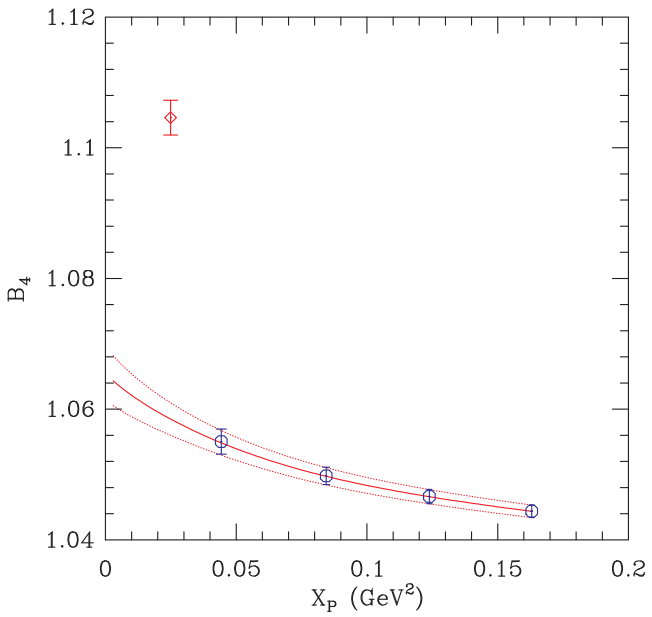


Figure 6.4: $B_4(\mu = 1/a)$ from the NNNLO Bayesian X-fit vs. X_P, on the fine ensemble F1, for $am_y = 0.03$. Notation is the same as in Fig. 6.3.

ID	C3	F1	F3	S1	U1
c_1	0.3971(11)	0.3673(15)	0.3611(13)	0.3577(23)	0.3426(31)
c_2	0.33(03)	0.45(04)	0.56(04)	−0.16(08)	−0.13(11)
c_3	−0.37(3)	−0.48(4)	−0.60(4)	0.02(6)	0.00(9)
c_4	−0.19(03)	−0.30(04)	−0.37(04)	0.30(09)	0.26(11)
c_5	0.13(2)	0.18(2)	0.23(3)	−0.16(5)	−0.15(6)
c_6	0.045(07)	0.054(07)	0.068(08)	−0.042(12)	−0.043(19)
χ^2	0.05(2)	0.13(4)	0.19(4)	0.12(7)	0.09(8)
B_3	0.4079(07)	0.3796(09)	0.3753(08)	0.3608(12)	0.3459(17)

Table 6.6: Fit parameters for B_3 X-fit. Notation is the same as in Table 6.4

ID	C3	F1	F3	S1	U1
c_1	1.1603(37)	1.1397(44)	1.1727(46)	1.0374(68)	1.0201(90)
c_2	−0.90(10)	−1.33(13)	−2.11(13)	0.27(22)	−0.15(28)
c_3	1.16(09)	1.65(12)	2.42(12)	0.42(18)	0.81(24)
c_4	0.60(09)	1.05(12)	1.66(12)	−0.43(23)	0.01(28)
c_5	−0.43(06)	−0.64(07)	−1.02(08)	0.23(12)	−0.01(17)
c_6	−0.144(21)	−0.191(22)	−0.310(23)	0.061(33)	−0.002(48)
χ^2	0.57(16)	1.57(36)	3.99(59)	0.26(27)	0.00(02)
B_4	1.1312(23)	1.1046(27)	1.1243(29)	1.0280(36)	1.0050(51)

Table 6.7: Fit parameters for B_4 X-fit. Notation is the same as in Table 6.4

ID	C3	F1	F3	S1	U1
c_1	0.9583(31)	0.9598(39)	0.9890(41)	0.8802(59)	0.8706(81)
c_2	−0.68(08)	−1.06(11)	−1.73(11)	0.23(19)	−0.15(26)
c_3	0.96(08)	1.40(10)	2.08(10)	0.42(15)	0.76(22)
c_4	0.51(07)	0.87(11)	1.39(11)	−0.32(20)	0.09(26)
c_5	−0.36(05)	−0.53(07)	−0.86(07)	0.17(11)	−0.05(15)
c_6	−0.122(18)	−0.160(20)	−0.261(20)	0.046(28)	−0.015(45)
χ^2	0.40(12)	1.09(27)	2.80(44)	0.14(17)	0.01(08)
B_5	0.9358(20)	0.9317(24)	0.9495(26)	0.8728(31)	0.8580(46)

Table 6.8: Fit parameters for B_5 X-fit. Notation is the same as in Table 6.4

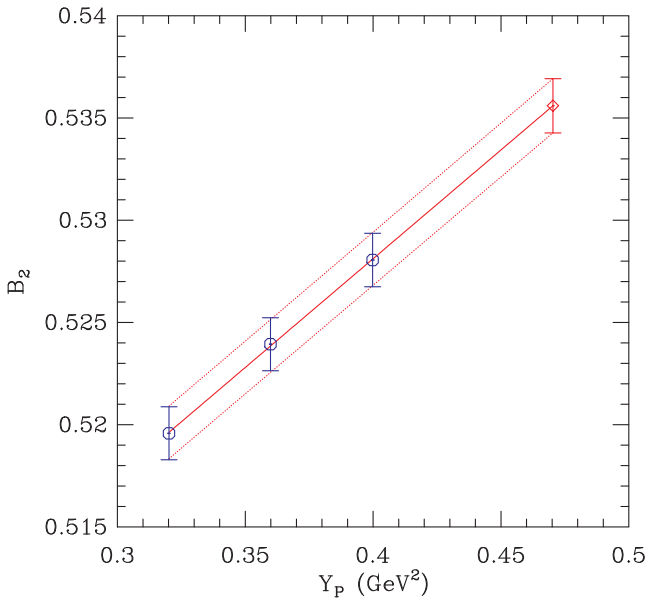


Figure 6.5: $B_2(\mu = 1/a)$ from the Y-fit vs. Y_P, on the fine ensemble F1. (Red) diamond is the linear extrapolation to the physical strange quark mass.

plots for the NNNLO Bayesian X-fit for the B_2 and B_4 on the fine ensemble F1. The extrapolation points are shifted from the fitting curve as the taste symmetry breaking discretization error is removed after the extrapolation to the physical down quark mass. The fitting results are collected in Tables. 6.4 – 6.8. Here we show the parameters for the BSM B-parameters as well as the for B_K.

Subsequent to the X-fit for the three heaviest values of m_y, we perform the chiral extrapolation to the physical strange quark masses varying $Y_P \propto m_y$. We call it the Y-fit. We expect that the B-parameters are smooth, analytic

ID	b_1	b_2	$\chi^2_{\rm aug}$/dof	$B_K(\mu = 1/a)$
C3	0.5148(71)	0.1345(120)	0.0035(11)	0.5780(53)
F1	0.4234(61)	0.2097(114)	0.0082(18)	0.5220(43)
F3	0.4305(68)	0.1759(135)	0.0050(15)	0.5132(48)
S1	0.4035(48)	0.2340(083)	0.0095(19)	0.5135(37)
U1	0.3894(85)	0.2329(172)	0.0041(14)	0.4989(65)

Table 6.9: Fit parameters for B_K Y-fit. The χ^2/dof is the uncorrelated χ^2 per degrees of freedom.

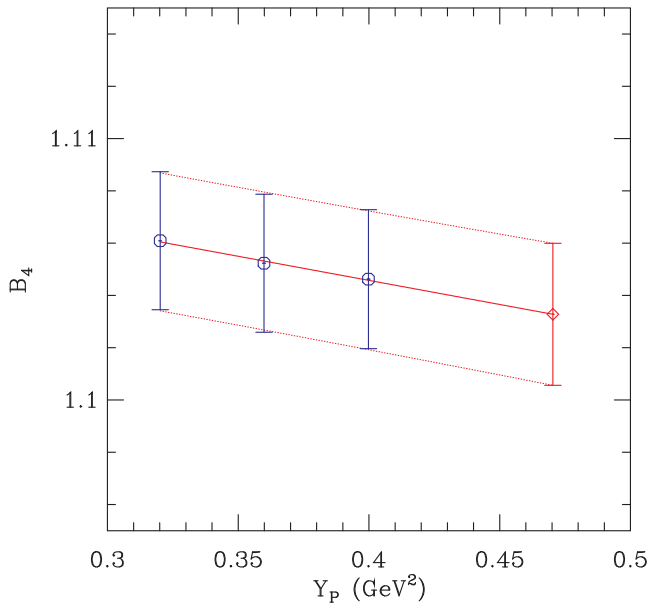


Figure 6.6: $B_4(\mu = 1/a)$ from the Y-fit vs. Y_P, on the fine ensemble F1. Notation is the same as in Fig. 6.5.

ID	b_1	b_2	χ^2/dof	$B_2(\mu = 1/a)$
C3	0.5358(09)	0.0980(20)	0.0146(13)	0.5819(10)
F1	0.4856(14)	0.1064(20)	0.0069(05)	0.5356(13)
F3	0.4808(12)	0.1045(22)	0.0102(09)	0.5299(12)
S1	0.4595(21)	0.1145(19)	0.0028(03)	0.5133(16)
U1	0.4385(26)	0.1061(48)	0.0024(04)	0.4884(25)

Table 6.10: Fit parameters for B_2 Y-fit. Notation is the same as in Table 6.9

ID	b_1	b_2	χ^2/dof	$B_3(\mu = 1/a)$
C3	0.3773(06)	0.0677(14)	0.0063(8)	0.4092(07)
F1	0.3511(10)	0.0713(14)	0.0030(4)	0.3846(09)
F3	0.3467(08)	0.0713(15)	0.0052(7)	0.3802(08)
S1	0.3319(14)	0.0805(13)	0.0013(2)	0.3698(12)
U1	0.3147(17)	0.0799(33)	0.0018(4)	0.3522(17)

Table 6.11: Fit parameters for B_3 Y-fit. Notation is the same as in Table 6.9

ID	b_1	b_2	χ^2/dof	$B_4(\mu = 1/a)$
C3	1.1505(23)	-0.0428(48)	0.0045(8)	1.1303(23)
F1	1.1120(31)	-0.0185(48)	0.0016(4)	1.1033(27)
F3	1.1295(35)	-0.0130(60)	0.0012(4)	1.1234(30)
S1	1.0151(45)	0.0359(50)	0.0002(1)	1.0320(35)
U1	0.9664(60)	0.0987(98)	0.0002(2)	1.0128(53)

Table 6.12: Fit parameters for B_4 Y-fit. Notation is the same as in Table 6.9

ID	b_1	b_2	χ^2/dof	$B_5(\mu = 1/a)$
C3	0.9278(20)	0.0174(42)	0.0080(10)	0.9360(20)
F1	0.9201(27)	0.0287(44)	0.0040(06)	0.9336(24)
F3	0.9357(30)	0.0341(53)	0.0035(06)	0.9518(27)
S1	0.8465(39)	0.0732(44)	0.0009(02)	0.8809(30)
U1	0.8046(55)	0.1367(93)	0.0001(01)	0.8689(47)

Table 6.13: Fit parameters for B_5 Y-fit. Notation is the same as in Table 6.9

functions of Y_P. Empirically, the linear fitting function explains the data well:

$$B_j(\text{Y-fit}) = b_1 + b_2 Y_P \,. \tag{6.22}$$

In Figs. 6.5 and 6.6 we show sample plots for the Y-fit. To avoid the small eigenvalue problem, we use uncorrelated fitting for the Y-fit. The fitting results are given in Tables. 6.9 – 6.13. In all cases, the fits are stable: values of the fit parameters are consistent through the different ensembles and small (uncorrelated) χ^2.

6.3. RG evolution

After the chiral extrapolation (X-fit and Y-fit), we have the B-parameters at different lattices with renormalization scale $\mu = 1/a$. In order to compare the results and extrapolate them to the continuum limit, we need to run the results to a common scale. In this paper, we run the results to 2 GeV and 3 GeV.

In the running, operator mixing arises in pairs:

$$(Q_2^{\text{Cont}}, Q_3^{\text{Cont}}) \text{ and } (Q_4^{\text{Cont}}, Q_5^{\text{Cont}}) \,. \tag{6.23}$$

Because it is operators that mix, and not the B-parameters themselves, we need to first bring the operators over a common denominator. Thus we define

$R_i(\mu) = N_i B_i(\mu)$ to remove the N_i dependence of the denominator of Eq. (5.23), and consider evolution of the $R_i(\mu)$ from $\mu = \mu_a$ to $\mu = \mu_b$,

$$B_j(\mu_b) = \sum_k \frac{1}{N_j} W^R(\mu_b, \mu_a)_{jk} N_k B_k(\mu_a)\,. \tag{6.24}$$

Since both the numerator and denominator evolve in R_i depend on μ and evolve independently, we can write the evolution kernel $W^R(\mu_b, \mu_a)$

$$W^R(\mu_b, \mu_a)_{jk} = \frac{W^Q(\mu_b, \mu_a)_{jk}}{[W^P(\mu_b, \mu_a)]^2}\,, \tag{6.25}$$

where where W^Q describes the evolution of the numerator, while W^P describes the evolution of the pseudoscalar density. The two-loop formula for the evolution kernels is

$$W(\mu_b, \mu_a) \approx \left[1 + \frac{\alpha_b}{4\pi} J\right]^{-1} V^{-1} \left(\frac{\alpha_b}{\alpha_a}\right)^{\gamma_D^{(0)}/2\beta_0} V \left[1 + \frac{\alpha_a}{4\pi} J\right], \tag{6.26}$$

$$J = \frac{\beta_1 \gamma^{(0)}}{2\beta_0^2} - V^{-1} M V, \tag{6.27}$$

$$M_{jk} = \frac{\left[V \gamma^{(1)} V^{-1}\right]_{jk}}{2\beta_0 + (\gamma_D^{(0)})_{kk} - (\gamma_D^{(0)})_{jj}}\,, \tag{6.28}$$

where V is the matrix that diagonalizes $\gamma^{(0)}$:

$$V \gamma^{(0)} V^{-1} = \gamma_D^{(0)}\,, \tag{6.29}$$

and $\left(\frac{\alpha_b}{\alpha_a}\right)^{\gamma_D^{(0)}/2\beta_0}$ is a diagonal matrix defined by

$$\left[\left(\frac{\alpha_b}{\alpha_a}\right)^{\gamma_D^{(0)}/2\beta_0}\right]_{ij} = \delta_{ij} \left(\frac{\alpha_b}{\alpha_a}\right)^{(\gamma_D^{(0)})_{ii}/2\beta_0}\,. \tag{6.30}$$

The expressions for the β_i and $\gamma^{(i)}$ of W^Q are given in Ref. [104].

In the case of $Q_{4,5}$ pair, the solution Eq. (6.26) has removable singularity when the denominator of Eq. (6.28) vanishes. In the singular case, we obtain the evolution kernel using the analytic continuation prescription introduced in Ref. [116]:

$$W(\mu_b, \mu_a) \approx V^{-1} \left(\frac{\alpha_b}{\alpha_a}\right)^{\gamma_D^{(0)}/2\beta_0} V + \frac{1}{4\pi} V^{-1} A V\,, \tag{6.31}$$

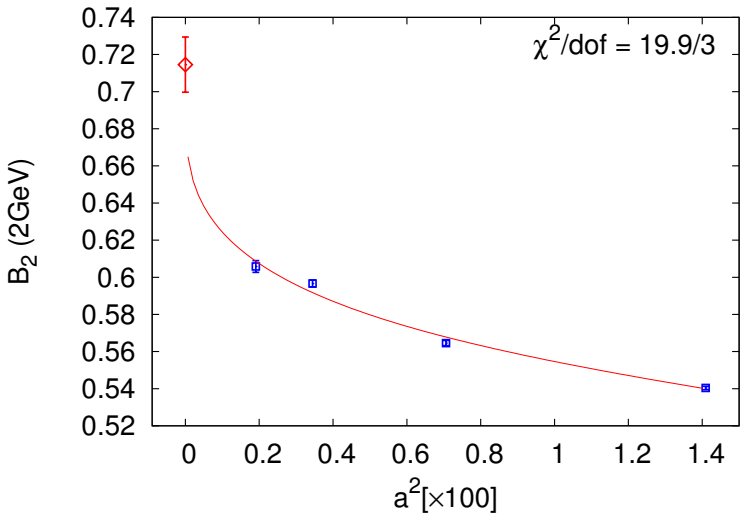


Figure 6.7: Continuum extrapolation of $B_2(\mu = 2GeV)$ including C3 ensemble, using the Bayesian constrained fitting with the fitting function given in Eq. (6.33)

where

$$A_{ij} = \frac{[V\gamma^{(1)}V^{-1}]_{ij}}{2\beta_0}\alpha_b \left(\frac{\alpha_b}{\alpha_a}\right)^{(\gamma_D^{(0)})_{jj}/2\beta_0} \ln\left(\frac{\alpha_b}{\alpha_a}\right) . \tag{6.32}$$

6.4. Continuum extrapolation

The final step is the continuum extrapolation of the results. The leading a and α_s dependence of B-parameters are known to be [102]:

$$B_j = d_1 + d_2(a\Lambda)^2 + d_3(a\Lambda)^2\alpha_s + d_4\alpha_s^2 + d_5(a\Lambda)^4 + \cdots , \tag{6.33}$$

where $\alpha_s = \alpha_s^{\overline{\rm MS}}(1/a)$. We perform continuum extrapolation of BSM B-parameters calculated on four different lattices, C3 – U1, with this form of 5-parameters taking $\Lambda = 300$MeV and constraining d_{2-5} using the Bayesian method with prior $d_i = 0 \pm 2$. However, it gives poor fitting results with large $\chi^2_{\rm aug}/{\rm dof} = 6.6 \sim 30$ for B_{2-5}. In Fig. 6.7 we show the continuum extrapolation of B_2 including the coarse lattice.

Thus we drop the results from the coarse lattice and fit for the three ensembles F1, S1 and U1. In this case, both the linear fitting (with $d_{1,2}$) and the constrained fitting (with d_{1-5}) work very well. In Fig. 6.8, we show an example of the continuum extrapolation for B_2. In Tables 6.14 and 6.15, we quote

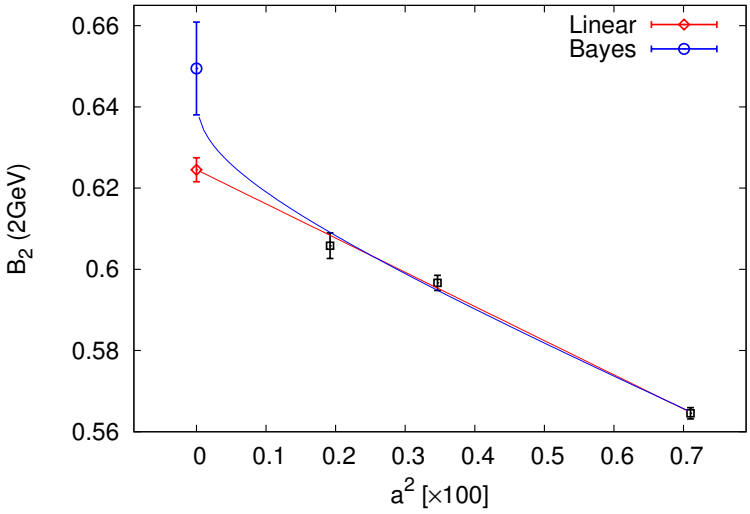


Figure 6.8: Continuum extrapolation of B_2 at 2GeV. (Red) diamond is the result from the linear fitting function; and (blue) circle is the result from the Bayesian constrained fitting with the fitting function given in Eq. (6.33).

the results from the linear extrapolation.

6.5. Conclusion

We presented a calculation of BSM B-parameters with HYP-smeared improved staggered fermions on the MILC asqtad lattices with $N_f = 2 + 1$ flavors. We use three different lattice spacings ($a \approx 0.045$, 0.06 and 0.09 fm) to obtain the continuum results. Chiral and continuum extrapolations are done using SU(2) staggered chiral perturbation theory. The preliminary results evaluated at 2 GeV and 3 GeV are reported in the Tables 6.14 and 6.15.

The next stage in our calculation is to quantify all sources of systematic error and so draw up a complete error budget. This requires results at other values of the light sea-quark masses to estimate residual m_ℓ dependence, and at other volumes to estimate finite volume effects. The latter can also be estimated using SChPT, and are expected to be small. We also plan to investigate whether the use of ratios which cancel chiral logarithms reduces errors in the analysis, and to compare the results to those from an analysis using SU(3) SChPT. We expect that, as for B_K, our dominant errors will come from to the use of one-loop matching and the continuum extrapolation.

Table 6.14: Preliminary results for BSM B-parameters and B_K at $\mu = 2\,\text{GeV}$. Continuum values are obtained using linear extrapolation. Only statistical errors are shown.

$B_j \diagdown$ Lat	C3	F1	S1	U1	Continuum
B_K	0.5672(52)	0.5295(43)	0.5362(38)	0.5318(70)	0.5383(66)
B_2	0.5404(09)	0.5646(14)	0.5967(19)	0.6058(31)	0.6245(30)
B_3	0.3689(06)	0.4148(10)	0.4594(14)	0.4805(24)	0.5032(22)
B_4	1.0965(23)	1.1260(28)	1.0911(37)	1.0942(57)	1.0698(56)
B_5	0.9278(20)	0.9381(25)	0.8875(31)	0.8720(49)	0.8432(48)

Table 6.15: Preliminary results for BSM B-parameters and B_K at $\mu = 3\,\text{GeV}$. Notation as in Table 6.14.

$B_j \diagdown$ Lat	C3	F1	S1	U1	Continuum
B_K	0.5478(50)	0.5114(42)	0.5179(37)	0.5137(67)	0.5199(64)
B_2	0.4779(08)	0.4993(12)	0.5277(17)	0.5358(28)	0.5524(26)
B_3	0.3152(05)	0.3496(08)	0.3840(12)	0.3997(20)	0.4174(19)
B_4	1.0462(22)	1.0750(26)	1.0421(36)	1.0452(55)	1.0222(54)
B_5	0.9132(19)	0.9272(24)	0.8824(31)	0.8714(48)	0.8450(47)

Part III

Art of Data Analysis

7. Basic Probability Theory

Physical experiments or Monte Carlo simulations generate data that usually follow a specific probability distribution. To analyze the results, statistical approach is needed In this section, we review basic concepts of probability, which are required to understand sophisticate data analysis.

7.1. Mean and variance

7.1.1. Probability and probability distribution

A *sample space* Ω is a set of all possible outcomes for a trial. For a coin flip trial, all possible outcomes are {Head, Tail}. A *random variable* X is a function mapping from sample space to a real number (or some labels of events, such as H or T). The random variable is called *discrete* if it is defined on a discrete sample space. If the underlying sample space is continuous, the random variable is called *continuous*.

If we have a subset of Ω such that $\{\omega \mid X(\omega) \leq x\}$, a probability can be written as $P(X \leq x)$. Sometimes this probability is expressed using *cumulative distribution function* (abbreviated as c.d.f.) $F_X(x) = P(X \leq x)$. *Probability density function* (sometimes called probability distribution function, abbreviated as p.d.f.) p_X is frequently used to describe probability of random variable X. For a discrete random variable X, p.d.f. is given by

$$p_X(x) = P(X = x). \tag{7.1}$$

Sum of $p_X(x)$ for all possible events is 1. For continuous case, $p_X(x)$ is determined through following properties:

$$p_X(x) \geq 0 \quad \text{for all } x \in \mathbb{R} \tag{7.2}$$

$$\int_{-\infty}^{\infty} p_X(x)dx = 1 \tag{7.3}$$

$$P(a \leq X \leq b) = \int_a^b p_X(x)dx. \tag{7.4}$$

Here are some examples of continuous p.d.f.:

$$p_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad \text{normal dist.} \tag{7.5}$$

$$p_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases} \qquad \text{uniform dist.} \tag{7.6}$$

7.1.2. Mean and variance

Expectation value of a random variable X is a sum (discrete) or an integral (continuous) of the random variable multiplied by probability density function. It is usually denoted by $E[X]$ in statistics. Physicists sometimes use bra-ket notation $\langle X \rangle$.

$$E[X] = \sum_{x \in \Omega} x p_X(x) \qquad \text{discrete case} \tag{7.7}$$

$$E[X] = \int_{-\infty}^{\infty} x p_X(x) dx \qquad \text{continuous case.} \tag{7.8}$$

Expectation value has following linear properties:

$$E[X+Y] = E[X] + E[Y] \tag{7.9}$$

$$E[cX] = cE[X] \tag{7.10}$$

where X and Y are random variables and c is a constant. If X and Y are independent, expectation value of XY can be split:

$$E[XY] = E[X]E[Y]. \tag{7.11}$$

Here we used *joint probability density function*

$$p_{X,Y}(x,y) = P(X = x, Y = y) \tag{7.12}$$

$$= p_X(x) \cdot p_Y(y) \quad \text{(if X and Y are indep.).} \tag{7.13}$$

Variation of a random variable X, is defined by expectation value of $(X - E[X])^2$. Expanding the square and using the linearity of expectation value, the variance $V(X)$ (sometimes $Var(X)$ is used to denote variance of X) becomes:

$$V(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2. \tag{7.14}$$

Standard deviation of a random variable is the square root of its variance:

$$\sigma_X \equiv \sqrt{V(X)}. \tag{7.15}$$

Variation has some useful properties:

$$V(cX) = c^2 V(X) \tag{7.16}$$
$$V(X+c) = V(X). \tag{7.17}$$

For independent random variables X, Y,

$$V(X+Y) = V(X) + V(Y). \tag{7.18}$$

These properties can be shown by the definition of variance Eq. (7.14). *Standard deviation* of a random variable X, is defined by $\sigma(X) = \sqrt{V(X)}$.

Covariance of two random variables X and Y is defined by

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]. \tag{7.19}$$

If X and Y are independent, $E[XY] = E[X]E[Y]$, covariance of the two random variables is zero.

Let us consider a sequence of random variables $X_1, X_2, \cdots, X_N$ that *independent and identically distributed* (abbreviated as i.i.d.). That means that they are mutually independent and have the same probability distribution. A new random variable $\overline{X}$, which is a kind of average, is defined by

$$\overline{X} = \frac{1}{N}\sum_{k=1}^{N} X_k. \tag{7.20}$$

Then the expectation value and the variance are:

$$E[\overline{X}] = E[X_i] = \mu, \tag{7.21}$$
$$V(\overline{X}) = \frac{1}{N} V(X_i) = \frac{\sigma^2}{N}, \tag{7.22}$$

where i can be chosen arbitrarily, μ is the expectation value of X_i, and σ^2 is the variance of X_i.

7.1.3. Sample mean and sample variance

For given samples $\{x_1, x_2, \cdots, x_N\}$ from a random variable X, *sample mean* is

$$\overline{x} = \frac{1}{N}\sum_{k=1}^{N} x_k. \tag{7.23}$$

This form is the same as the mean of i.i.d. random variables. However, *sample variance* is somewhat different:

$$\sigma_x^2 = \frac{1}{N-1}\sum_{k=1}^{N}(x_k - \overline{x})^2 = \frac{1}{N-1}\left(\sum_{k=1}^{N} x_k^2 - N\overline{x}^2\right). \tag{7.24}$$

The difference $N-1$ comes from the fact that the $E[X]$ is not known so it is replaced by sample mean $\overline{x}$. This new denominator is called Bessel's correction. To be clear, let us show $E[\sigma_X^2] = \sigma^2$ where σ_X^2 is a variance of i.i.d. random variables defined by

$$\sigma_X^2 = \frac{1}{N-1}\left(\sum_{k=1}^{N} X_k^2 - N\overline{X}^2\right), \tag{7.25}$$

and σ^2 is variance of the random variable X_i.

$$(N-1)E[\sigma_X^2] = E[\sum_{k=1}^{N} X_k^2] - NE[\overline{X}^2] = NE[X_i^2] - NE[\overline{X}^2] \tag{7.26}$$

where i can be chosen arbitrarily, and the last equality holds for i.i.d. random variables X_i. Using the fact that $V(Y) = E[Y^2] - (E[Y])^2$ for a random variable Y, the following two equations can be written:

$$E[X_i^2] = V(X_i) + (E[X_i])^2 = \sigma^2 + \mu^2, \tag{7.27}$$

$$E[\overline{X}^2] = V(\overline{X}) + (E[\overline{X}])^2 = \frac{\sigma^2}{N} + \mu^2. \tag{7.28}$$

Inserting these two equations into right hand side of the Eq. (7.26), one finds

$$(N-1)E[\sigma_X^2] = (N-1)\sigma^2. \tag{7.29}$$

This is what we wanted to show.

Sample variance of mean can be derived in the same way:

$$\sigma_{\overline{x}}^2 = \frac{1}{N(N-1)}\sum_{k=1}^{N}(x_k - \overline{x})^2. \tag{7.30}$$

This is a combined result of the variance of mean (Eq. (7.22)) and Bessel's correction.

Also a *sample covariance* and a *sample covariance of mean* can be considered. Let x_i be the samples of a random variable X and y_i be the samples of

a random variable Y. The sample covariance C_{xy} is

$$C_{xy} = \frac{1}{N-1}\sum_{k=1}^{N}(x_k - \overline{x})(y_k - \overline{y}) \tag{7.31}$$

and the sample covariance of mean $C_{\bar{x}\bar{y}}$ is

$$C_{\bar{x}\bar{y}} = \frac{1}{N(N-1)}\sum_{k=1}^{N}(x_k - \overline{x})(y_k - \overline{y}). \tag{7.32}$$

7.1.4. Fundamental theorems of probability

In probability theory, there are two fundamental theorems: the *law of large numbers* and the *central limit theorem*.

According to the law of large numbers, the average of large number of trials is close to the expectation value. It has two forms: the *weak law of large numbers* and the *strong law of large numbers*. The weak one describes the probabilistic convergence of sample average toward the expected value. The strong one, which implies the weak law of large numbers, states that the sample average of a random variable converges to the expectation value with probability 1:

$$P\left(\lim_{N\to\infty}\frac{1}{N}(X_1 + X_2 + \cdots + X_N) = \mu\right) = 1, \tag{7.33}$$

where X_i are i.i.d. random variables with finite expectation value μ and finite variance. A simple proof with extra assumption $E[X_i^4] < \infty$ can be found in Refs. [117, 118].

Central limit theorem states that the average of sufficiently large number of samples of a random variable are distributed according to normal distribution:

$$\lim_{N\to\infty} P\left(\frac{a\sigma}{\sqrt{N}} < (\overline{X} - \mu) < \frac{b\sigma}{\sqrt{N}}\right) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}dx \tag{7.34}$$

where X_i are i.i.d. random variables with $E[X_i] = \mu$ and $V(X_i) = \sigma^2$. This means that if we define a new random variable

$$Z_N = \frac{X_1 + X_2 + \cdots + X_N - N\mu}{\sigma\sqrt{N}} = \frac{\overline{X} - \mu}{\sigma/\sqrt{N}}, \tag{7.35}$$

then Z_N tends to standard normal distribution:

$$\lim_{N\to\infty} Z_N \sim \mathcal{N}(0,1). \tag{7.36}$$

Here the notation $X \sim \mathcal{N}(\mu, \sigma^2)$ means that the random variable X tends to

normal distribution with expectation value μ and variance σ^2. The proof can be found in Ref. [117].

7.2. Special distributions

7.2.1. Normal distribution

The *normal distribution* is a continuous distribution whose probability distribution function is given by

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{7.37}$$

where μ is the mean value and σ^2 the variance. It is also called the Gaussian distribution. In the case of $\mu = 0$ and $\sigma^2 = 1$, it is called the *standard normal distribution* : $\mathcal{N}(0, 1)$. The cumulative distribution function of the normal distribution is given in terms of the error function(erf):

$$\begin{aligned} F(x;\, \mu, \sigma^2) &= \int_{-\infty}^{x} p(t)dt \\ &= \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right], \quad x \in \mathbb{R}. \end{aligned} \tag{7.38}$$

If a random variable X is distributed according to the normal distribution with mean μ and variance σ^2, it is usually denoted by

$$X \sim \mathcal{N}(\mu, \sigma^2). \tag{7.39}$$

The moment of the normal distribution is

$$\mathrm{E}\left[(X-\mu)^p\right] = \begin{cases} 0 & \text{if } p \text{ is odd,} \\ \sigma^p\,(p-1)!! & \text{if } p \text{ is even.} \end{cases} \tag{7.40}$$

A linear combination of normally distributed independent random variables X and Y, with means μ_X, μ_Y and variance σ_X^2, σ_Y^2 is also normally distributed:

$$X + Y \sim \mathcal{N}(\mu_X + \mu_Y,\, \sigma_X^2 + \sigma_Y^2)\,. \tag{7.41}$$

The proof is follows. If we define a new random variable $Z = X + Y$, then the probability distribution of Z can be written as

$$p_Z(z) = \int_{-\infty}^{\infty} p_Y(z-x) p_X(x) dx\,, \tag{7.42}$$

where we multiplied the probability of $X = x$ with $Y = z - x$ to obtain the probability of $Z = X + Y = z$. Using the explicit form of $p_X(x)$ and $p_Y(z-x)$ given in Eq. (7.37), we find

$$p_Z(z) = \frac{1}{\sqrt{2\pi(\sigma_X^2 + \sigma_Y^2)}} \exp\left[-\frac{(z - (\mu_X + \mu_Y))^2}{2(\sigma_X^2 + \sigma_Y^2)}\right]. \tag{7.43}$$

Since $p_Z(z)$ is a probability distribution function of a normal random variable with mean $\mu_X + \mu_Y$ and variance $\sigma_X^2 + \sigma_Y^2$, this proves Eq. (7.41).

7.2.2. χ^2-distribution and noncentral χ^2-distribution

If the random variables $Z_1, Z_2, \cdots, Z_k$ are independent, standard normal random variables ($Z_i \sim \mathcal{N}(0,1)$), then a new random variable Y defined by

$$Y = \sum_{i=1}^{k} Z_i^2 \tag{7.44}$$

tends to χ^2-distribution with k degrees of freedom. It is usually denoted by

$$Y \sim \chi_k^2. \tag{7.45}$$

The probability distribution function of the χ^2-distribution is

$$p(x;k) = \begin{cases} \frac{1}{2^{k/2}\Gamma(k/2)}\, x^{k/2-1} e^{-x/2} & (x \geq 0); \\ 0 & (\text{otherwise}). \end{cases} \tag{7.46}$$

where k is the degrees of freedom and $\Gamma(k/2)$ is the gamma function. A derivation for this is given in Appendix C. The cumulative distribution function can be easily calculated:

$$\begin{aligned} F(x;k) &= \int_{-\infty}^{x} p(t;k)dt \\ &= \frac{\gamma(k/2, x/2)}{\Gamma(k/2)} \\ &= P(k/2, x/2) = 1 - Q(k/2, x/2). \end{aligned} \tag{7.47}$$

Here the terms $Q(k/2, x/2)$ and $P(k/2, x/2)$ are regularized gamma functions and the term $\gamma(k/2, x/2)$ denotes the lower incomplete gamma function. For definitions of the gamma functions, see Appendix B.

Using Eq. (7.46), one can find the mean and variance of the χ^2-distribution

as follows

$$E[Y] = k \tag{7.48}$$
$$V[Y] = 2k. \tag{7.49}$$

If the random variables $X_1, X_2, \cdots, X_k$ are independent, normal random variables with mean μ_i and variance σ_i^2, then their regularized sum,

$$\sum_{i=1}^{k} \left(\frac{X_i - \mu_i}{\sigma_i} \right)^2 \sim \chi_k^2, \tag{7.50}$$

is distributed as χ^2 distribution with degrees of freedom k. If we define a new random variable Y' as following,

$$Y' = \sum_{i=1}^{k} \left(\frac{X_i}{\sigma_i} \right)^2, \tag{7.51}$$

then it is distributed according to the noncentral χ^2-distribution with degrees of freedom k and noncentrality parameter λ defined by

$$\lambda = \sum_{i=1}^{k} \left(\frac{\mu_i}{\sigma_i} \right)^2. \tag{7.52}$$

The probability distribution function is given by

$$p(x; k, \lambda) = \frac{1}{2} e^{-(x+\lambda)/2} \left(\frac{x}{\lambda} \right)^{k/4-1/2} I_{k/2-1}(\sqrt{\lambda x}) \tag{7.53}$$

where $I_\nu(z)$ is a modified Bessel function of the first kind [119]. This probability distribution $p(x; k, \lambda)$ becomes $p(x; k)$ of Eq. (7.46). The mean and variance of the noncentral χ^2-distribution are

$$E[Y'] = k + \lambda \tag{7.54}$$
$$V[Y'] = 2(k + 2\lambda). \tag{7.55}$$

8. Error Analysis

There are two types of errors - *systematic error* and *statistical error*. Let us consider pipes that a factory produces, and we want to know the expectation value of the pipe's length. We are measuring the pipes using a ruler whose minimum division is $1mm$. We sampled 100 pipes. Its mean length was $10cm$ and its standard deviation was $0.5cm$. If we say that the expectation value of the pipe's length is $10cm$, it is not a exact value and there are two sources of errors. One is the minimum division of the ruler and it generates a systematic error $0.5mm$. The other is finite number of samplings (100 pipes) and it generates a statistical error $\frac{0.5cm}{\sqrt{100}}$.

In many cases, the expectation value of a measurement is not what we ultimately want to know. Usually we use the measured values as an input to functions. Let us consider a function:

$$y = f(x). \tag{8.1}$$

If x is a measured value having errors, y also has some error. In this section, We review two methods that estimate the error of y originated from the statistical error of x.

8.1. Propagation of error

If the function form is known and if it is differentiable, *propagation of error* is a good method to estimate the error. If the mean value of random variables is $\overline{X}$ and the standard deviation of $\overline{X}$ is $\sigma_{\overline{X}}$, then the error propagates as follow:

$$f(\overline{X} \pm \sigma_{\overline{X}}) = f(\overline{X}) \pm \left.\frac{\partial f(X)}{\partial X}\right|_{\overline{X}} \cdot \sigma_{\overline{X}} + \mathcal{O}(\sigma_{\overline{X}}^2) \tag{8.2}$$

Here, $X = \overline{X} \pm \sigma_{\overline{X}}$ means that X fluctuates around between $\overline{X} - \sigma_{\overline{X}}$, $\overline{X} + \sigma_{\overline{X}}$ and the $\sigma_{\overline{X}}$ is

$$\sigma_X^2 = \left\langle (X - \langle X \rangle)^2 \right\rangle, \tag{8.3}$$

$$\sigma_{\overline{X}}^2 = \left\langle (\overline{X} - \langle \overline{X} \rangle)^2 \right\rangle = \frac{1}{N}\sigma_X^2. \tag{8.4}$$

Because σ_X is finite and its expection value is not depend on N, $\sigma_{\overline{X}}$ decreases as N goes large. Therefore the higher order term in Eq. (8.2) is $\mathcal{O}(\frac{1}{N})$ and can be ignored if N is large.

Eq. (8.2) shows that the expected value is $y = f(\overline{X})$ and its error is:

$$\sigma_f = f'(\overline{X}) \cdot \sigma_{\overline{X}} \tag{8.5}$$

This result can be obtained by the other way. Let us define the error of function f as

$$\sigma_f^2 = \left\langle \left[f(\overline{X}) - \langle f(\overline{X}) \rangle \right]^2 \right\rangle. \tag{8.6}$$

Since the difference $(\overline{X} - \langle X \rangle)$ is small,

$$\mathcal{O}\left(\overline{X} - \langle X \rangle\right) = \mathcal{O}\left(\sigma_{\overline{X}}\right) = \mathcal{O}\left(\frac{1}{\sqrt{N}}\right), \tag{8.7}$$

Eq. (8.6) can be written in other form. First, note that $f(\overline{X})$ can be expanded around $\langle X \rangle$ and the expectation value $\langle f(\overline{X}) \rangle$ is

$$\left\langle f(\overline{X}) \right\rangle = \left\langle f(\langle X \rangle) + f'\left(\langle X \rangle\right) \cdot \left(\overline{X} - \langle X \rangle\right) + \mathcal{O}\left(\left(\overline{X} - \langle X \rangle\right)^2\right) \right\rangle \simeq f(\langle X \rangle). \tag{8.8}$$

The difference between functions also can be expanded:

$$f(\overline{X}) - f(\langle X \rangle) = f'\left(\langle X \rangle\right) \cdot \left(\overline{X} - \langle X \rangle\right) + \mathcal{O}\left(\left(\overline{X} - \langle X \rangle\right)^2\right). \tag{8.9}$$

Using this expansion, Eq. (8.6) becomes

$$\begin{aligned} \sigma_f^2 = \left[f'\left(\langle X \rangle\right)\right]^2 \cdot \langle \left(\overline{X} - \langle X \rangle\right)^2 \rangle &= \left[f'\left(\langle X \rangle\right)\right]^2 \cdot \sigma_{\overline{X}}^2 \\ &\simeq \left[f'\left(\overline{X}\right)\right]^2 \cdot \sigma_{\overline{X}}^2, \end{aligned} \tag{8.10}$$

which is the same result of Eq. (8.5).

Let us consider a multi-variable function $f(X, Y)$. The error of this new function also can be obtained by the same way:

$$\begin{aligned} \sigma_f^2 &= \left\langle \left[f(\overline{X}, \overline{Y}) - \langle f(\overline{X}, \overline{Y}) \rangle \right]^2 \right\rangle \\ &\simeq \left\langle \left[f(\overline{X}, \overline{Y}) - f(\langle X \rangle, \langle Y \rangle) \right]^2 \right\rangle. \end{aligned} \tag{8.11}$$

Expanding $f(\overline{X}, \overline{Y})$ around $(\langle X \rangle, \langle Y \rangle)$ yields

$$f(\overline{X}, \overline{Y}) - f(\langle X \rangle, \langle Y \rangle) = \frac{\partial f}{\partial X} \cdot \left(\overline{X} - \langle X \rangle\right) + \frac{\partial f}{\partial Y} \cdot \left(\overline{Y} - \langle Y \rangle\right) + \cdots. \tag{8.12}$$

Using this expansion, Eq. (8.11) becomes

$$\begin{aligned}\sigma_f^2 &= \left(\frac{\partial f}{\partial X}\right)^2 \langle(\overline{X} - \langle X\rangle)^2\rangle + \left(\frac{\partial f}{\partial Y}\right)^2 \langle(\overline{Y} - \langle Y\rangle)^2\rangle \\ &\quad +2\frac{\partial f}{\partial X}\frac{\partial f}{\partial Y}\langle(\overline{X} - \langle X\rangle)(\overline{Y} - \langle Y\rangle)\rangle + \cdots \\ &\simeq \left(\frac{\partial f}{\partial X}\right)^2 \sigma_{\overline{X}}^2 + \left(\frac{\partial f}{\partial Y}\right)^2 \sigma_{\overline{Y}}^2 + 2\frac{\partial f}{\partial X}\cdot\frac{\partial f}{\partial Y}\cdot Cov(\overline{X},\overline{Y}) \qquad (8.13)\end{aligned}$$

where the covariance between $\overline{X}$ and $\overline{Y}$, $Cov(\overline{X},\overline{Y})$, is defined in Eq. (7.19).

Error propagation method is a good estimator of error. It is easy to understand and easy to use. However, this method is feasible only if the functional form is known and differentiable. There are another error estimation methods, which dose not require analytic functional forms: the resampling methods.

8.2. Resampling methods

Let us consider a set of random samples following specific distribution:

$$\{x_1, x_2, \cdots, x_N\}, \qquad (8.14)$$

and a function that takes a set of random samples and yields a number:

$$y = f(\{x_1, x_2, \cdots, x_N\}) \qquad (8.15)$$

For example, the set of x can be a data from an experiment or a simulation and the function can be $f(\overline{x})$ or χ^2-minimizer. If we want to know the error of y, we need to generate many sets of samples and make a set of y to find the standard deviation of y:

$$\begin{aligned} y_1 &= f(\{x_1^1, x_2^1, \cdots, x_N^1\}), \\ y_2 &= f(\{x_1^2, x_2^2, \cdots, x_N^2\}), \\ &\vdots \\ y_M &= f(\{x_1^M, x_2^M, \cdots, x_N^M\}), \qquad (8.16)\end{aligned}$$

$$\sigma_y^2 = \frac{1}{M-1}\sum_{k=1}^{M}(\overline{y} - y_k)^2. \qquad (8.17)$$

Generally, we do not have such many sets of samples, but only one set of samples. To estimate the error of y, therefore, we have to make sets of samples using one set of samples. This is called *resampling*. There are two common resampling methods - bootstrapping and jackknifing.

8.2.1. Bootstrap method

Bootstrapping is a resampling method introduced by Efron in 1979 [120]. Because we do not know the distribution function of a set of samples (Eq. (8.14)), let us just assume that the set of samples is the population of the distribution. Then sets of random samples can be sampled from the population. Randomly taking N samples from the population generates a bootstrap sample. The resampling process may take a sample, at random, repeatedly. For example, a bootstrap sample may include x_1 multiple times and do not include x_3. Let us make M sets of bootstrap samples that consists of N samples:

$$\{x_1^1, x_2^1, \cdots, x_N^1\}, \quad \{x_1^2, x_2^2, \cdots, x_N^2\}, \quad \cdots, \quad \{x_1^M, x_2^M, \cdots, x_N^M\} \tag{8.18}$$

Inputting each set into the function (Eq. (8.15)) generates a set of function values:

$$\{y_1^B, y_2^B, \cdots, y_M^B\}. \tag{8.19}$$

Here, the superscript B stands for bootstrap. This set of values provides us the standard deviation of y,

$$\sigma_y^2 = \frac{1}{M-1} \sum_{k=1}^{M} (\overline{y^B} - y_k^B)^2, \tag{8.20}$$

where $\overline{y^B}$ is an average of y_k^B over k.

8.2.2. Jackknife method

Jackknife method is a resampling method introduced by Quenouille(1956) [121] and developed by Tukey(1958) [122]. The most popular jackknife method samples by deleting one element of the set, Eq. (8.14). It yields N sets of samples that consists of $N-1$ samples:

$$\begin{aligned}
&\{x_2, x_3, x_4, x_5, \cdots, x_N\}, \\
&\{x_1, x_3, x_4, x_5, \cdots, x_N\}, \\
&\qquad \vdots \\
&\{x_1, x_2, x_3, \cdots, x_{N-3}, x_{N-2}, x_N\}, \\
&\{x_1, x_2, x_3, \cdots, x_{N-3}, x_{N-2}, x_{N-1}\}.
\end{aligned} \tag{8.21}$$

Inputting each set into the function (Eq. (8.15)) generates a set of function values:

$$\{y_1^J, y_2^J, \cdots, y_N^J\}. \tag{8.22}$$

Here, the superscript J stands for jackknife. This set of values provides us the estimated value of y and its standard deviation as follows,

$$y = \overline{y^J} = \frac{1}{N}\sum_{k=1}^{N} y_k^J, \tag{8.23}$$

$$\sigma_y^2 = \frac{N-1}{N}\sum_{k=1}^{N}(\overline{y^J} - y_k^J)^2. \tag{8.24}$$

Note that the overall factor of Eq. (8.24) is not $\frac{1}{N-1}$ but $\frac{N-1}{N}$. Since the sets of jackknife samples are similar each other, y_k^J are sharply distributed around $\overline{y^J}$. Therefore, the standard deviation of y should be bigger than the standard deviation of y_k^J. Later in this chapter, we show that this definition of standard deviation yields the same result of the error propagation if the function, Eq. (8.15), is a function of averaged values.

The difference between the function value of the original samples (Eq. (8.15)) and the jackknife estimate (Eq. (8.23)) depends on the function form. Appendix D shows that if the function is the sample variance of mean, the difference becomes $\mathcal{O}(\frac{1}{N})$. If the function is a function of averaged values, $f = f(\overline{x})$, the difference is $\mathcal{O}(\frac{1}{N^2})$. The mean value of i^{th} jackknife sample is

$$x_i^J = \frac{1}{N-1}\sum_{k\neq i} x_k = \overline{x} + \frac{1}{N-1}(\overline{x} - x_i). \tag{8.25}$$

Then y_i^J is

$$\begin{aligned} y_i^J &= f(x_i^J) \\ &= f(\overline{x}) + \frac{f'(\overline{x})}{N-1}(\overline{x} - x_i) + \frac{f''(\overline{x})}{2(N-1)^2}(\overline{x} - x_i)^2 + \mathcal{O}\left(\frac{1}{N^3}\right). \end{aligned} \tag{8.26}$$

Using this equality, $\frac{1}{N}\sum_{i=1}^{N} f(x_i^J) = \overline{y^j}$ can be shown:

$$\begin{aligned} \frac{1}{N}\sum_{i=1}^{N} f(x_i^J) &= \frac{1}{N}\sum_{i=1}^{N} f(\overline{x}) + \frac{f'(\overline{x})}{N(N-1)}\sum_{i=1}^{N}(\overline{x} - x_i) \\ &\quad + \frac{f''(\overline{x})}{2N(N-1)^2}\sum_{i=1}^{N}(\overline{x} - x_i)^2 + \mathcal{O}\left(\frac{1}{N^3}\right) \\ &= f(\overline{x}) + \frac{f''(\overline{x})}{2N(N-1)}\sigma_x^2 + \mathcal{O}\left(\frac{1}{N^3}\right) \\ &= f(\overline{x}) + \mathcal{O}\left(\frac{1}{N^2}\right). \end{aligned} \tag{8.27}$$

It is easy to show that the estimated standard deviation (Eq. (8.24)) is the same as the result of error propagation (Eq. (8.5)). Let us show that the estimated standard deviation of jackknife method is the same as that of error propagation method:

$$\begin{aligned}\sigma_y^2 &= \frac{N-1}{N}\sum_{i=1}^{N}\left(f(x_i^J) - \frac{1}{N}\sum_{i=1}^{N} f(x_i^J)\right)^2 \\ &= \frac{N-1}{N}\sum_{i=1}^{N}\left(\frac{f'(\overline{x})}{N-1}(\overline{x}-x_i) + \frac{f''(\overline{x})}{2(N-1)^2}\left\{(\overline{x}-x_i)^2 - \sigma_x^2\right\} + \mathcal{O}\left(\frac{1}{N^3}\right)\right)^2 \\ &= \frac{[f'(\overline{x})]^2}{N(N-1)}\sum_{i=1}^{N}(\overline{x}-x_i)^2 + \frac{f'(\overline{x})f''(\overline{x})}{N(N-1)^2}\sum_{i=1}^{N}(\overline{x}-x_i)^3 + \mathcal{O}\left(\frac{1}{N^3}\right) \\ &= \left[f'(\overline{x})\right]^2\sigma_{\overline{x}}^2 + \mathcal{O}\left(\frac{1}{N^2}\right). \end{aligned} \tag{8.28}$$

If the x_i are distributed symmetrically around $\overline{x}$, the sum $\sum_{i=1}^{N}(\overline{x}-x_i)^3$ is zero. In that case, estimated standard deviation of jackknife method is much closer to that of error propagation:

$$\sigma_y^2 = \left[f'(\overline{x})\right]^2\sigma_{\overline{x}}^2 + \mathcal{O}\left(\frac{1}{N^3}\right). \tag{8.29}$$

Here we reviewed only the deleted-one jackknife method, which makes a jackknife sample by removing one element. Generally, deleted-k jackknife method, which makes a jackknife sample by removing k elements, is possible.

8.3. Calculating error of error

Sometimes the object is a statistical error itself, such as standard deviation or covariance. In this case, the statistical error of the error is needed. The error of error can be calculated by the resampling methods or the error propagation.

Let us consider the error of sample covariance. After generating samples by the single elimination jackknife method, sample covariance of each sample can be calculated. The set of sample covariances of the jackknife samples yields the error of sample covariance through the jackknife error estimation, Eq. (8.24).

Error propagation also yields the error of error. Let $\{x_k(i)|k = 1, 2, \cdots, N\}$ be a set of samples of random variables $X(i)$. The covariance matrix C of sample mean is

$$C_{ij} = \frac{1}{N(N-1)}\sum_{k=1}^{N}\Big(x_k(i) - \overline{x}(i)\Big)\Big(x_k(j) - \overline{x}(j)\Big), \tag{8.30}$$

where $\overline{x}(i)$ is an average of $x_k(i)$ over k. When the terms are rearranged, it can be regarded as a mean value of some variables $c_{ij}^{(k)}$ over k:

$$C_{ij} = \frac{1}{N}\sum_{k=1}^{N}\left[\frac{\Big(x_k(i)-\overline{x}(i)\Big)\Big(x_k(j)-\overline{x}(j)\Big)}{N-1}\right] \tag{8.31}$$

$$= \frac{1}{N}\sum_{k=1}^{N} c_{ij}^{(k)}, \tag{8.32}$$

where $c_{ij}^{(k)}$ is defined by

$$c_{ij}^{(k)} = \frac{\Big(x_k(i)-\overline{x}(i)\Big)\Big(x_k(j)-\overline{x}(j)\Big)}{N-1}. \tag{8.33}$$

The statistical error of C_{ij}, $\sigma(C_{ij})$, can be calculated by the sample variance of $c_{ij}^{(k)}$:

$$\sigma^2(C_{ij}) = \frac{1}{N(N-1)}\sum_{k=1}^{N}\Big(c_{ij}^{(k)}-\overline{c}_{ij}\Big)^2. \tag{8.34}$$

Here, $\overline{c}_{ij}$ is an average of $c_{ij}^{(k)}$ over k, which is C_{ij} (See Eq. (8.32)). One can easily rearrange the variance as

$$\sigma^2(C_{ij}) = \frac{1}{N(N-1)}\sum_{k=1}^{N}\Big(c_{ij}^{(k)}-C_{ij}\Big)^2 \tag{8.35}$$

$$= \frac{1}{N-1}\left[\frac{1}{N}\sum_{k=1}^{N}\Big(c_{ij}^{(k)}\Big)^2-\Big(C_{ij}\Big)^2\right]. \tag{8.36}$$

In this example, one may not see why this is an error propagation. Let us consider a case calculating the error of standard deviation $\sigma_{\overline{x}}$,

$$\sigma_{\overline{x}} = \sqrt{\frac{1}{N(N-1)}\sum_{k=1}^{N}\Big(x_k-\overline{x}\Big)^2}. \tag{8.37}$$

The term inside the square-root is also can be regarded as an averaged value of s_i,

$$\frac{1}{N(N-1)}\sum_{k=1}^{N}\Big(x_k-\overline{x}\Big)^2 = \frac{1}{N}\sum_{k=1}^{N} s_i \equiv \overline{s}, \tag{8.38}$$

where s_i is defined by

$$s_i = \frac{(x_k - \overline{x})^2}{N-1}. \tag{8.39}$$

Then $\sigma_{\overline{x}}$ becomes a function of an averaged value,

$$\sigma_{\overline{x}} = f(\overline{s}) = \sqrt{\overline{s}}, \tag{8.40}$$

whose statistical error can be calculated by the error propagation Eq. (8.5). In the case of the sample covariance of mean, the function was identity, $f(x) = x$.

8.4. Dealing with Jackknife samples

Sometimes, one has to do data analysis only with the jackknife data samples, which are the averages of single elimination jackknife samples of original data samples. The procedure making jackknife data samples does not lose any information, which indicates that one can recover the original samples. In some cases, one has only the measured results[1] from the jackknife samples. If the *measurement* that processes the jackknife samples is differentiable, one can approximately obtain the measured results of the bootstrap samples. In this section, we present methods which covert the measured results of jackknife samples into those of the original samples or the bootstrap samples.

8.4.1. From jackknife samples to original samples

Here, we show how to recover the original data samples from the jackknife data samples. Let $\{x_k(i)|k = 1, 2, \cdots, N\}$ be the samples of random variables $X(i)$ with $i = 1, 2, \cdots, D$. D is the number of random variables, and it can be consider as a dimension of a random sample, which will be explained below. The set of random samples can be written as

$$\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N\} \tag{8.41}$$

where N is the number of samples, and $\mathbf{x}_i$ is a data sample vector of i^{th} data,

$$\mathbf{x}_i = (x_i(1), x_i(2), \cdots, x_i(D)). \tag{8.42}$$

[1] Here, *measured results* denotes the processed results from the jackknife data. Generally, the *processing* includes the least χ^2 fitting, and in that case, the measured results are the fitting results.

From the original data, Eq. (8.41), N jackknife sets can be generated:

$$\begin{aligned}
&\{\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \cdots, \mathbf{x}_N\}, \\
&\{\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_4, \cdots, \mathbf{x}_N\}, \\
&\qquad \vdots \\
&\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{N-2}, \mathbf{x}_{N-1}\}.
\end{aligned} \tag{8.43}$$

For each jackknife set, there is a jackknife mean vector $\overline{\mathbf{x}}_i^J$,

$$\overline{\mathbf{x}}_i^J \equiv \frac{1}{N-1}\sum_{k\neq i}\mathbf{x}_k \tag{8.44}$$

$$= \frac{N}{N-1}\overline{\mathbf{x}} - \frac{1}{N-1}\mathbf{x}_i, \tag{8.45}$$

where the mean vector $\overline{\mathbf{x}}$ is

$$\overline{\mathbf{x}} = \frac{1}{N}\sum_{k=1}^{N}\mathbf{x}_k = \frac{1}{N}\sum_{k=1}^{N}\overline{\mathbf{x}}_k^J. \tag{8.46}$$

If one has only a set of jackknife mean vectors

$$\{\overline{\mathbf{x}}_1^J, \overline{\mathbf{x}}_2^J, \cdots, \overline{\mathbf{x}}_N^J\}, \tag{8.47}$$

the original data $\mathbf{x}_i$ can be obtained using the Eq. (8.45),

$$\mathbf{x}_i = N\overline{\mathbf{x}} - (N-1)\overline{\mathbf{x}}_i^J. \tag{8.48}$$

For example, let us consider a case calculating the statistical error of covariance matrix of sample mean $\overline{\mathbf{x}}$ using jackknife method. The covariance matrix of sample mean of the m^{th} jackknife set is

$$C_{ij}^{J,m} = \frac{1}{N-2}\left[\frac{1}{N-1}\sum_{k\neq m}\Big(x_k(i) - \overline{x}_m^J(i)\Big)\Big(x_k(j) - \overline{x}_m^J(j)\Big)\right], \tag{8.49}$$

where $\overline{\mathbf{x}}_k^J$ is defined in Eq. (8.44). This covariance matrix can be rewritten in terms of jackknife mean vectors using Eq. (8.48):

$$\begin{aligned}
C_{ij}^{J,m} = \frac{1}{N-2}\Bigg[&\frac{1}{N-1}\sum_{k\neq m}\Big\{(N-1)\Big[\overline{x}(i) - \overline{x}_k^J(i)\Big] + \Big[\overline{x}(i) - \overline{x}_m^J(i)\Big]\Big\} \\
&\times\Big\{(N-1)\Big[\overline{x}(j) - \overline{x}_k^J(j)\Big] + \Big[\overline{x}(j) - \overline{x}_m^J(j)\Big]\Big\}\Bigg]
\end{aligned} \tag{8.50}$$

Combining these results with Eq. (8.23) and Eq. (8.24) yields the jackknife estimate of the covariance matrix C_{ij} and its error $\sigma(C_{ij})$ by

$$C_{ij} = \frac{1}{N}\sum_{m=1}^{N} C_{ij}^{J,m} \tag{8.51}$$

$$\sigma(C_{ij}) = \sqrt{\frac{N-1}{N}\left[\sum_{m=1}^{N}\left(C_{ij}^{J,m} - C_{ij}\right)^2\right]}. \tag{8.52}$$

8.4.2. From jackknife results to bootstrap results

In some cases, one has only the measurements from the jackknife samples. Here, *measurements* denote the processed results from the jackknife data. Generally, the *processing* includes the least χ^2 fitting, and in that case, the measured results are the fitting results. If the *measurement* that processes the jackknife samples is differentiable, one can approximately obtain the measured results of the bootstrap samples. Here, we show the approximation method that produces the bootstrap results.

Let us consider a function of a set of samples:

$$u = f\Big(\{x_1, x_2, \cdots, x_N\}\Big). \tag{8.53}$$

The function $f(\{x\})$ is a measurement that processes a given set of data samples. It could include the least χ^2 fitting procedure. Let us assume that $f(\{x\})$ is a function an average,

$$u = f\Big(\{x_1, x_2, \cdots, x_N\}\Big) = f(\overline{x}), \tag{8.54}$$

where $\overline{x} = \frac{1}{N}\sum_i x_i$. This assumption is not necessary but it makes the calculation simple. The following calculation can be generalized for a function an average data points vector and a covariance matrix,

$$f\Big(\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N\}\Big) = f\left(\overline{\mathbf{x}}, C\right), \tag{8.55}$$

where $\overline{\mathbf{x}}$ is an average data points vector and C is a covariance matrix. The final result in Eq. (8.61) still holds. Hence, we can apply the result to the least χ^2 fitting. Note that if the function $f(\{x\})$ is a linear fitting process it is guaranteed the differentiability of the function, $f(\{x\})$.

The Jackknife average data, y_i, and Bootstrap average data, z_i, is defined by

$$y_i = \frac{1}{N-1}\sum_{k\neq i} x_k = \overline{x} + \frac{1}{N-1}(\overline{x} - x_i) \tag{8.56}$$

$$z_i = \frac{1}{N}\sum_{k\in B_i} x_k = \overline{x} + \left(\frac{1}{N}\sum_{k\in B_i} x_k - \overline{x}\right), \tag{8.57}$$

where B_i is a set of N random integers in $\{1, 2, \cdots, N\}$, which defines the Bootstrap samples. If we assume that the set of original data samples, $\{x_i\}$ is the population of the base random variable, the *Central Limit Theorem* implies

$$\left(\frac{1}{N}\sum_{k\in B_i} x_k - \overline{x}\right) \sim \mathcal{O}\left(\frac{1}{\sqrt{N}}\right). \tag{8.58}$$

The Jackknife result and Bootstrap result is defined by

$$v_i = f(y_i) = f(\overline{x}) + \frac{f'(\overline{x})}{N-1}(\overline{x} - x_i) + \frac{f''(\overline{x})}{2(N-1)^2}(\overline{x} - x_i)^2 + \mathcal{O}(\frac{1}{N^3}) \tag{8.59}$$

$$w_i = f(z_i) = f(\overline{x}) + f'(\overline{x})\left(\frac{1}{N}\sum_{k\in B_i} x_k - \overline{x}\right) + \mathcal{O}(\frac{1}{N}). \tag{8.60}$$

Here, v_i is the Jackknife result, which is what we have, and w_i is the Bootstrap result, which is what we want to know, and note that $f(\overline{x}) = \overline{v} + \mathcal{O}(\frac{1}{N^2})$. By eliminating x_k and $\overline{x}$ in Eq. (8.60) using Eq. (8.59), we find the following relation,

$$\begin{aligned} w_i &= \overline{v} + \frac{N-1}{N}\sum_{k\in B_i}(\overline{v} - v_k) + \mathcal{O}(\frac{1}{N}) \\ &= \frac{1}{N}\sum_{k\in B_i}\Big[\overline{v} + (N-1)(\overline{v} - v_k)\Big] + \mathcal{O}(\frac{1}{N}). \end{aligned} \tag{8.61}$$

Although the best way is to bootstrap from the raw data, one could approximately obtain the results of the bootstrap samples.

9. Least χ^2 Fitting

Experiments or simulations generate data, and usually, there are theories that explain the data. The theory can have free parameters, such as masses and normalization factors, which have to be determined by the data. The procedure finding appropriate parameter values from the data is called the *fitting*.

There are many kinds of fitting methods. The choice of a suitable fitting method depends on the characteristics of the data. Here we review the *least* χ^2 *fitting*, which finds fitting parameters that minimize χ^2.

9.1. Theory of least χ^2 fitting

Let us consider the probability distribution of our theory under the condition of the data y is observed:

$$P(\text{theory}|\text{data}) \equiv P(\lambda|\overline{y}). \tag{9.1}$$

Since the fitting parameters λ can represent the fitting function, we use λ as the theory. We also use the notation $\overline{y}$ to represent the data. Here we used *conditional probability* $P(A|B)$, which means the probability of some event A under the condition of some event B is given. $\overline{y}$ means the mean value of data. The best fitting parameters can be defined by the values that maximizes the probability distribution, Eq. (9.1).

Bayes' theorem states that

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \tag{9.2}$$

provided $P(B)$ is not zero. Using this theorem the probability of theory, Eq. (9.1), can be written as

$$P(\lambda|\overline{y}) = \frac{P(\overline{y}|\lambda)P(\lambda)}{P(\overline{y})}. \tag{9.3}$$

Since the data does not depend on the theory, let us assume that $P(\overline{y})$ is a constant. If we assume that every values of parameters are equivalent (no values are preferred) before we see the data, $P(\lambda)$ is also a constant. Under these assumptions, $P(\lambda|\overline{y})$ is proportional to the $P(\overline{y}|\lambda)$,

$$P(\lambda|\overline{y}) \propto P(\overline{y}|\lambda). \tag{9.4}$$

As a results, the probability of the fitting parameters can be maximized by maximizing $P(\overline{y}|\lambda)$, which is the probability that the data will be observed when the theory is true. Let us define the $P(\overline{y}|\lambda)$ as the *likelihood function*,

$$L = P(\overline{y}|\lambda). \tag{9.5}$$

9.1.1. Uncorrelated χ^2

Let us consider data $\overline{y}(t)$ with $t = 1, 2, \cdots, T$. For fixed t, $\overline{y}(t)$ is an average of many independent data $y_i(t)$ over i. In this section, we assume that $y_i(t)$ is independent in t for fixed i. In other words, we assume that $\overline{y}(t)$ are uncorrelated. Next section deals with the correlated case.

If we assume that a theory or a fitting function $f(t)$ is true, it gives expectation value of $\overline{y}(t)$ for each t. Under the assumption, in other words, the average $\overline{y}(t)$ should converges to the function value $f(t)$ as we do the experiment or simulation many times. Furthermore, the central limit theorem tells us that the distribution of $\overline{y}(t)$ is a normal distribution. Since the expectation value and the shape of distribution is given, the variance is the only missing variable. Let us approximate the variance by the sample variance of $\overline{y}(t)$. It completes the distribution of $\overline{y}(t)$:

$$\overline{y}(t) \sim \mathcal{N}(f(t), \sigma^2_{\overline{y}(t)}). \tag{9.6}$$

Because the distribution of $\overline{y}(t)$ is specified, the likelihood function of whole fitting can be derived. Experiments or Monte Carlo simulations give us the mean value $\overline{y}(t)$ and its variance $\sigma^2_{\overline{y}(t)}$ for each $t = 1, 2, \cdots, T$. If we assume that the theory is true, the p.d.f of $\overline{y}(t)$ is

$$p\big[\overline{y}(t)|f(t)\big] = \frac{1}{\sqrt{2\pi}\sigma_{\overline{y}(t)}} \exp\left[-\frac{\{\overline{y}(t) - f(t)\}^2}{2\sigma^2_{\overline{y}(t)}}\right]. \tag{9.7}$$

For whole fitting, the likelihood function can be defined by the joint probability distribution function of $\{\overline{y}(1), \overline{y}(2), \cdots, \overline{y}(T)\}$:

$$\begin{aligned} L &= p\big[\overline{y}(1), \overline{y}(2), \cdots, \overline{y}(T)|f(t)\big] \\ &\propto \exp\left[-\frac{1}{2}\sum_{t=1}^{T} \frac{\{\overline{y}(t) - f(t)\}^2}{\sigma^2_{\overline{y}(t)}}\right]. \end{aligned} \tag{9.8}$$

To maximize likelihood function L, we should determine fitting parameters that minimize χ^2 which is defined as

$$\chi^2 \equiv \sum_{t=1}^{T} \frac{[\overline{y}(t) - f(t)]^2}{\sigma^2_{\overline{y}(t)}}. \tag{9.9}$$

Least χ^2 fitting procedure is finding the parameter values, of the theory or fitting function $f(t)$, that minimize χ^2.

It is called χ^2 because it is distributed according to χ^2-distribution. If the random variables $Z_1, Z_2, \cdots, Z_k$ are independent, standard normal random variables ($Z_i \sim \mathcal{N}(0,1)$), then a new random variable Q defined by

$$Q = \sum_{i=1}^{k} Z_i^2 \tag{9.10}$$

tends to χ^2-distribution with k degrees of freedom (denoted by χ_k^2). If we assume that $f(t)$ is the expectation value of $\overline{y}(t)$, the term $\frac{[\overline{y}(t)-f(t)]^2}{\sigma^2_{\overline{y}(t)}}$ in Eq. (9.9) is square of standard normal random variable because $\overline{y}(t) \sim N(f(t), \sigma_{\overline{y}(t)})$. Although χ^2 is sum of T terms, if $f(t)$ has m free fitting parameters, m terms among $[\overline{y}(t) - f(t)]^2$ can be zero. Therefore, χ^2 tends to χ^2-distribution with ν degrees of freedom, where $\nu = T - m$. The expectation value of χ^2-distribution with ν degrees of freedom is ν and its variance is 2ν. Hence we expect that χ^2/ν to be

$$\frac{\chi^2}{\nu} = 1 \pm \sqrt{\frac{2}{\nu}}, \tag{9.11}$$

which means that the expectation value of χ^2/ν is 1, and the standard deviation is $\sqrt{2/\nu}$. Here we use the mean and variance of the χ^2-distribution given in Sec. 7.2.2.

9.1.2. Correlated χ^2

Let us consider data $\overline{y}(t)$ with $t = 1, 2, \cdots, T$. We have many data $y_i(t)$ for each t and $\overline{y}(t)$ is their average. Here we assume that the data is not independent in t. In order to deal with dependent case, we need multivariate random variable.

A multivariate random variable is a vector whose components are random variables :

$$\mathbf{X} = (X_1, X_2, \cdots, X_n). \tag{9.12}$$

Also there is a *multidimensional central limit theorem*, which is an extended version of central limit theorem [123]. It states that the average of sufficiently large number of samples of a multivariate random variable are distributed according to multivariate normal distribution. This means that if we define a new multivariate random variable

$$\mathbf{Z}_N = \frac{\mathbf{X}_1 + \mathbf{X}_2 + \cdots + \mathbf{X}_N - N\boldsymbol{\mu}}{\sqrt{N}} = \sqrt{N}(\overline{\mathbf{X}} - \boldsymbol{\mu}), \tag{9.13}$$

then $\mathbf{Z}_N$ is distributed according to the multivariate normal distribution with

zero mean when N is large enough:

$$\lim_{N\to\infty} \mathbf{Z}_N \sim \mathcal{N}(\mathbf{0}, \Sigma), \tag{9.14}$$

where $\mathbf{X}_i$ are independent and identically distributed (i.i.d.) multivariate random variables, $\boldsymbol{\mu}$ is a mean vector whose components are expectation value of $\mathbf{X}_i$, and Σ is covariance matrix of $\mathbf{X}_i$,

$$\Sigma = E\big[(\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})^T\big]\,. \tag{9.15}$$

Here, a multivariate random variable $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ is said to be distributed according to *multivariate normal distribution* if every linear combination of its components

$$Y = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n \tag{9.16}$$

is distributed according to normal distribution. The probability density function of $\mathbf{X}$ is

$$p_{\mathbf{X}}(\mathbf{t}) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{t}-\boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{t}-\boldsymbol{\mu})\right] \tag{9.17}$$

where $\mathbf{X}$ denotes the random variable distributed according to the distribution $p_{\mathbf{X}}(\mathbf{t})$, and $|\Sigma|$ is the determinant of Σ. This is an generalization of the probability distribution in one-dimension Eq. (7.37) to a multi-dimension.

Let us consider the data samples of a multivariate random variables:

$$\mathbf{y}_i = (y_i(1), y_i(2), \cdots, y_i(T))\,. \tag{9.18}$$

Multidimensional central limit theorem says that an average vector

$$\overline{\mathbf{y}} = (\overline{y}(1), \overline{y}(2), \cdots, \overline{y}(T)) \tag{9.19}$$

is distributed according to multivariate normal distribution. If we assume that the theory (fitting function) is true, it gives expectation value of the random vector, $E[\overline{\mathbf{y}}] \equiv \boldsymbol{f} = (f(1), f(2), \cdots, f(T))$. The covariance of the distribution can be approximated by the sample covariance of $\overline{\mathbf{y}}$.

The distribution that $\overline{\mathbf{y}}$ should follow is specified - multidimensional normal distribution with mean vector

$$\mathbf{f} = (f(1), f(2), \cdots, f(T)) \tag{9.20}$$

and with covariance matrix C which is estimated by the sample covariance matrix of $\overline{\mathbf{y}}$,

$$C_{tt'} = \frac{1}{N(N-1)} \sum_{i=1}^{N} \Big(y_i(t) - \overline{y}(t)\Big)\Big(y_i(t') - \overline{y}(t')\Big). \tag{9.21}$$

Here, $\overline{y}(t)$ is an average of $y_i(t)$ over i,

$$\overline{y}(t) = \frac{1}{N}\sum_{i=1}^{N} y_i(t). \tag{9.22}$$

As a results, we can define the likelihood function in the correlated case using the probability distribution of $\overline{\mathbf{y}}(t)$, Eq. (9.17),

$$L = P(\overline{\mathbf{y}}|\boldsymbol{\lambda}) = \mathcal{N}\exp\left[-\frac{1}{2}(\overline{\mathbf{y}}-\mathbf{f})^T C^{-1}(\overline{\mathbf{y}}-\mathbf{f})\right]. \tag{9.23}$$

To maximize the likelihood, one should minimize the exponent,

$$\chi^2 = (\overline{\mathbf{y}}-\mathbf{f})^T C^{-1}(\overline{\mathbf{y}}-\mathbf{f}) \tag{9.24}$$

$$= \sum_{t=1}^{T}\sum_{t'=1}^{T}\Big(\overline{y}(t)-f(t)\Big)(C^{-1})_{tt'}\Big(\overline{y}(t')-f(t')\Big). \tag{9.25}$$

Least χ^2 fitting procedure is finding the parameter values of the fitting function $f(t)$, that minimize χ^2.

The correlated χ^2 is distributed according to χ^2-distribution. To see that the correlated χ^2, Eq. (9.24), tends to χ^2-distribution, let M be a nonsingular matrix that satisfies

$$MCM^T = I. \tag{9.26}$$

Because we are assuming that $\overline{\mathbf{y}}$ is distributed according to multidimensional normal distribution with mean vector $\mathbf{f}$ and covariance matrix C,

$$(\overline{\mathbf{y}}-\mathbf{f}) \sim \mathcal{N}(\mathbf{0}, C), \tag{9.27}$$

if we define a new random variable $\mathbf{Y} = M(\overline{\mathbf{y}}-\mathbf{f})$, then it is distributed according to multidimensional normal distribution with mean

$$E[\mathbf{Y}] = ME[(\overline{\mathbf{y}}-\mathbf{f})] = \mathbf{0} \tag{9.28}$$

and with covariance matrix

$$E[\mathbf{Y}\mathbf{Y}^T] = E[M(\overline{\mathbf{y}}-\mathbf{f})(\overline{\mathbf{y}}-\mathbf{f})^T M] = MCM^T = I. \tag{9.29}$$

Since the covariance matrix is identity, the components of $\mathbf{Y}$, Y_i, are independent and standard normal random variables ($Y_i \sim \mathcal{N}(0,1)$). The correlated χ^2,

Eq. (9.24), can be rewritten in terms of $\mathbf{Y}$:

$$\begin{aligned}(\overline{\mathbf{y}}-\mathbf{f})^T C^{-1}(\overline{\mathbf{y}}-\mathbf{f}) &= \mathbf{Y}^T(M^T)^{-1}C^{-1}M^{-1}\mathbf{Y} = \mathbf{Y}^T(MCM^T)^{-1}\mathbf{Y} \\ &= \mathbf{Y}^T\mathbf{Y} \\ &= \sum_{i=1}^{T} Y_i^2 \end{aligned} \tag{9.30}$$

Therefore the correlated χ^2 is distributed as $\sum_{i=1}^{T} Y_i^2$, which is a definition of χ^2-distribution.

Although the correlated χ^2 is sum of T terms, if $f(t)$ has m free fitting parameters, m terms among $[\overline{y}(t)-f(t)]$ can be zero. Therefore, χ^2 tends to χ^2-distribution with ν degrees of freedom, where $\nu = T-m$. The expectation value of χ^2-distribution with ν degrees of freedom is ν and its variance is 2ν. Hence we expect that correlated χ^2/ν to be

$$\frac{\chi^2}{\nu} = 1 \pm \sqrt{\frac{2}{\nu}}. \tag{9.31}$$

The derivation of the theorems or the distributions in this section can be found in [123].

Here, there are two assumptions : (1)the fitting function describes the data good enough and (2)the number of data samples is large enough to estimate the covariance matrix precisely. In other words, we assumed that the expectations value of average data points, $\bar{y}(t) = f(t)$ and the sample covariance matrix defined in Eq. (9.21) is the true covariance matrix of data. In some cases, practically, these two assumptions are not satisfied.

Let us consider a inexact fitting function:

$$E[\bar{\mathbf{y}}] = \boldsymbol{\mu} \neq \mathbf{f}\,.$$

In that case, $\mathbf{Y} = \sqrt{N}[\bar{\mathbf{y}}-\mathbf{f}]$ is distributed according to $\mathcal{N}(\boldsymbol{\rho},\Gamma)$ where Γ is the true covariance matrix of the sample ($\frac{1}{N}\Gamma \simeq C$) and $\boldsymbol{\rho} = \sqrt{N}[\mu-\mathbf{f}]$. Here, we use the same notation as in Ref. [123]. In this case, $[\chi^2/(N-1)][(N-\nu)/\nu]$ is distributed as a non-central F distribution of $F_{\nu,N-\nu}$, which is defined in Ref. [123], and its non-centrality parameter is

$$\kappa \equiv \boldsymbol{\rho}^T\Gamma^{-1}\boldsymbol{\rho} = (\boldsymbol{\mu}-\mathbf{f})(N\ \Gamma^{-1})(\boldsymbol{\mu}-\mathbf{f})\,.$$

In Ref. [123], it is proved that the limiting distribution of χ^2 as $N\to\infty$ is the χ^2-distribution if $\boldsymbol{\mu} = \mathbf{f}$.

In the case of small number of data samples, the finite sample size effect should be considered. Considering finite sample size, the multivariate statistical

theory predicts the following [119]:

$$\begin{aligned} E[\chi^2] &= (d+\kappa)\left[1+\frac{d+1}{N}+\mathcal{O}(\frac{1}{N^2})\right] \\ V[\chi^2] &= 2(d+2\kappa)\left[1+\frac{1}{N}\left(2d+4+\frac{(d+\kappa)^2}{d+2\kappa}\right)+\mathcal{O}(\frac{1}{N^2})\right] \end{aligned} \tag{9.32}$$

where κ is non-centrality parameter.

9.1.3. Quality of the fit

Provided that the theory or fitting function describes the data well enough, the χ^2/ν is expected to have a value around 1, where ν is the degrees of freedom. If the minimized χ^2/ν is too large or too small, one may suspect that something is going wrong. Hence the χ^2/ν can be used as a value that indicates the quality of the fit.

When the χ^2/ν is too large, the first thing one should suspect is the theory or fitting function. A large χ^2/ν means that the mean of data is placed far away from the value of the fitting function. In other words, the theory does not describe the data well. Even though the expected value of χ^2/ν is 1, the probability that the χ^2/ν has a large value is not zero. Therefore there is a chance that the theory is right and the mean of data is placed at a improbable point.

Although the fitting process finds the fitting parameters that minimizes χ^2, too small value of χ^2/ν is not welcomed. When the variance of the data is over estimated, a small value of χ^2/ν would be detected. The fraudulent data or theory can be under consideration. Also there is a chance that a improbable thing happens.

Since the probability distribution of χ^2 is known, one can calculate the probability that the χ^2 is larger than the minimized value, which is the result of the fitting. This probability is expected to be 0.5 and if it is larger than or smaller than a certain value (such as 0.95 or 0.05), one may reject the fit result.

Let us consider a uncorrelated case. The probability density function of χ^2 is given by

$$p_\nu(x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2-1} e^{-x/2}, \tag{9.33}$$

where ν is the degrees of freedom and $\Gamma(\nu/2)$ is the gamma function. The cu-

mulative distribution function can be easily calculated:

$$\begin{aligned} F_\nu(x) &= \int_{-\infty}^{x} p_\nu(t)dt \\ &= \frac{\gamma(\nu/2, x/2)}{\Gamma(\nu/2)} \\ &= P(\nu/2, x/2) = 1 - Q(\nu/2, x/2). \end{aligned} \tag{9.34}$$

Here the terms $Q(\nu/2, x/2)$ and $P(\nu/2, x/2)$ are regularized gamma functions and the term $\gamma(\nu/2, x/2)$ denotes the lower incomplete gamma function. The value of $Q(\nu/2, x/2)$ is the probability that the χ^2 is larger than the minimized value. For example, when the degrees of freedom is 4, Q is smaller than 0.05 if the χ^2/ν is larger than 2.37. That means the probability that χ^2/ν gets a larger value than 2.37 is 0.05.

9.1.4. Uncertainty of fitting parameters

Since the data have statistical error, the fitting parameters also have statistical error. The error can be calculated by the direct integration of the expectation value, Eq. (9.38). Because the probability distribution of the fitting parameters is usually peaked near its maximum, VEGAS [124] is an effective method to integrate. One of the most robust ways to estimate the error of fitting parameters is resampling method (Section 8.2). However, there is an another way to estimate the error [125], which is an approximation of the direct integration method.

General χ^2 is

$$\chi^2 = \sum_{i,j} \Big(\overline{y}(i) - f(i;\boldsymbol{\lambda})\Big)(C^{-1})_{ij}\Big(\overline{y}(j) - f(j;\boldsymbol{\lambda})\Big), \tag{9.35}$$

where $\overline{y}(i)$ is the mean value of data at $t = i$ and $\boldsymbol{\lambda}$ is a vector of fitting parameters. Here C is a covariance matrix of $\overline{y}$, which can be estimated from a set of samples by

$$C_{ij} = \frac{1}{N(N-1)} \sum_{k=1}^{N} \Big(y_k(i) - \overline{y}(i)\Big)\Big(y_k(j) - \overline{y}(j)\Big). \tag{9.36}$$

Since the fitted parameters fluctuate as the data $\overline{y}$ fluctuate, the covariance of the fitted parameters $\overline{\boldsymbol{\lambda}}$ can be defined by

$$Cov(\overline{\lambda}_a, \overline{\lambda}_b) = \big\langle (\overline{\lambda}_a - \lambda_a^*)(\overline{\lambda}_b - \lambda_b^*) \big\rangle_{P(\boldsymbol{\lambda}|\overline{\mathbf{y}})} \tag{9.37}$$

$$= \frac{1}{\int [d\overline{\mathbf{y}}] P(\overline{\mathbf{y}}|\boldsymbol{\lambda})} \int [d\overline{\mathbf{y}}] P(\overline{\mathbf{y}}|\boldsymbol{\lambda})(\overline{\lambda}_a - \lambda_a^*)(\overline{\lambda}_b - \lambda_b^*), \tag{9.38}$$

where λ_a^* denotes the true mean of the parameter λ_a and $P(\boldsymbol{\lambda}|\overline{\mathbf{y}})$ is the probability distribution of $\overline{y}$ under the condition of a theory or fitting parameters are given. In correlated χ^2 fitting, $P(\boldsymbol{\lambda}|\overline{\mathbf{y}})$ is defined in Eq. (9.23).

For simplicity, we use some new notations:

$$\overline{y}_i \equiv \overline{y}(i), \quad f(i;\boldsymbol{\lambda}) \equiv f_i(\boldsymbol{\lambda}). \tag{9.39}$$

Also we will use summation convention over repeated indices.

Let $\overline{\boldsymbol{\lambda}}$ be the vector of fitted parameters that minimizes the χ^2. Then the gradient of χ^2 at $\overline{\boldsymbol{\lambda}}$ must be zero:

$$0 = \left.\frac{\partial \chi^2}{\partial \lambda_a}\right|_{\overline{\boldsymbol{\lambda}}} = 2 \left.\frac{\partial f_i(\boldsymbol{\lambda})}{\partial \lambda_a}\right|_{\overline{\boldsymbol{\lambda}}} (C^{-1})_{ij} \Big(f_j(\overline{\boldsymbol{\lambda}}) - \overline{y}_j \Big), \tag{9.40}$$

where we used the symmetric property of covariance matrix, $(C^{-1})_{ij} = (C^{-1})_{ji}$. The hessian matrix, denoted H, is

$$\begin{aligned} H_{ab} = \left.\frac{\partial^2 \chi^2}{\partial \lambda_a \partial \lambda_b}\right|_{\overline{\boldsymbol{\lambda}}} &= 2 \left.\frac{\partial f_i(\boldsymbol{\lambda})}{\partial \lambda_a \partial \lambda_b}\right|_{\overline{\boldsymbol{\lambda}}} (C^{-1})_{ij} \Big(f_j(\overline{\boldsymbol{\lambda}}) - \overline{y}_j \Big) \\ &\quad + 2 \left.\frac{\partial f_i(\boldsymbol{\lambda})}{\partial \lambda_a}\right|_{\overline{\boldsymbol{\lambda}}} (C^{-1})_{ij} \left.\frac{\partial f_j(\boldsymbol{\lambda})}{\partial \lambda_b}\right|_{\overline{\boldsymbol{\lambda}}}. \end{aligned} \tag{9.41}$$

Note that

$$\frac{\partial f_i(\overline{\boldsymbol{\lambda}})}{\partial \overline{\lambda}_a} = \left.\frac{\partial f_i(\boldsymbol{\lambda})}{\partial \lambda_a}\right|_{\overline{\boldsymbol{\lambda}}}. \tag{9.42}$$

Let us expand the fitted parameter $\overline{\boldsymbol{\lambda}}$ around the true value $\boldsymbol{\lambda}^*$.

$$\begin{aligned} \overline{\lambda}_a(\overline{\mathbf{y}}) &\simeq \lambda_a^* + \sum_i \left.\frac{\partial \overline{\lambda}_a}{\partial \overline{y}_i}\right|_{\mathbf{y}^*} (\overline{y}_i - y_i^*) \\ &\simeq \lambda_a^* + \sum_i \left.\frac{\partial \overline{\lambda}_a}{\partial \overline{y}_i}\right|_{\overline{\mathbf{y}}} (\overline{y}_i - y_i^*) \end{aligned} \tag{9.43}$$

where $y_i^* = f_i(\boldsymbol{\lambda}^*)$, $\mathbf{y}^*$ is a vector whose components are y_i^* and where $\overline{\mathbf{y}}$ is a vector whose components are $\overline{y}_i$. Using this approximation, the covariance $Cov(\overline{\lambda}_a, \overline{\lambda}_b)$ in Eq. (9.38) becomes

$$\begin{aligned} & Cov(\overline{\lambda}_a, \overline{\lambda}_b) \\ &\simeq \frac{1}{\int [d\overline{\mathbf{y}}] P(\overline{\mathbf{y}}|\boldsymbol{\lambda})} \int [d\overline{\mathbf{y}}] P(\overline{\mathbf{y}}|\boldsymbol{\lambda}) \sum_{kl} \left.\frac{\partial \overline{\lambda}_a}{\partial \overline{y}_k}\right|_{\mathbf{y}^*} \left.\frac{\partial \overline{\lambda}_b}{\partial \overline{y}_l}\right|_{\mathbf{y}^*} (\overline{y}_k - y_k^*)(\overline{y}_l - y_l^*) \qquad (9.44) \\ &\simeq \sum_{kl} \left.\frac{\partial \overline{\lambda}_a}{\partial \overline{y}_k}\right|_{\mathbf{y}^*} \left.\frac{\partial \overline{\lambda}_b}{\partial \overline{y}_l}\right|_{\mathbf{y}^*} Cov(\overline{y}_k, \overline{y}_l), \end{aligned} \tag{9.45}$$

where $Cov(\overline{y}_k, \overline{y}_l)$ is the covariance of the mean data, defined by

$$\begin{aligned} Cov(\overline{y}_k, \overline{y}_l) &= \langle (\overline{y}_k - y_k^*)(\overline{y}_l - y_l^*) \rangle_{P(\overline{y})} && (9.46) \\ &= \frac{1}{\int [d\overline{\mathbf{y}}] P(\overline{\mathbf{y}}|\boldsymbol{\lambda})} \int [d\overline{\mathbf{y}}] P(\overline{\mathbf{y}}|\boldsymbol{\lambda}) (\overline{y}_k - y_k^*)(\overline{y}_l - y_l^*). && (9.47) \end{aligned}$$

If a set of samples is given, $Cov(\overline{y}_k, \overline{y}_l)$ can be estimated by sample covariance, C_{kl}, defined in Eq. (9.36).

To find $\frac{\partial \overline{\lambda}_a}{\partial \overline{y}_k}$, one should differentiate Eq. (9.40) with respect to $\overline{y}_i$:

$$\begin{aligned} 0 &= 2 \frac{\partial^2 f_j(\overline{\boldsymbol{\lambda}})}{\partial \overline{\lambda}_a \partial \overline{\lambda}_b} \cdot \frac{\partial \overline{\lambda}_b}{\partial \overline{y}_i} \cdot (C^{-1})_{jk} \cdot (f_k(\overline{\boldsymbol{\lambda}}) - \overline{y}_k) \\ &\quad + 2 \frac{\partial f_j(\overline{\boldsymbol{\lambda}})}{\partial \overline{\lambda}_a} \cdot (C^{-1})_{jk} \cdot \frac{\partial f_k(\overline{\boldsymbol{\lambda}})}{\partial \overline{\lambda}_b} \cdot \frac{\partial \overline{\lambda}_b}{\partial \overline{y}_i} - 2 \frac{\partial f_j(\overline{\boldsymbol{\lambda}})}{\partial \overline{\lambda}_a} \cdot (C^{-1})_{jk} \delta_{ki}. \quad (9.48) \end{aligned}$$

Note that the matrix C is not a function of $\overline{\mathbf{y}}$ because it is the result obtained by integrating over $\overline{\mathbf{y}}$, as one can see in Eq. (9.46).

Using the hessian matrix, Eq. (9.41), the equation Eq. (9.48) can be simplified by

$$0 = H_{ab} \cdot \frac{\partial \overline{\lambda}_b}{\partial \overline{y}_i} - 2 \frac{\partial f_j(\overline{\boldsymbol{\lambda}})}{\partial \overline{\lambda}_a} \cdot (C^{-1})_{ji}. \quad (9.49)$$

It is rearranged by

$$\frac{\partial \overline{\lambda}_a}{\partial \overline{y}_i} = 2 (H^{-1})_{ab} \cdot \frac{\partial f_j(\overline{\boldsymbol{\lambda}})}{\partial \overline{\lambda}_b} \cdot (C^{-1})_{ji}. \quad (9.50)$$

Finally, the covariance of fitted parameters, Eq. (9.44), can be rewritten as

$$\begin{aligned} &Cov(\overline{\lambda}_a, \overline{\lambda}_b) \\ &= 4 (H^{-1})_{ac} \cdot \frac{\partial f_k(\overline{\boldsymbol{\lambda}})}{\partial \overline{\lambda}_c} (C^{-1})_{ki} \cdot (H^{-1})_{bd} \cdot \frac{\partial f_l(\overline{\boldsymbol{\lambda}})}{\partial \overline{\lambda}_d} (C^{-1})_{lj} \cdot C_{ij} \\ &= 4 (H^{-1})_{ac} \cdot \frac{\partial f_k(\overline{\boldsymbol{\lambda}})}{\partial \overline{\lambda}_c} (C^{-1})_{kl} \cdot \frac{\partial f_l(\overline{\boldsymbol{\lambda}})}{\partial \overline{\lambda}_d} \cdot (H^{-1})_{bd} \quad (9.51) \end{aligned}$$

When $a = b$, the square root of the covariance is the error of fitted parameters. If the fitting is good, in other words, $(f_i(\overline{\boldsymbol{\lambda}}) - \overline{y}_i)$ is small, the hessian matrix, Eq. (9.41), can be approximated as

$$H_{ab} \simeq 2 \frac{\partial f_i(\overline{\boldsymbol{\lambda}})}{\partial \overline{\lambda}_a} \cdot (C^{-1})_{ij} \cdot \frac{\partial f_j(\overline{\boldsymbol{\lambda}})}{\partial \overline{\lambda}_b} \quad (9.52)$$

In this case, the covariance of fitted parameters is simply estimated by

$$Cov(\overline{\lambda}_a, \overline{\lambda}_b) \simeq 2(H^{-1})_{ab}. \tag{9.53}$$

Usually, this is a very good approximation for the covariance of fitted parameters.

9.2. Constrained fitting

In section 9.1, we assumed that every values of fitting parameters are equivalent (no values are preferred). However, some parameters can be constrained by physics, e.g., the mass of a physical particle must be positive. The fitting becomes easier to converge if the physical constraints are included in the fitting. In 2002, Lepage *et al.* introduced a new fitting method to lattice community [115]. It is called the constrained fitting. Since it is based on the Bayesian statistics, it is also called the Bayesian method. It provides the way to include the physical constraints in the fitting procedure.

There are two different ways to approach to statistical inference: frequentist and Bayesian statistics. The frequentist considers the probability as a limiting frequency of many trials, which is a classical approach that defines the probability. For example, the probability of head for a coin tossing is the number of heads divided by the total number of coin tossing for infinitely many trials. In contrast, the probability in Bayesian statistics is a degrees of personal belief for an event. For example, we can say that the probability of head for a coin tossing is 1/2 before we try the coin tossing by our prior knowledge about the coin tossing. It can be different from that of frequentist's for unfair coins. However, the Bayesian statistics can say a probability based on our prior knowledge, before we actually try many number of coin tossing. After some trials, the Bayesian statistics updates their prior knowledge about the coin tossing and it converges to the probability of frequentist's, after long-run trials.

Let us rewrite the probability distribution of the theory, Eq. (9.1), including background information I:

$$P(\boldsymbol{\lambda}|\overline{\mathbf{y}}, I) = \frac{P(\overline{\mathbf{y}}|\boldsymbol{\lambda}, I)P(\boldsymbol{\lambda}|I)}{P(\overline{\mathbf{y}}|I)}. \tag{9.54}$$

In Bayesian statistics, $P(\boldsymbol{\lambda}|\overline{\mathbf{y}}, I)$ is called the *posterior probability*. Since the data does not depend on the theory, $P(\overline{\mathbf{y}}|I)$ can be considered as a constant. We know that the likelihood function, $P(\overline{\mathbf{y}}|\boldsymbol{\lambda})$ is nothing but

$$P(\overline{\mathbf{y}}|\boldsymbol{\lambda}) = \mathcal{N}' \exp\left[-\frac{1}{2}\chi^2\right], \tag{9.55}$$

where χ^2 is defined in Eq. (9.35). Given background information on the theory,

$P(\boldsymbol{\lambda}|I)$ is not a constant any more. The physics can constrain the values of fitting parameters in a finite range. The $P(\boldsymbol{\lambda}|I)$ is called *prior probability.*

Let the physical ranges of some parameters are known to be

$$\lambda_k = \tilde{\lambda}_k \pm \tilde{\sigma}_k. \tag{9.56}$$

If we assume that the prior distribution can be approximated by a normal distribution, it can be written as

$$P(\boldsymbol{\lambda}) = \mathcal{N}'' \exp\left[-\frac{1}{2}\chi^2_{\text{prior}}\right], \tag{9.57}$$

where χ^2_{prior} is defined by

$$\chi^2_{\text{prior}} = \sum_k \frac{(\lambda_k - \tilde{\lambda}_k)^2}{\tilde{\sigma}_k^2}. \tag{9.58}$$

Armed with these distributions, the posterior distribution is

$$\begin{aligned} L &= P(\boldsymbol{\lambda}|\overline{\mathbf{y}}) \\ &= \mathcal{N} \cdot \exp\left[-\frac{1}{2}\chi^2\right] \cdot \exp\left[-\frac{1}{2}\chi^2_{\text{prior}}\right] \qquad (9.59) \\ &\equiv \mathcal{N} \cdot \exp\left[-\frac{1}{2}\chi^2_{\text{aug}}\right], \qquad (9.60) \end{aligned}$$

where χ^2_{aug} is

$$\chi^2_{\text{aug}} = \chi^2 + \chi^2_{\text{prior}}. \tag{9.61}$$

To maximize the posterior distribution, one should minimize the augmented chi-square, χ^2_{aug}. The rest is the same as the ordinary χ^2 fitting.

One weakness of this method is the arbitrariness of the prior distribution. There is no compelling argument that the prior distribution can be approximated by normal distribution. The maximum entropy principle states that given mean and variance, the most ignorant form of the prior distribution is the normal distribution [126]. However, the prior information Eq. (9.56) is not a probabilistic mean and variance, but just a physical range of the parameters. Hence the maximum entropy principle does not give a precise distribution for the prior condition. The uncertainty of the prior distribution will be tested by changing the prior distribution, usually the variance of the distribution, and the effect of changing the prior distribution will be considered as a systematic error.

9.3. Finding fitting parameters

The χ^2 fitting is a procedure that determines the fitting parameters that minimizes χ^2. In general, the number of fitting parameters is so large that numerical methods which minimize a multidimensional function are needed.

If the fit function is linear to the fitting parameters, the fitting can be done by a simple matrix inversion. For general non-linear functions, however, it is much more complicated.

9.3.1. Fitting data to linear functions

The general χ^2 is defined in Eq. (9.24):

$$\chi^2 = (\overline{\mathbf{y}} - \mathbf{f})^T C^{-1} (\overline{\mathbf{y}} - \mathbf{f}). \tag{9.62}$$

Here, C is a covariance matrix, $\overline{\mathbf{y}}$ is a vector whose components are $\overline{y}(t)$, which is the average of data at t and $\mathbf{f}$ is a vector of $f(t)$, which are the function values at t. Let us consider a case that the fitting function $f(t)$ is linear to the fitting parameters,

$$f(t) = \sum_{a=1}^{m} \lambda_a F_a(t), \tag{9.63}$$

where m is the number of fitting parameters, λ_a are the fitting parameters and where $F_a(t)$ is a function of t. To find the minimum of χ^2, one needs to solve the equation:

$$\frac{\partial \chi^2}{\partial \lambda_a} = 0. \tag{9.64}$$

Applying to the Eq. (9.62) gives

$$\left(\frac{\partial \mathbf{f}}{\partial \lambda_a} \right)^T C^{-1} (\mathbf{f} - \overline{\mathbf{y}}) = 0. \tag{9.65}$$

Using the derivative of Eq. (9.63), $\frac{\partial \mathbf{f}}{\partial \lambda_a} = \mathbf{F}_a$, it can be rewritten as follows,

$$\sum_{b=1}^{m} (\mathbf{F}_a)^T \cdot C^{-1} \cdot \mathbf{F}_b \lambda_b = (\mathbf{F}_a)^T \cdot C^{-1} \cdot \overline{\mathbf{y}}, \tag{9.66}$$

$$A \cdot \boldsymbol{\lambda} = \mathbf{h}, \tag{9.67}$$

where

$$A_{ab} = (\mathbf{F}_a)^T \cdot C^{-1} \cdot \mathbf{F}_b \tag{9.68}$$

$$h_a = (\mathbf{F}_a)^T \cdot C^{-1} \cdot \overline{\mathbf{y}}. \tag{9.69}$$

Finally, one gets the best fitting parameters by inverting the matrix A,

$$\boldsymbol{\lambda} = A^{-1}\mathbf{h}. \tag{9.70}$$

9.3.2. Fitting data to nonlinear functions

For nonlinear fitting functions, there is no elegant way to find the best fitting parameters. One should employee a multidimensional function minimizer, such as Amoeba method, Conjugate gradient method or Newton method. These methods are general methods that find the minimum point of given function. In section 11, we review these methods in detail.

10. Covariance Fitting of Highly Correlated Data

Least χ^2 fitting is a popular method to determine unknown parameters in a theory from the numerical data such as lattice QCD results. One caveat is that the least χ^2 fitting works only if the fitting function is precise enough. This requirement becomes a troublesome problem if there is high correlation between data points because the high correlation generates small eigenvalues in the covariance matrix. When some eigenvalues of a covariance matrix are small, even a tiny error of fitting function can produce large χ^2 and spoil the fitting procedure. In some cases, the fitting function cannot be precise enough due to theoretical uncertainty such as truncated higher order terms in a perturbative series expansion.

10.1. Trouble with correlated data fitting

In this section, I will show the small eigenvalues of a covariance matrix can amplify the small error of fitting function and can yield large χ^2. Let us rewrite the χ^2 defined in Eq. (9.24) using the eigenvalue decomposition of the covariance matrix:

$$C^{-1} = \sum_{k=1}^{D} \frac{1}{\lambda_k} |\mathbf{v}_k\rangle\langle\mathbf{v}_k| \tag{10.1}$$

$$\chi^2 = (\overline{\mathbf{y}} - \mathbf{f})^T C^{-1} (\overline{\mathbf{y}} - \mathbf{f}) \tag{10.2}$$

$$= \langle \overline{\mathbf{y}} - \mathbf{f} | C^{-1} | \overline{\mathbf{y}} - \mathbf{f} \rangle \tag{10.3}$$

$$= \sum_{k=1}^{D} \frac{1}{\lambda_k} \left| \langle \overline{\mathbf{y}} - \mathbf{f} | \mathbf{v}_k \rangle \right|^2 \tag{10.4}$$

where D is the dimension of the covariance matrix and we use bra-ket notation for the convenience. If the fitting function is inexact, the average data point $\bar{\mathbf{y}}$ may deviate from the fitting function value $\mathbf{f}$ more than we expected. If the deviation in the direction of eigenvectors corresponding to the small eigenvalues is larger than the small eigenvalues,

$$\left| \langle \overline{\mathbf{y}} - \mathbf{f} | \mathbf{v}_l \rangle \right|^2 >> \lambda_l \,, \tag{10.5}$$

parameter	value
sea quarks	asqtad staggered fermions
valence quarks	HYP staggered fermions
gluons	Symanzik improved gluon action
geometry	$20^3 \times 64$
number of confs	671
number of meas	9 per conf
am_l/am_s	0.01/0.05
$1/a$	1662 MeV
α_s	0.3286 at $\mu = 1/a$
am_x, am_y	$0.005 \times n$ $(n = 1, 2, 3, \ldots, 10)$

Table 10.1: Parameters for the numerical study on the coarse (C3) ensemble. m_l is the light sea quark mass, m_s the strange sea quark mass, m_x the light valence quark mass, and m_y the strange valence quark mass. Here, "conf" and "meas" represent gauge configuration and measurement, respectively.

the χ^2 have larger values than we expected.

To be concrete, let us walk through a specific example of B_K. To demonstrate the problem, we choose the B_K data on a coarse (C3) ensemble out of MILC asqtad lattices using the notation of Ref. [46]. This ensemble is particularly a good sample, because it has relatively large statistics. The input parameters for the C3 ensemble is summarized in Table 10.1. In the X-fit, we fix $am_y = 0.05$ and select 4 data points of $am_x = 0.005$, 0.010, 0.015, 0.020 to fit to the functional form suggested by the SU(2) SChPT as in Ref. [46]. Hence, the covariance matrix C_{ij} is a 4×4 matrix.

$$C_{ij} = \begin{bmatrix} 1.42, & 0.661, & 0.398, & 0.274 \\ 0.661, & 0.392, & 0.271, & 0.204 \\ 0.398, & 0.271, & 0.205, & 0.165 \\ 0.274, & 0.204, & 0.165, & 0.138 \end{bmatrix} \times 10^{-5} \tag{10.6}$$

Its eigenvalues are

$$\lambda_i = \{\ 1.95 \times 10^{-5},\ 1.92 \times 10^{-6},\ 7.58 \times 10^{-8},\ 1.11 \times 10^{-9}\}\,. \tag{10.7}$$

The components of the matrix C_{ij} are between 1.42×10^{-5} and 1.38×10^{-6}. In the meanwhile, the smallest eigenvalue is smaller than the components by three orders of magnitude.

i	1	2	3	4
a_i	1.021(4)	0.5655(14)	0.1061(3)	0.01442(3)

Table 10.2: Eigenmode projection coefficients for $\bar{y}$.

Now let us look into the eigenvectors:

$$\mathbf{v}_1 = \begin{bmatrix} 0.837 \\ 0.429 \\ 0.276 \\ 0.200 \end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix} -0.508 \\ 0.387 \\ 0.542 \\ 0.546 \end{bmatrix}, \ \mathbf{v}_3 = \begin{bmatrix} 0.202 \\ -0.739 \\ 0.0725 \\ 0.639 \end{bmatrix}, \ \mathbf{v}_4 = \begin{bmatrix} -0.0378 \\ 0.347 \\ -0.790 \\ 0.503 \end{bmatrix}. \tag{10.8}$$

The eigenvector $\mathbf{v}_4$ corresponds to the smallest eigenvalue. This eigenmode dominates the χ^2 fitting completely. Let us expand $\bar{\mathbf{y}}$ in terms of eigenvectors as follows,

$$\bar{\mathbf{y}} = \sum_{i=1}^{4} a_i \mathbf{v}_i \,, \tag{10.9}$$

where a_i is the eigenmode projection coefficient. In Table 10.2, we show a_i for the data $\bar{y}$. The eigenmode $\mathbf{v}_1$ and $\mathbf{v}_2$ describes most of the average value of the data $\bar{\mathbf{y}}$ and $\mathbf{v}_4$ gives very small contribution to the average value of the data.

The fitting functional form suggested by the SU(2) staggered chiral perturbation theory (SChPT) is linear as follows,

$$f_{\rm th}(X) = \sum_{a=1}^{P} c_a F_a(X) \tag{10.10}$$

where c_a are the low energy constants (LECs) and F_a is a function of X which represents collectively X_P (pion squared mass of $\bar{x}x$), Y_P (pion squared mass of $\bar{y}y$), and so on. The details on F_a and X are given in SW-1.

In Figure 10.1, we show the fitting results with the full covariance matrix. As one can see, the fitting curve does not pass through the data points. Hence, the quality of fitting looks poor to our eyes. The χ^2 value is

$$\chi^2 = 7.2 \pm 5.4 \,.$$

If the degrees of freedom is comparable to the number of samples ($d \approx N$), the leading deviation of the T^2 distribution from the χ^2 distribution becomes of order $\mathcal{O}(1)$ and so we can not use the χ^2 distribution in this case, which is also pointed out in Ref. [127] in the context of distorted normal distribution. In Ref. [128], the sample size effect is systematically explained in terms of $1/N$ expansion.

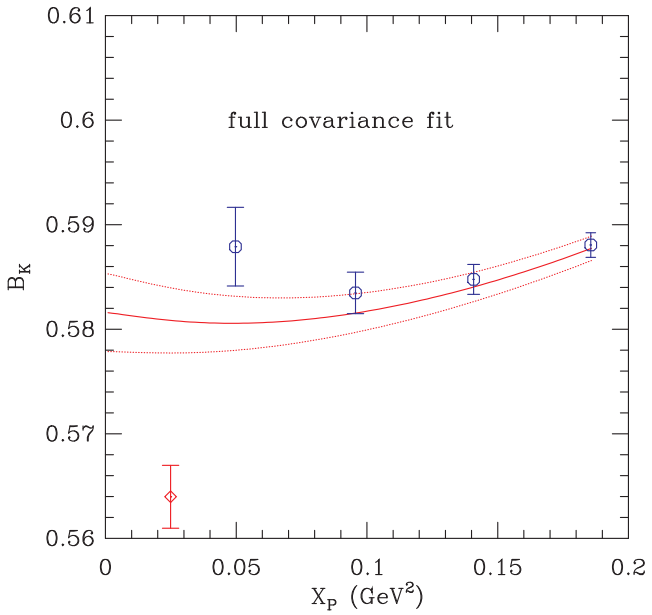


Figure 10.1: $B_K(1/a)$ vs. X_P on the C3 ensemble. The fit type is 4X3Y-NNLO in the SU(2) analysis. We fix $am_y = 0.05$. The red line represents the results of fitting with the full covariance matrix. The red diamond corresponds to the B_K value obtained by extrapolating m_x to the physical light valence quark mass after setting all the pion multiplet splittings to zero.

However, in our example of B_K, the degrees of freedom $d = 1$ and the number of data samples $N = 671$. Hence, the leading correction of the χ^2 distribution to the χ^2 distribution will be negligibly small ($\approx 0.15\%$) where the finite sample size effect in χ^2 is given in Eq. (9.32). In our example, the non-centrality parameter can be estimated by $\kappa \approx \chi^2 - d = 6.2$. Then, we can obtain $V[\chi^2]$ as follows,

$$V[\chi^2] \approx 2(d + 2\kappa) = 26.8 \tag{10.11}$$

Hence, the error of χ^2 is supposed to be $\sqrt{26.8} = 5.2$, which is reasonably consistent with the measured value 5.4. The κ is the non-centrality parameter which represents how much the fitting function deviates from the true mean values. In our example, $\kappa = 6.2$ turns out to be a rather large value which comes from the fact that the small deviation of our fitting function from the true value due to the truncation of the higher order terms in the series expansion of the SU(2) SChPT can be amplified dramatically if there are small eigenvalues in the covariance matrix.

Let us decompose the fitting function in terms of eigenmodes as follows,

$$\mathbf{f} = \sum_{i=1}^{4} b_i \mathbf{v}_i \tag{10.12}$$

To see how much the fitting function deviates from the data in a specific eigenmode, we define the difference, δ_i, as follow,

$$\delta_i = |a_i - b_i|. \tag{10.13}$$

As summarized in Table 10.3, the difference, δ_i, is 7.22×10^{-3} for $\mathbf{v}_1$, 2.40×10^{-3} for $\mathbf{v}_2$, 3.28×10^{-4} for $\mathbf{v}_3$, whereas it is only 9.69×10^{-6} for $\mathbf{v}_4$. Hence, the procedure of the least χ^2 fitting works hard for the coefficient of $\mathbf{v}_4$ but work less precisely for the coefficients of $\mathbf{v}_1$ and $\mathbf{v}_2$ mainly because the eigenvalue λ_4 is significantly smaller than λ_1 and λ_2. The irony is that the average data points have very small overlap with $\mathbf{v}_4$ while most of the average data points are dominated by $\mathbf{v}_1$ and $\mathbf{v}_2$. In this sense, the failure of the full covariance fitting is obviously due to the fact that the least χ^2 fitting tries to determine the coefficient of $\mathbf{v}_4$ component of the data very precisely but lose precision in determining the coefficients of the $\mathbf{v}_1$ and $\mathbf{v}_2$ components. As a consequence, the fitting curve misses the data points and the quality of fitting looks poor to our eyes.

i	1	2	3	4
b_i	1.014(4)	0.5679(11)	0.1058(3)	0.01443(3)
a_i	1.021(4)	0.5655(14)	0.1061(3)	0.01442(3)
$10^5 \cdot \delta_i$	722(270)	240(90)	32.8(123)	0.969(362)

Table 10.3: Eigenmode decomposition of $\mathbf{f}$ for the full covariance fitting.

10.2. Prescriptions

Here we present possible solutions for the problem raised in the previous section: diagonal approximation, cutoff method, Bayesian method and eigenmode shift (ES) method.

10.2.1. Diagonal approximation

One simple solution to the problem is to use the diagonal approximation (uncorrelated fitting) [127]. In this method, we neglect the off-diagonal covariance as follows:

$$C_{ij} = 0 \qquad \text{if } i \neq j\,. \tag{10.14}$$

In this way, the small eigenvalue problem disappears. However, as the size of covariance matrix becomes larger and larger, the number of ignored terms in the covariance matrix increases. Furthermore, since we modify the covariance matrix in this method, the minimized χ^2 dose not follow the χ^2 distribution and there is no way to estimate the quality of the fit. Hence, in the case of large degrees of freedom, one should be careful to use this method.

Let us get back to the example of B_K X-fit in C3 ensemble. The fitting results using the diagonal approximation are shown in Figure 10.2.

The fitting function $\mathbf{f}$ in the diagonal approximation is decomposed into eigenmodes of the full covariance matrix in Table 10.4. In this fit, the difference, δ_i, is 3.80×10^{-4} for v_1, 4.26×10^{-4} for v_2, 5.23×10^{-4} for v_3 and 4.85×10^{-4} for v_4. Here, note that the diagonal approximation method removes the small eigenvalues and so it takes all the eigenmodes, equally. As a result, the differences for all directions are less than or equal to 5.23×10^{-4}. Hence, the fitting looks quite reasonable to our eyes as one can see in Figure 10.2.

10.2.2. Cutoff method

Another way to deal with the small eigenvalue problem is to exclude the eigenmodes corresponding to the small eigenvalues from the fitting. A popular and systematic way of chopping away the eigenmodes is to set up such a cutoff that we project out the eigenvectors of those eigenvalues smaller than the cutoff

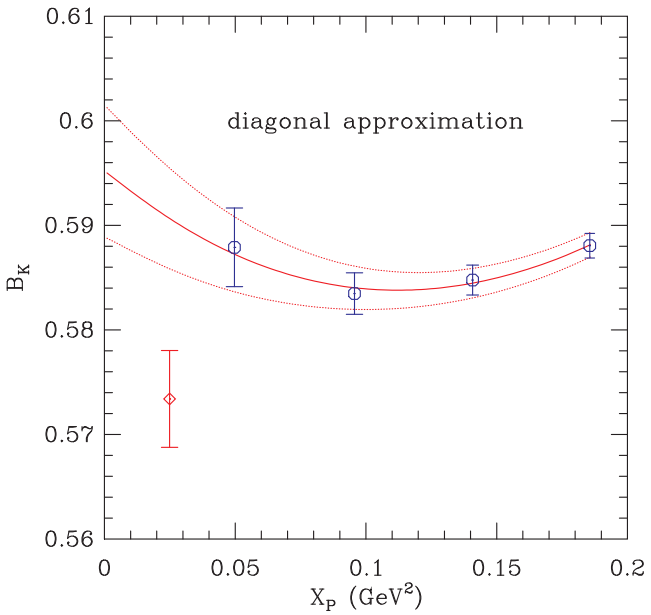


Figure 10.2: $B_K(1/a)$ vs. X_P on the C3 ensemble. All the parameters are the same as in Figure 10.1. Here, the red line represents the results of the uncorrelated fitting using the diagonal approximation.

i	1	2	3	4
b_i	1.021(4)	0.5659(13)	0.1056(3)	0.01490(18)
a_i	1.021(4)	0.5655(14)	0.1061(3)	0.01442(3)
$10^5 \cdot \delta_i$	38.0(142)	42.6(159)	52.3(195)	48.5(181)

Table 10.4: Eigenmode decomposition of $\mathbf{f}$ for the fitting with diagonal approximation.

into a null space of the inverse covariance matrix C_{ij}^{-1}:

$$\frac{1}{\lambda_k} = 0 \qquad \text{if } \lambda_k < \lambda_{\text{cut}} \tag{10.15}$$

$$C^{-1} = \sum_{k=1}^{D} \frac{1}{\lambda_k} |\mathbf{v}_k\rangle\langle\mathbf{v}_k| \tag{10.16}$$

$$\tag{10.17}$$

We call this the cutoff method. When we decompose the covariance matrix into eigenvalues and eigenvectors, we usually use the SVD (singular value decomposition) algorithm. Hence it is also called SVD method. A number of lattice QCD groups [129, 130, 131] use this method in their fitting.

However, a major drawback of the cutoff method is that we cannot give the physical meaning to the quality of fit, which is normally reflected in the minimized value of χ^2. In Section 10.2.5.2, we show that the probability distribution of χ^2 defined in cutoff method is the χ^2 distribution with degrees of freedom equal to $D - P - R$. Here, D is the number of data points, P is the number of fitting parameters and R is the number of removed eigenmodes. However, even though we know the distribution of the χ^2 in cutoff method, we cannot measure the quality of fit through the minimized value of χ^2. This is because we remove some eigenmodes and the χ^2 in the cutoff method is orthogonal to the removed eigenmodes. As you know, the physical χ^2 has $D - P$ degrees of freedom, while the χ^2 of the cut-off method possesses $D - P - R$ degrees of freedom. Unfortunately, the missing degrees of freedom, R are physical.

The situation becomes even worse when we use the resampling method, such as jackknife method or bootstrap method. When the size of covariance matrix is large, we might lose a control over the number of the small eigenvalues that we remove. In other words, in one jackknife sample, we may remove two small eigenvalues and in another jackknife sample, we may remove three of them. During the procedure, the definition of χ^2 is shifting from one to another.

Now let us walk through an example to demonstrate how it works. In our example of B_K, we have three parameters to determine from the fit. Hence, it is possible to remove only one eigenmode v_4 out of the four by setting $1/\lambda_4 = 0$. In Figure 10.3, we show the results of the covariance fitting using the cutoff method. It is amusing to see how good it works. The results are consistent with those in Figure 10.2. In Table 10.5 we decompose the fitting function $\mathbf{f}$ obtained using the cutoff method into eigenmodes of the full covariance matrix. In this case, the difference, δ_i, is zero for $i = 1, 2, 3$ and is 8.21×10^{-4} in v_4. In this method, we do not care about the v_4 eigenmode at all. As a result, the fitting quality looks quite good to our eyes, which is quite consistent with that of the diagonal approximation.

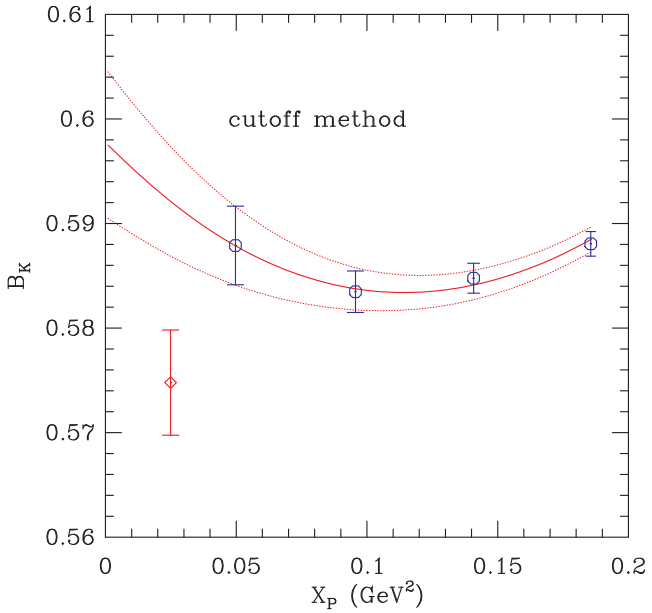


Figure 10.3: $B_K(1/a)$ vs. X_P on the C3 ensemble. All the parameters are the same as in Figure 10.2. Here, the red line represents the results of the covariance fitting after removing the smallest eigenvalue using the cutoff method.

i	1	2	3	4
b_i	1.021(4)	0.5655(14)	0.1061(3)	0.01524(30)
a_i	1.021(4)	0.5655(14)	0.1061(3)	0.01442(3)
$10^5 \cdot \delta_i$	0(0)	0(0)	0(0)	82.1(307)
β_i	0.7589(18)	0.2328(18)	0.008190(69)	0.0001690(63)

Table 10.5: Eigenmode decomposition of $\mathbf{f}$ for the fitting with the cutoff method.

10.2.3. Eigenmode shift method

So far, all the methods are based on the philosophy that we may manipulate or modify the covariance matrix. This philosophy is very dangerous in a sense that the modification results in a missing information in physics. The missing information is highly physical. Hence, it is not a good idea to modify the covariance matrix. The only degrees of freedom that we have is to modify the fitting function, but not the covariance matrix.

We know that the whole trouble comes from the inexact fitting function that has small error in the direction of eigenmodes corresponding to the small eigenvalues. Hence, we can think of a new fitting function $\mathbf{f}'$ defined as follows,

$$\mathbf{f}'(X) = \mathbf{f}(X) + \sum_k \eta_k \mathbf{v}_k \,, \tag{10.18}$$

where the summation over the k runs over the eigenmodes corresponding to the small eigenvalues and η_i are tiny parameters which can be determined using the Bayesian method. The Bayes theorem [126] says that

$$P(A|\mathbb{D}, I) \;\propto\; P(\mathbb{D}|A, I) \times P(A|I) \tag{10.19}$$

Here, A represents our theoretical hypothesis, $\mathbb{D}$ is the data, and I corresponds to the background information. Note that $P(\mathbb{D}|A, I)$ means the probability that we obtain the data set of $\mathbb{D}$ if A and I are given to us. We know the conditional likelihood function of $P(\mathbb{D}|A, I)$, which is nothing but

$$P(\mathbb{D}|A, I) \;\propto\; \exp\left(-\frac{\chi^2}{2}\right) \tag{10.20}$$

$$\chi^2 \;=\; [\bar{\mathbf{y}} - \mathbf{f}']^T C^{-1} [\bar{\mathbf{y}} - \mathbf{f}'] \tag{10.21}$$

as explained in Ref. [126].

In addition, if we impose constraints on the prior that $a_\eta - \sigma_\eta \lesssim \eta \lesssim a_\eta + \sigma_\eta$, then the prior becomes the following:

$$P(A|I) \;\propto\; \exp\left(-\frac{\chi^2_{\rm prior}}{2}\right) \tag{10.22}$$

$$\chi^2_{\rm prior} \;=\; \sum_k \frac{(\eta_k - a_{\eta_k})^2}{\sigma_{\eta_k^2}} \tag{10.23}$$

i	1	2	3	4
b_i	1.021(4)	0.5655(14)	0.1061(3)	0.01524(30)
a_i	1.021(4)	0.5655(14)	0.1061(3)	0.01442(3)
$10^5 \cdot \delta_i$	1.88(70)	0.624(233)	0.0855(319)	81.9(306)

Table 10.6: Eigenmode decomposition of $\mathbf{f}$ for the fitting with the ES method.

Then, we obtain the posterior pdf as follows:

$$P(A|\mathbb{D}, I) \propto \exp\left(-\frac{\chi^2_{\rm aug}}{2}\right) \tag{10.24}$$

$$\chi^2_{\rm aug} = \chi^2 + \chi^2_{\rm prior} \tag{10.25}$$

The Bayesian principle [126] is to determine the fitting parameters by maximizing the posterior pdf: $P(A|\mathbb{D}, I)$. This is equivalent to minimizing the $\chi^2_{\rm aug}$. We call this the eigenmode shift (ES) method [53, 54].

Let us switch the gear to the B_K example. Since the $\mathbf{v}_4$ eigenmode is troublesome, we can introduce a shifting parameter on the $\mathbf{v}_4$ direction. From the SChPT, the neglected highest order term in the $f(X)$ is

$$X^2(\ln(X))^2 \approx 0.006$$

where $X = X_P/\Lambda^2 \approx 0.02$. The only constraint on the coefficient c_4 of this term is that $c_4 = 0 \pm 1$, but we do not know the sign of c_4. Hence, we set the prior condition by $a_\eta = 0$ and $\sigma_\eta = 0.006$.

Then we perform the full covariance fitting with the extra fitting parameter, η, by minimizing $\chi^2_{\rm aug}$ (the Bayesian principle). The obtained η is

$$\eta = -0.00082(31)\,.$$

When we do the extrapolation to the physical pion mass, we use only the $f(X)$ function, dropping out the η term, which is too small to make any difference at any rate.

In Figure 10.4, we show the fitting results obtained using the ES method. In our example, tiny correction proportional to η makes the fitting results that pass through the average data point.

In Table 10.6, we show the eigenmode decomposition of $\mathbf{f}$ when we use the ES method in fitting. As one can see in the table, the δ_i is smaller by order of magnitude compared with the diagonal approximation for $i = 1, 2, 3$. The only non-trivial component is δ_4, which is taken care of by the shift parameter η. As a result, in Table 10.7, the δ_4 for $\mathbf{f}'$ becomes negligibly small by the shift

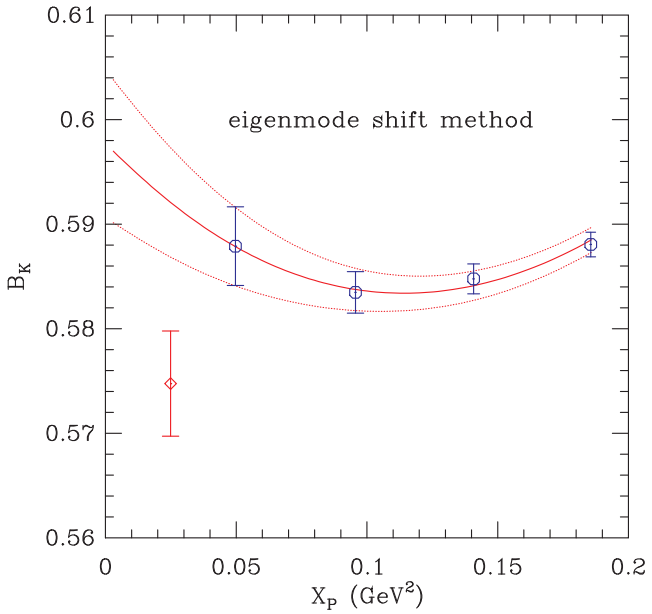


Figure 10.4: $B_K(1/a)$ vs. X_P on the C3 ensemble. All the parameters are the same as in Figure 10.2. Here, the red line represents the results of the eigenmode shift method.

i	1	2	3	4
b_i	1.021(4)	0.5655(14)	0.1061(3)	0.01442(3)
a_i	1.021(4)	0.5655(14)	0.1061(3)	0.01442(3)
$10^5 \cdot \delta_i$	1.88(70)	0.624(233)	0.0855(319)	0.00252(94)

Table 10.7: Eigenmode decomposition of $\mathbf{f}'$ for the fitting with the ES method.

parameter η. The success of the fitting can be checked against the hypothesis that $|\eta| \lesssim 0.006$. The results of fitting say that $\eta = -0.00082(31)$, which is highly consistent with the hypothesis. In this sense, the Bayesian prior is quite reasonable and self-consistent.

Let us turn to the issue of quality of fitting and physical interpretation of χ^2_{aug}. In a naive sense of physical interpretation, we may count the prior condition as one of data points, and we consider the shift parameter η as a new parameter in the fitting. Hence, in this interpretation, the effective number of data points is $\tilde{D} = D + 1$, and the effective number of unknown parameters of the fitting function is $\tilde{P} = P + 1$. Accordingly, the effective degrees of freedom becomes $\tilde{d} = \tilde{D} - \tilde{P} = D - P = d$. Therefore, the χ^2_{aug} follows the χ^2 distribution with $\chi^2_{\text{aug}} \approx \tilde{d} \pm \sqrt{2\tilde{d}} = d \pm \sqrt{2d}$, as explained in Section 10.2.5.3.

However, in our Bayesian prior information, we did not use the statistical information for a_η and σ_η, but the optimal range of possible value of η to set the a_η and σ_η. (i.e. $a_\eta \neq \mathcal{E}(\eta)$ and $\sigma_\eta \neq \sqrt{\mathcal{V}(\eta)}$). Hence, the choice of a_η and σ_η could be larger or smaller (overestimated or underestimated) than the statistical value of them. As a consequence, our estimate of $\chi^2_{\rm aug}$ could be smaller or larger than the normal value of $\chi^2_{\rm aug} \approx \tilde{d}$. Hence, in this approach of Bayesian method, we cannot rely on the strict statistical interpretation of $\chi^2_{\rm aug}$. One good news is that the probability interpretation and the model selection is still possible with $\chi^2_{\rm aug}$, on the basis of the Bayesian statistics [126]. In other words, we can say that the fitting parameter values that gives smaller $\chi^2_{\rm aug}$ is the better results, based on the Bayesian statistics. This is an important point because it justifies the fitting procedure that finds the most probable fitting parameters by minimizing $\chi^2_{\rm aug}$.

In our example of B_K, the $\chi^2_{\rm aug}$/dof for the ES method is 0.019(14).

In the limit of $\sigma_\eta \to \infty$, we can remove the prior condition completely, which we call unconstrained ES method. In Section 10.2.3.1, we prove that the unconstrained ES method is equivalent to the cutoff method. However, the original ES method is quite different from the cutoff method in the following sense. The effective number of degrees of freedom for the ES method is $\tilde{d} = \tilde{D} - \tilde{P} = 1$ while that for the cutoff method is $\tilde{d} = (D-1) - P = 0$. In addition, the ES method is rigorously based on the Bayesian method and is subject to the probability interpretation. However, the cutoff method does not allow for the probability interpretation mainly because it modifies the covariance matrix. In addition, in the ES method, we modify the fitting function by the shift parameter η, which we can monitor and gives us an estimate of how much we are changing the fitting function. In the case of the cutoff method, we do not know how much of the fitting function we are dumping into the null space of the covariance matrix. In order to illustrate the difference between the ES method and the cutoff method, we provide a pedagogical and heuristic example in Section 10.2.6, in which the ES method works well, but the cutoff method and the diagonal approximation fail manifestly.

10.2.3.1. Equivalence of cutoff method and unconstrained ES method

Unconstrained ES method is the ES method whose shifting parameter, η, is not constrained by the Bayesian prior. It is the same as the ES method whose prior condition is set to $\sigma_\eta = \infty$. In this section, we would like to prove that the unconstrained ES method is equivalent to the cutoff method.

The χ^2 of the cutoff method can be written in the following form:

$$\chi^2_{\rm cut} = \langle \bar{\mathbf{y}} - \mathbf{f} | C'^{-1} | \bar{\mathbf{y}} - \mathbf{f} \rangle \,. \tag{10.26}$$

Here, $\mathbf{f}$ is a vector representing the fitting function value,

$$f_i = f(X_i)\,,$$

$\bar{\mathbf{y}}$ is the D-dimensional vector of average data points and C'^{-1} is the inverse covariance matrix of the cutoff method. If R number of eigenmodes, denoted by $S+1 \leq i \leq D$ with $S \equiv D-R$, are removed from the covariance matrix by the cutoff method, C'^{-1} can be written as follows,

$$\begin{aligned} C'^{-1} &= C^{-1} - \sum_{i=S+1}^{D} \frac{1}{\lambda_i} |\mathbf{v}_i\rangle\langle\mathbf{v}_i| \\ &= \sum_{i=1}^{S} \frac{1}{\lambda_i} |\mathbf{v}_i\rangle\langle\mathbf{v}_i| \qquad (10.27) \\ C^{-1} &= \sum_{k=1}^{D} \frac{1}{\lambda_k} |\mathbf{v}_k\rangle\langle\mathbf{v}_k|\,. \end{aligned}$$

Using the eigenmode decomposition given in Eq. (10.27), we can rewrite $\chi^2_{\rm cut}$ as

$$\chi^2_{\rm cut} = \sum_{i=1}^{S} \frac{1}{\lambda_i} \left[\langle \bar{\mathbf{y}} - \mathbf{f} | \mathbf{v}_i \rangle\right]^2\,, \qquad (10.28)$$

while χ^2 is defined by

$$\begin{aligned} \chi^2 &= \langle \bar{\mathbf{y}} - \mathbf{f} | C^{-1} | \bar{\mathbf{y}} - \mathbf{f} \rangle \\ &= \sum_{i=1}^{D} \frac{1}{\lambda_i} \left[\langle \bar{\mathbf{y}} - \mathbf{f} | \mathbf{v}_i \rangle\right]^2\,. \qquad (10.29) \end{aligned}$$

Here, C^{-1} is the inverse of the full covariance matrix.

The χ^2 of the unconstrained ES method can be written in the following form:

$$\chi^2_{\rm UES} = \langle \bar{\mathbf{y}} - \mathbf{f}' | C^{-1} | \bar{\mathbf{y}} - \mathbf{f}' \rangle\,, \qquad (10.30)$$

where $\mathbf{f}'$ is defined by

$$\mathbf{f}' = \mathbf{f} + \sum_{i=S+1}^{D} \eta_i \mathbf{v}_i\,. \qquad (10.31)$$

The $\chi^2_{\rm UES}$ can be expanded as follows,

$$\begin{aligned}\chi^2_{\rm UES} &= \chi^2 - 2\sum_{i=S+1}^{D}\frac{1}{\lambda_i}\eta_i\langle\bar{\mathbf{y}}-\mathbf{f}|\mathbf{v}_i\rangle + \sum_{i=S+1}^{D}\frac{1}{\lambda_i}\eta_i^2 \\ &= \chi^2 - \sum_{i=S+1}^{D}\frac{1}{\lambda_i}\left[\langle\bar{\mathbf{y}}-\mathbf{f}|\mathbf{v}_i\rangle\right]^2 + \sum_{i=S+1}^{D}\frac{1}{\lambda_i}\Big[\eta_i - \langle\bar{\mathbf{y}}-\mathbf{f}|\mathbf{v}_i\rangle\Big]^2 \\ &= \chi^2_{\rm cut} + \sum_{i=S+1}^{D}\frac{1}{\lambda_i}\Big[\eta_i - \langle\bar{\mathbf{y}}-\mathbf{f}|\mathbf{v}_i\rangle\Big]^2 . \end{aligned} \tag{10.32}$$

Minimizing $\chi^2_{\rm UES}$ yields

$$\eta_i = \langle\bar{\mathbf{y}}-\mathbf{f}|\mathbf{v}_i\rangle\,, \tag{10.33}$$

and it removes the last term in Eq. (10.32). Hence, the minimizing $\chi^2_{\rm cut}$ gives the same results as those of the minimizing $\chi^2_{\rm UES}$. This proves our claim.

10.2.4. Bayesian method

When we obtain the fitting function, we use the staggered chiral perturbation theory to expand it in powers of p^2, a^2, and m_q. In the series expansion, we must truncate the higher order terms because we cannot include an arbitrary number of terms in the fitting function. One constraint is that we have only 4 data points of B_K for the SU(2) analysis. Hence, the fitting function can have at most 3 unknown parameters, if we want to perform the normal least χ^2 fitting based on the multivariate statistics theory. This means that we can include all the next to leading order (NLO) terms and one additional term at the next to next to leading order (NNLO). It is the fitting function of Eq. (10.10) that we obtain following this premature logical path. This looks fine as long as the truncated terms at the higher order are under control such that the full covariance fitting works well. However, in our case of the SU(2) analysis on B_K, the full covariance fitting fails manifestly because the data have such a high precision that the truncated terms of the higher order are required to fit the data. We cannot add higher order terms to the fitting function in a normal sense of the multivariate statistical theory. Hence, the situation is checkmate as it is. The question is how we can get out of this trap. A natural answer is the Bayesian method [115, 126].

In the Bayesian method, the prior condition behaves in the fitting as if it is one of the data points as explained in the previous subsection. Hence, it is possible to add n higher order terms as long as we impose m prior conditions on the fitting with $n \leq m$. In practice, we impose the same number of prior conditions as that of the higher order terms added to the fitting function as

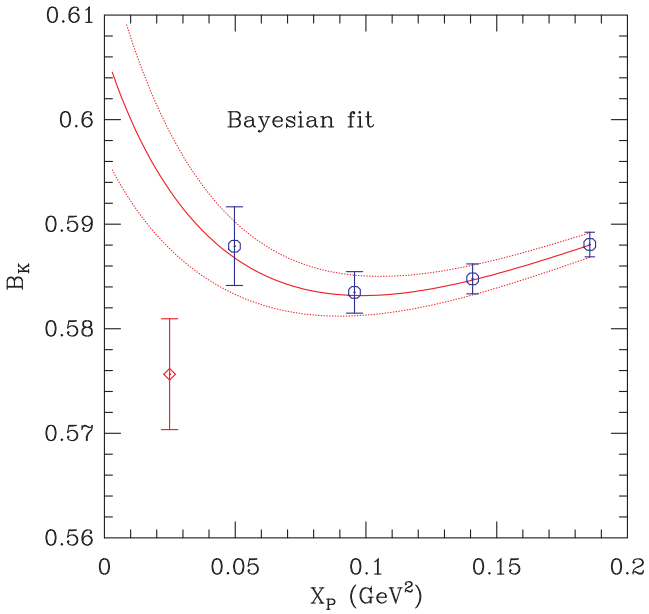


Figure 10.5: $B_K(1/a)$ vs. X_P on the C3 ensemble. All the parameters are the same as in Figure 10.2. Here, the red line represents the results of the Bayesian method.

follows. The fitting function has three additional terms at higher order:

$$f_{\rm th}^{\rm B}(X) \;=\; f_{\rm th}(X) + c_4^b\ X^2\,(\ln X)^2 + c_5^b\ X^2\,(\ln X) + c_6^b\ X^3\,. \qquad (10.34)$$

We impose the prior conditions on the fitting through the prior probability as in the previous subsection:

$$\chi^2_{\rm prior} = \sum_{k=4}^{6} \frac{(c_k^b - a_{k;B})^2}{\sigma^2_{k;B}} \qquad (10.35)$$

Since we know that $c_k^b = 0 \pm 1$, we may choose $a_{k;B} = 0$ and $\sigma_{k;B} = 1$. In the Bayesian method, we use $\chi^2_{\rm aug}$ instead of χ^2, in the analysis, which is defined as

$$L \;\equiv\; \log(P(A|\mathbb{D}, I)) \qquad (10.36)$$

$$\chi^2_{\rm aug} \;=\; (-2)L = \chi^2 + \chi^2_{\rm prior} \qquad (10.37)$$

where $P(A|\mathbb{D}, I)$ is the posterior pdf [126]. The Bayesian principle is that we determine the fitting parameters such that they maximize the posterior pdf or minimize $\chi^2_{\rm aug}$. The main advantage of the Bayesian method is that it allows for probability interpretation and model selection as explained in Ref. [126].

i	1	2	3	4
b_i	1.020(5)	0.5659(13)	0.1060(3)	0.01442(3)
a_i	1.021(4)	0.5655(14)	0.1061(3)	0.01442(3)
$10^5 \cdot \delta_i$	110(41)	36.5(136)	5.00(187)	0.147(55)
β_i	0.7583(16)	0.2334(16)	0.008194(69)	0.0001515(12)

Table 10.8: Eigenmode decomposition of $f_{\rm th}^{\rm B}$ for the fitting with the Bayesian method.

In Figure 10.5, we show the fitting results obtained using the Bayesian method. The fitting has $\chi^2_{\rm aug} = 1.09(81)$. The effective number of the data points is $\tilde{D} = 4 + 3 = 7$ and the number of the unknown parameters is $\tilde{P} = 6$. Hence, the effective number of degrees of freedom is $\tilde{d} = \tilde{D} - \tilde{P} = 1 = d$, the same as the full covariance fitting. In Table 10.8, we decompose the fitting function $f_{\rm th}^{\rm B}$ obtained using the Bayesian method in terms of the eigenmodes of the full covariance matrix. The fitting looks fine to our eyes.

A major advantage of the Bayesian method is that it does NOT touch the covariance matrix at all unlike the cutoff method and the diagonal approximation.

In the Bayesian method, we need to gauge the sensitivity of the fitting to the prior condition. In our case, we change the prior condition as follows:

$$\sigma_{k;B} = 1 \to 2 \tag{10.38}$$

while we keep $a_{k;B}$ unchanged. The fitting results are changed as follows.

$$B_K = 0.5757(53) \to 0.5772(58) \tag{10.39}$$

$$\chi^2_{\rm aug} = 1.09(81) \to 0.31(23) \tag{10.40}$$

The mean value of B_K is shifted by 0.28σ while the error bar increases by 9%. Hence, the difference between the ES method and the Bayesian method is 0.18σ, which is smaller than the systematic error due to the ambiguity in the prior condition. Since both the ES and Bayesian methods are based on the Bayes theorem, they are equivalent to each other in that sense. When we obtain the higher order terms in Eq. (10.34), we use the continuum chiral perturbation theory but not the staggered chiral perturbation theory. Hence, the functional form of each higher order term is approximate and not exact. From this standpoint, we cannot claim that the Bayesian method is better than the ES method. Therefore, we decide to quote the difference between the results of ES and Bayesian methods as our systematic error due to the ambiguity in the covariance fitting in the SW-2, and to choose the results of the Bayesian method as the central value.

10.2.5. Probability distribution of minimized χ^2

10.2.5.1. Distribution of χ^2 for the full covariance fitting

As an introduction, we review the probability distribution of the minimized value of χ^2 for the full covariance fitting. Here, we assume that we have large number of data samples so that the sample covariance matrix is well determined. Generally, the fitting function is inaccurate and the inaccuracy leads us to the non-central χ^2-distribution. The non-central χ^2-distribution is defined as follows.

Let us consider a set of independent random variables,

$$\{X_1, X_2, X_3, \cdots, X_k\}\,.$$

The random variables, X_i, follow the normal distribution with mean μ_i and variance 1, denoted by $X_i \sim \mathcal{N}(\mu_i, 1)$. If we define a new random variable Q as follows,

$$Q = \sum_{i=1}^{k} X_i^2\,, \tag{10.41}$$

then Q is distributed according to the non-central χ^2-distribution with degrees of freedom k and with non-centrality parameter

$$\kappa = \sum_{i=1}^{k} \mu_i^2\,. \tag{10.42}$$

The χ^2 of the full covariance fitting can be written in the following form:

$$\chi^2 = \langle \bar{\mathbf{y}} - \mathbf{f} | C^{-1} | \bar{\mathbf{y}} - \mathbf{f} \rangle\,. \tag{10.43}$$

Here, $\mathbf{f}$ is a vector representing the fitting function value,

$$f_i = f(X_i)\,,$$

$\bar{\mathbf{y}}$ is the D-dimensional vector of average data points and C^{-1} is the inverse of the covariance matrix. The expectation value and covariance of the vector $\bar{\mathbf{y}} - \mathbf{f}$ are

$$\mathcal{E}[\bar{\mathbf{y}} - \mathbf{f}] = \boldsymbol{\phi} \tag{10.44}$$

$$\mathcal{C}\big[\bar{\mathbf{y}} - \mathbf{f}\big] = \mathcal{C}\big[\bar{\mathbf{y}}\big] = C\,. \tag{10.45}$$

Here, the vector $\boldsymbol{\phi}$ represents the error of the fitting function. If the fitting function is exact, $\boldsymbol{\phi} = \mathbf{0}$. According to the multi-dimensional central limit theorem, the average of data, $\bar{y}$, is distributed as multivariate normal distribution. Hence, $(\bar{\mathbf{y}} - \mathbf{f})$ follows the multivariate normal distribution with mean vector $\boldsymbol{\phi}$

and the covariance matrix C, $(\bar{\mathbf{y}} - \mathbf{f}) \sim \mathcal{N}(\boldsymbol{\phi}, C)$.

Let M be a non-singular square matrix satisfying

$$MCM^T = I\,, \tag{10.46}$$

where I is an identity matrix. If we define a new vector $\mathbf{Y}$ by

$$\mathbf{Y} = M(\bar{\mathbf{y}} - \mathbf{f})\,, \tag{10.47}$$

then the expectation value and covariance of Y is

$$\begin{aligned} \mathcal{E}[\mathbf{Y}] &= M\boldsymbol{\phi} && (10.48)\\ \mathcal{C}\big[\mathbf{Y}\big] &= MCM^T = I\,. && (10.49) \end{aligned}$$

Since M is non-singular, the transformed vector $\mathbf{Y}$ is also distributed according to the multivariate normal distribution (Proof is given in Ref. [123]). Hence, $\mathbf{Y}$ is distributed according to the multivariate normal distribution with mean $M\boldsymbol{\phi}$ and covariance matrix I. Note that identity covariance matrix implies mutual independence of Y_i.

In terms of $\mathbf{Y}$, the χ^2, given in Eq. (10.43), can be rewritten by

$$\begin{aligned} \chi^2 &= \langle \mathbf{Y}|\mathbf{Y}\rangle \\ &= \sum_{i=1}^{D} Y_i^2\,. \end{aligned} \tag{10.50}$$

Because Y_i are independently distributed as normal distribution with mean $[M\boldsymbol{\phi}]_i$ and variance 1, χ^2 is distributed according to the non-central χ^2-distribution. If the number of fitting parameters is P, degrees of freedom of the distribution is $D - P$. The non-centrality parameter of the distribution is

$$\begin{aligned} \kappa &= \sum_{k=1}^{D} \Big([M\boldsymbol{\phi}]_k\Big)^2 \\ &= (M\boldsymbol{\phi})^T(M\boldsymbol{\phi}) = \boldsymbol{\phi}^T M^T M\boldsymbol{\phi} = \boldsymbol{\phi}^T C^{-1}\boldsymbol{\phi}\,, \end{aligned} \tag{10.51}$$

where we used $M^T M = C^{-1}$ at the last equality.

10.2.5.2. Distribution of χ^2 for the cutoff method

As described in Section 10.2.3.1, the χ^2 of the cutoff method can be written in the following form:

$$\chi^2_{\rm cut} = \langle \bar{\mathbf{y}} - \mathbf{f}|C'^{-1}|\bar{\mathbf{y}} - \mathbf{f}\rangle\,, \tag{10.52}$$

where C'^{-1} is the inverse covariance matrix of the cutoff method, which can be written by

$$C'^{-1} = C^{-1} - \sum_{i=S+1}^{D} \frac{1}{\lambda_i} |\mathbf{v}_i\rangle\langle\mathbf{v}_i| \tag{10.53}$$

$$C^{-1} = \sum_{k=1}^{D} \frac{1}{\lambda_k} |\mathbf{v}_k\rangle\langle\mathbf{v}_k| \,.$$

Here, R is the number of removed eigenmodes, and $S = D - R$. Let M be a non-singular square matrix satisfying

$$MCM^T = I \,, \tag{10.54}$$

where I is an identity matrix. Using the eigenmode decomposition, it is

$$M = \sum_{k=1}^{D} \frac{1}{\sqrt{\lambda_k}} |\mathbf{e}_k\rangle\langle\mathbf{v}_k| \tag{10.55}$$

$$M^{-1} = \sum_{k=1}^{D} \sqrt{\lambda_k} |\mathbf{v}_k\rangle\langle\mathbf{e}_k| \,, \tag{10.56}$$

where $\mathbf{e}_k$ is a unit vector in k-direction. In terms of the new vector $\mathbf{Y} = M(\bar{\mathbf{y}} - \mathbf{f})$, the $\chi^2_{\rm cut}$ becomes

$$\chi^2_{\rm cut} = \langle\mathbf{Y}|(M^T)^{-1} C'^{-1} M^{-1}|\mathbf{Y}\rangle$$

$$= \sum_{k=1}^{D} Y_k^2 - \sum_{i=S+1}^{D} Y_i^2 \tag{10.57}$$

$$= \sum_{k=1}^{S} Y_k^2 \,. \tag{10.58}$$

Here, we used the following relation,

$$(M^T)^{-1} C'^{-1} M^{-1} = I - \sum_{i=S+1}^{D} |\mathbf{e}_i\rangle\langle\mathbf{e}_i| \,. \tag{10.59}$$

As mentioned in Section 10.2.5.1, Y_i are independently distributed as normal distribution with mean $[M\boldsymbol{\phi}]_i$ and variance 1, $Y_i \sim \mathcal{N}\big([M\boldsymbol{\phi}]_i, 1\big)$. Hence, Eq. (10.57) shows that $\chi^2_{\rm cut}$ is distributed according to the non-central χ^2-distribution with degrees of freedom $D - P - R$ and with non-centrality pa-

rameter

$$\begin{aligned}\kappa &= \sum_{k=1}^{D}\Big([M\boldsymbol{\phi}]_k\Big)^2 - \sum_{i=S+1}^{D}\Big([M\boldsymbol{\phi}]_i\Big)^2 \\ &= \sum_{k=1}^{S}\frac{1}{\lambda_k}\Big[\langle \mathbf{v}_k|\boldsymbol{\phi}\rangle\Big]^2, \end{aligned} \tag{10.60}$$

where $\boldsymbol{\phi}$, which is a vector representing the error of fitting function, is defined in Eq. (10.44). Here, D is the number of data points, P is the number of fitting parameters and R is the number of removed eigenmodes. If we assume that the fitting function is exact ($\boldsymbol{\phi} = \mathbf{0}$), χ^2_{cut} follows the χ^2-distribution with degrees of freedom: $S - P = D - P - R$.

10.2.5.3. Distribution of χ^2 for the ES method

In Section 10.2.3.1, we show that unconstrained ES method is the same as the cutoff method. The probability distribution of the minimized value of χ^2 in cutoff method is derived in Section 10.2.5.2 and it is the χ^2-distribution with degrees of freedom $D - P - R$, if we assume that the fitting function is good enough. Hence, the probability distribution of the minimized value of χ^2 in unconstrained ES method is the χ^2-distribution with degrees of freedom $D-P-R$. Here, R is the number of shifting eigenvectors that modify the fitting function.

The deformation of the degrees of freedom in unconstrained ES method can be considered as a consequence of adding new fitting parameters that control the shifting eigenvectors. We add R number of shifting parameters, $\{\eta_i\}$ with $i = 1, 2, \ldots, R$, and it increases the number of fitting parameters to $P + R$. As a result, the degrees of freedom becomes $D - P - R$.

In the normal ES method, we constrain the shifting parameters by Bayesian method. Hence, distribution of the augmented χ^2 in ES method can be interpreted that of in the Bayesian constrained fitting.

By the Bayes theorem,

$$P(A|\mathbb{D}, I) \propto P(\mathbb{D}|A, I) \times P(A|I) \tag{10.61}$$

Here, $P(\mathbb{D}|A, I)$ is the likelihood function, which can be expressed as

$$P(\mathbb{D}|A, I) \propto \exp\left(-\frac{\chi^2}{2}\right) \tag{10.62}$$

In addition, let us assume that the prior probability $P(A|I)$ can be written as

$$P(A|I) \propto \exp\left(-\frac{\chi^2_{\text{prior}}}{2}\right) \tag{10.63}$$

The effective number of degrees of freedom in the ES method is $\tilde{d} = \tilde{D} - \tilde{P} = D - P = d$, where $\tilde{D} = D + R$ and $\tilde{P} = P + R$ since there are R number of the prior conditions imposed on η_i and there exists as much increase in the number of fitting parameters by R. Hence, the $\chi^2_{\rm aug} = \chi^2 + \chi^2_{\rm prior}$ must follow the normal χ^2 distribution with $\chi^2_{\rm aug} = \tilde{d} \pm \sqrt{2\tilde{d}} = d \pm \sqrt{2d}$. Therefore, the $\chi^2_{\rm aug}$ of the ES method follows the same normal χ^2 distribution as that of the full covariance fitting.

10.2.6. An example of fitting with random data

In this section, we will show an example to see how the fitting methods, given in section 10.2, work.

First, let us explain how the data has been generated. The true mean μ_i of the data is

$$\mu_i = f_{\rm true}(x_i) = \frac{1}{1 - x_i} \tag{10.64}$$

We choose 7 data points ($D = 7$) such that $x_i = 0.2 \times \dfrac{(i-1)}{6}$ with $i = 1, 2, 3, \ldots, 7$. We also choose the eigenvalues of the true covariance matrix as follows,

$$\lambda_i^{\Gamma} = 10^{-\frac{2}{3}(i-1)} \tag{10.65}$$

for $i = 1, 2, 3, \ldots, 7$. Then, we generate the random numbers to construct the eigenvectors which are orthonormal to each other by construction. Hence, we can obtain the true covariance matrix Γ_{ij} from the eigenvalues and eigenvectors. Note that the covariance matrix for the sample mean C_{ij} is related to the Γ_{ij} as follows:

$$C_{ij} = \frac{1}{N}\Gamma_{ij}$$

Hence, the eigenvalues of C_{ij} is smaller by a factor of N as follows.

$$\lambda_i^C \simeq \frac{1}{N}\lambda_i^{\Gamma}$$

For the notational convenience, we drop out the superscipt for the eigenvalues in the discussion in this section.

Next, we generate the data y_i following the probability distribution: $\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}, \Gamma)$. Here, we use the numerical algorithm introduced in Ref. [132] to generate the data of y_i according to the pdf:

$$P(\mathbf{y}|\boldsymbol{\mu}, \Gamma) = Z \exp[-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu}) \cdot \Gamma^{-1} \cdot (\mathbf{y} - \boldsymbol{\mu})] \tag{10.66}$$

$$Z = \frac{1}{(2\pi)^D \det(\Gamma)^{1/2}} \tag{10.67}$$

In this way, we collect 1000 random data samples for y_i (i.e. $N = 1000$).

We know that the true fitting function can be rewritten as follows,

$$f_{\text{true}}(x) = \frac{1}{1-x} = 1 + x + x^2 + \cdots + x^n + \cdots \tag{10.68}$$

We choose the trial fitting function as follows:

$$f_{\text{trial}}(x) = c_1 + c_2 x + c_3 x^2 \tag{10.69}$$

In addition, we also choose the trial fitting function for the Bayesian method as

$$f_{\text{Bayes}}(x) = c_1 + c_2 x + c_3 x^2 + c_4^b x^3 \,, \tag{10.70}$$

where we will impose the prior condition on c_4^b.

We fit the data using the fitting methods such as the full covariance fitting, diagonal approximation, cut-off method, ES method, and Bayesian method. In the cut-off method, we remove lowest two eigenmodes by imposing the cut-off ($\lambda_i/\lambda_1 \geq 1.0 \times 10^{-3}$), where λ_1 is the largest eigenvalue. In the ES method, we introduce two shift parameters η_1 and η_2 in the direction of the two smallest eigenmodes. We impose the following prior condition on η_i.

$$\begin{aligned} \eta_i &\sim \mathcal{N}(a_i, \sigma_i^2) \\ a_i &= 0 \\ \sigma_i &= 0.001 \end{aligned} \tag{10.71}$$

Here, the highest terms of the truncated are of order $\mathcal{O}(x^3)$. By assuming that the coefficient is $\mathcal{O}(1)$, we can estimate the size of truncated terms approximately as

$$\sigma_i \approx 1.0 \times x^3 \approx 1.0 \times (0.1)^3 = 0.001 \,,$$

where we take the average of x_i as the x value.

In the Bayesian method, we impose the following prior condition on c_4^b.

$$\begin{aligned} c_4^b &\sim \mathcal{N}(a_4, \sigma_4^2) \\ a_4 &= 0 \\ \sigma_4 &= 1 \end{aligned} \tag{10.72}$$

In Figure 10.6, we show the fitting results obtained using various fitting methods. As one can see in the plot, it is very clear that there is something seriously wrong with the cut-off method. In Table 10.9, we present results for various fitting methods. The results of the full covariance fitting is what we want to achieve in the fitting business ultimately. As one can see in the table, the diagonal approximation and the cut-off method fail, manifestly. In other words, errors of the c_i parameters are dramatically larger than those of the full co-

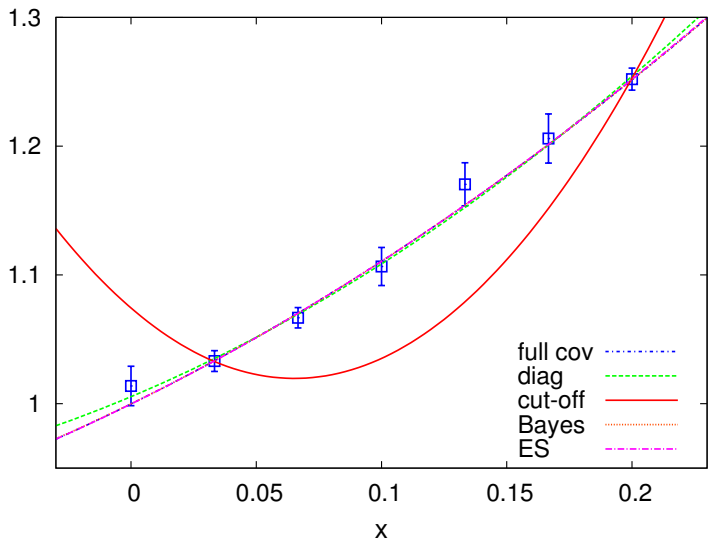


Figure 10.6: Comparison of various fitting methods: full covariance fitting (full cov), diagonal approximation (diag), cut-off method (cut-off), Bayesian method (Bayes) and eigenmode shift method (ES).

variance fitting, which indicates a symptom of a complete failure in fitting. In addition, we notice that both the ES method and the Bayesian method work very well in good agreement with the expectation, while the ES method looks marginally better in this example. In the case of the ES method, we obtain the following results for the shift parameters:

$$\begin{aligned} \eta_1 &= 0.00015 \pm 0.00154 \\ \eta_2 &= -0.00023 \pm 0.00039\,, \end{aligned}$$

which are highly self-consistent with the prior condition.

This is only an example with random samples. However, it shows that the ES method is totally different from the cut-off method, although the unconstrained ES method is the same as the cut-off method. It also warns us that we should be careful when we use the diagonal approximation or the cut-off method.

fit type	c_1	c_2	c_3	χ^2/dof
full cov	0.9998(19)	0.957(41)	1.51(24)	0.65(81)
diag	1.0055(105)	0.818(211)	2.13(97)	0.45(96)
cut-off	1.0743(808)	-1.67(285)	12.8(123)	0.70(118)
Bayes	0.9999(19)	0.950(42)	1.59(27)	0.63(79)
ES	0.9998(19)	0.954(42)	1.53(24)	0.60(76)

Table 10.9: Fitting results of various fitting methods: The fitting details are explained in the text.

11. Multidimensional Function Minimizer

Suppose that a function depending on one or more parameters is given and one wants to find the minimum point. If the number of independent parameters is one, specialized one-dimensional minimization algorithms([133]) would be the best option. If the number of parameters is less than 8, Amoeba method, Conjugate gradient method or Newton method are the possible options. In this section, we review these methods. If the number of parameters is large or the function space has many local minima, one may use Simulated annealing method. [133]

Amoeba method is a general purpose multidimensional function minimizer that only requires function values, not its derivatives. Also there are algorithms that using derivatives to find the minimum of the function, such as conjugate gradient function minimizer or Newton function minimizer. Here, we review these three minimizers. Conjugate gradient minimizer requires first-order derivatives and Newton minimizer requires up to second-order derivatives.

One thing one should note is that in the fitting procedure, the parameters should be normalized before using these function minimizing algorithms. If the parameters have very different order of magnitudes, it might be a drawback to finding the solution. One way to deal with these parameters is that divide the parameters by normalization constant (i.e. initial guess of the parameter) before inputting to the minimizer and multiply the normalization constant after minimization procedure. If the order of parameter values are varying in the minimization procedure, one should normalize the parameters for every iteration of the minimization procedure. In that case, the parameter values of the previous iteration would be a good normalization constant.

Basically these methods cannot distinguish the local minimum and the global minimum. Sometimes the minimizer falls into a local minimum. To avoid the local minimum, one may find many local minima from many other initial starting point and take the lowest point. If one believes the initial guess is precise enough, one may try two different kinds of minimizations such as amoeba method and Newton method. It might help to avoid local minimum if one takes the parameters that yields lower function value.

11.1. Amoeba method

Amoeba method (sometimes called *downhill simplex method*) is a general multidimensional function minimization algorithm proposed by John Nelder and

Roger Mead in 1965. [134]. This method requires only the function values, not the derivatives. Even the methods using derivatives are usually faster than this method, this method is easy to implement and broadly applicable.

A simplex is a generalized notion of triangle in two dimensional space and tetrahedron in three dimensional space. If the parameter space is n dimensional, one can make a simplex with $n+1$ vertices. Amoeba method generates a simplex around the initial starting point and updates the vertex having the largest function value (the highest point). First, it reflects the highest point through the plane specified by the other n points. If the reflection is successful (if the function value of the reflected point is the lowest), it expands (go further) along the direction when the expansion gives smaller function value. Then replace the highest point by the new point and do the same procedure with highest point of new simplex. If the reflection is a failure, shrink (go backward) the new point along the direction. If the contraction is successful, replace the highest point by the new point. If not, shrink all the vertices except the lowest nearer to the lowest point and do the procedure again.

It stops if it reaches the maximum number of iterations of all the vertices converges enough to a point. Sometimes it can converge to a point that is not the minimum. Therefore it is a good idea to restart the minimization procedure with replaced initial starting point by the converged point. If the converged point is a minimum, the new minimization procedure converges to the same point quickly so it is not expensive.

11.2. Conjugate gradient algorithm

To understand the conjugate gradient function minimizer, one should understand the conjugate gradient algorithm first. Conjugate gradient algorithm is an iterative method for solving following linear equations,

$$A\mathbf{x} = \mathbf{b} \tag{11.1}$$

where A is an hermitian, positive definite ($^{\forall}\mathbf{x} \neq 0, \quad \mathbf{x}^{\dagger}A\mathbf{x} > 0$) $n \times n$ complex matrix and $\mathbf{x}$, $\mathbf{b}$ are vectors in n dimensional complex vector space. A and $\mathbf{b}$ are known, $\mathbf{x}$ is unknown. [135] This method is one of the most successful methods that solves the linear equation when A is sparse and large-dimension matrix.

To solve the linear algebraic equation, $\mathbf{x}$ is updated as follow.

$$\mathbf{x}_{(i+1)} = \mathbf{x}_{(i)} + \alpha_{(i)}\mathbf{d}_{(i)} \tag{11.2}$$

where $\alpha_{(i)}$ is a constant, $\mathbf{d}_{(i)}$ is a direction vector and $\mathbf{x}_{(0)}$ is an initial vector. We also update $\mathbf{d}_{(i)}$ by

$$\mathbf{d}_{(i+1)} = \mathbf{r}_{(i+1)} + \beta_{(i+1)}\mathbf{d}_{(i)} \tag{11.3}$$

where $\beta_{(i)}$ is a constant, $\mathbf{d}_{(0)} = \mathbf{r}_{(0)}$ and $\mathbf{r}_{(i)}$ is a residual defined by

$$\mathbf{r}_{(i)} = \mathbf{b} - A\mathbf{x}_{(i)}. \tag{11.4}$$

There are two conditions for determining the constants $\alpha_{(i)}$ and $\beta_{(i)}$.

$$\mathbf{d}^{\dagger}_{(i)}\mathbf{r}_{(j)} = 0 \qquad (i < j) \tag{11.5}$$

$$\mathbf{d}^{\dagger}_{(i)}A\mathbf{d}_{(j)} = 0 \qquad (i \neq j). \tag{11.6}$$

These conditions determine the constants $\alpha_{(i)}$ and $\beta_{(i)}$, which make the iterative process converges to the solution within n, the dimension of A.

11.2.1. Calculation of $\alpha_{(i)}$

The first condition Eq. (11.5) determine the $\alpha_{(i)}$. The equation Eq. (11.5) gives

$$\mathbf{d}^{\dagger}_{(i)}\mathbf{r}_{(i+1)} = 0. \tag{11.7}$$

Combining Eq. (11.4) and Eq. (11.2) yields

$$\mathbf{r}_{(i+1)} = \mathbf{r}_{(i)} - \alpha_{(i)}A\mathbf{d}_{(i)}. \tag{11.8}$$

Inserting Eq. (11.8) into Eq. (11.7) makes

$$\mathbf{d}^{\dagger}_{(i)}\mathbf{r}_{(i)} - \alpha_{(i)}\mathbf{d}^{\dagger}_{(i)}A\mathbf{d}_{(i)} = 0. \tag{11.9}$$

The $\alpha_{(i)}$ is

$$\alpha_{(i)} = \frac{\mathbf{d}^{\dagger}_{(i)}\mathbf{r}_{(i)}}{\mathbf{d}^{\dagger}_{(i)}A\mathbf{d}_{(i)}} \tag{11.10}$$

This equation can be written in another form. We know

$$\begin{aligned} \mathbf{d}^{\dagger}_{(i)} &= \mathbf{r}^{\dagger}_{(i)} + \beta^{*}_{(i)}\mathbf{d}^{\dagger}_{(i-1)} \\ \mathbf{d}^{\dagger}_{(i)}\mathbf{r}_{(i)} &= \mathbf{r}^{\dagger}_{(i)}\mathbf{r}_{(i)} + \beta^{*}_{(i)}\mathbf{d}^{\dagger}_{(i-1)}\mathbf{r}_{(i)} \\ \mathbf{d}^{\dagger}_{(i)}\mathbf{r}_{(i)} &= \mathbf{r}^{\dagger}_{(i)}\mathbf{r}_{(i)} \end{aligned} \tag{11.11}$$

where we used Eq. (11.5) for $\mathbf{d}^{\dagger}_{(i-1)}\mathbf{r}_{(i)} = 0$. Using this result,

$$\alpha_{(i)} = \frac{\mathbf{r}^{\dagger}_{(i)}\mathbf{r}_{(i)}}{\mathbf{d}^{\dagger}_{(i)}A\mathbf{d}_{(i)}}. \tag{11.12}$$

One easily finds that the $\alpha_{(i)}$ is real.

11.2.2. Calculation of $\beta_{(i+1)}$

From the second condition Eq. (11.6), following equation can be derived:

$$\mathbf{d}^{\dagger}_{(i+1)} A \mathbf{d}_{(i)} = 0. \tag{11.13}$$

Inserting Eq. (11.3) into Eq. (11.13) yields

$$[\mathbf{r}^{\dagger}_{(i+1)} + \beta^{*}_{(i+1)} \mathbf{d}^{\dagger}_{(i)}] A \mathbf{d}_{(i)} = 0.$$

The $\beta^{*}_{(i+1)}$ is

$$\beta^{*}_{(i+1)} = -\frac{\mathbf{r}^{\dagger}_{(i+1)} A \mathbf{d}_{(i)}}{\mathbf{d}^{\dagger}_{(i)} A \mathbf{d}_{(i)}}. \tag{11.14}$$

Let us simplify Eq. (11.14). From Eq. (11.3), the following equation can be derived:

$$\mathbf{r}^{\dagger}_{(i+1)} \mathbf{d}_{(i)} = \mathbf{r}^{\dagger}_{(i+1)} \mathbf{r}_{(i)} + \beta_{(i)} \mathbf{r}^{\dagger}_{(i+1)} \mathbf{d}_{(i-1)}. \tag{11.15}$$

Using the first condition, Eq. (11.5), $\mathbf{r}^{\dagger}_{(i+1)} \mathbf{d}_{(i)}$ and $\mathbf{r}^{\dagger}_{(i+1)} \mathbf{d}_{(i)}$ are zero. That implies

$$\mathbf{r}^{\dagger}_{(i+1)} \mathbf{r}_{(i)} = 0. \tag{11.16}$$

Multiplying $\mathbf{r}^{\dagger}_{(i+1)}$ to Eq. (11.8) yields,

$$\mathbf{r}^{\dagger}_{(i+1)} \mathbf{r}_{(i+1)} = \mathbf{r}^{\dagger}_{(i+1)} \mathbf{r}_{(i)} - \alpha_{(i)} \mathbf{r}^{\dagger}_{(i+1)} A \mathbf{d}_{(i)}.$$

Using Eq. (11.16), the first term of the right hand side can be deleted:

$$\mathbf{r}^{\dagger}_{(i+1)} A \mathbf{d}_{(i)} = -\frac{1}{\alpha_{(i)}} \mathbf{r}^{\dagger}_{(i+1)} \mathbf{r}_{(i+1)}.$$

From this result, one finds the $\beta_{(i+1)}$ of Eq. (11.14) becomes

$$\begin{aligned} \beta^{*}_{(i+1)} &= \frac{1}{\alpha_{(i)}} \frac{\mathbf{r}^{\dagger}_{(i+1)} \mathbf{r}_{(i+1)}}{\mathbf{d}^{\dagger}_{(i)} A \mathbf{d}_{(i)}} \\ &= \frac{\mathbf{r}^{\dagger}_{(i+1)} \mathbf{r}_{(i+1)}}{\mathbf{r}^{\dagger}_{(i)} \mathbf{r}_{(i)}} = \beta_{(i+1)}, \end{aligned} \tag{11.17}$$

where we used Eq. (11.12) replacing $\alpha_{(i)}$. One can easily see that the $\beta_{(i+1)}$ is real.

11.2.3. Convergence

Conjugate gradient method solves the linear algebraic equation iteratively within n (the dimension of A) iterations. A set of A-conjugate(A-orthogonal, $\mathbf{d}^\dagger A\mathbf{d} = 0$) vectors is used as direction vectors. Taking each direction vectors only once, the method will reach the exact solution. Therefore only n iterations are needed to get the solution.

From the first condition, Eq. (11.5),

$$\begin{aligned} \mathbf{d}_{(i)}^\dagger \mathbf{r}_{(j)} &= 0 \qquad (i<j) \\ [\mathbf{r}_{(i)}^\dagger + \beta_{(i)}^* \mathbf{d}_{(i-1)}^\dagger]\mathbf{r}_{(j)} &= 0 \qquad (i<j) \\ \mathbf{r}_{(i)}^\dagger \mathbf{r}_{(j)} + \beta_{(i)}^* \mathbf{d}_{(i-1)}^\dagger \mathbf{r}_{(j)} &= 0 \qquad (i<j). \end{aligned} \tag{11.18}$$

Using Eq. (11.5), the term $\mathbf{d}_{(i-1)}^\dagger \mathbf{r}_{(j)}$ can be deleted. It generates the result:

$$\mathbf{r}_{(i)}^\dagger \mathbf{r}_{(j)} = 0 \qquad (i<j). \tag{11.19}$$

This equation shows the orthogonality of residual vectors. Then one can makes a set of n-linearly independent residual vectors.

$$\{\mathbf{r}_{(0)}, \mathbf{r}_{(1)}, \cdots, \mathbf{r}_{(n-1)}\} \quad : \quad n\text{-linearly indep. vectors} \tag{11.20}$$

The vector $\mathbf{d}_{(i)}$ can be generated by adding a new vector $\mathbf{r}_{(i)}$, Eq. (11.3). Therefore the set of direction vectors spans n-dimensional complex vector space:

$$\{\mathbf{d}_{(0)}, \mathbf{d}_{(1)}, \cdots, \mathbf{d}_{(n-1)}\} \quad \text{spans} \quad \mathbb{C}^n \tag{11.21}$$

From the second condition, Eq. (11.6), each direction vectors are A-conjugate.

Let us define a error vector $\mathbf{e}_{(i)}$,

$$\mathbf{e}_{(i)} = \mathbf{x}_{(i)} - \mathbf{x} \tag{11.22}$$

where $\mathbf{x}$ is the exact solution of $A\mathbf{x} = \mathbf{b}$. If $\mathbf{e}_{(i)} = 0$, the $\mathbf{x}_{(i)}$ and the exact solution is the same. We will show the convergence by showing $\mathbf{e}_{(n)} = 0$. Because the error vector is a vector in n-dimensional complex vector space, the vector can be represented using Eq. (11.21):

$$\mathbf{e}_{(0)} = \sum_{j=0}^{n-1} \delta_{(j)} \mathbf{d}_{(j)} \tag{11.23}$$

where $\delta_{(j)}$ are some constants. Multiplying $(\mathbf{d}^\dagger_{(k)}A)$ yields

$$\begin{aligned}\mathbf{d}^\dagger_{(k)}A\mathbf{e}_{(0)} &= \sum_{j=0}^{n-1}\delta_{(j)}\mathbf{d}^\dagger_{(k)}A\mathbf{d}_{(j)} \\ &= \delta_{(k)}\mathbf{d}^\dagger_{(k)}A\mathbf{d}_{(k)} \end{aligned} \tag{11.24}$$

where we used the second condition Eq. (11.6), $(\mathbf{d}^\dagger_{(i)}A\mathbf{d}_{(j)} = 0(i \neq j))$. Therefore $\delta_{(k)}$ is

$$\begin{aligned}\delta_{(k)} &= \frac{\mathbf{d}^\dagger_{(k)}A\mathbf{e}_{(0)}}{\mathbf{d}^\dagger_{(k)}A\mathbf{d}_{(k)}} \\ &= \frac{\mathbf{d}^\dagger_{(k)}A[\mathbf{e}_{(0)} + \sum_{i=0}^{k-1}\alpha_{(i)}\mathbf{d}_{(i)}]}{\mathbf{d}^\dagger_{(k)}A\mathbf{d}_{(k)}}, \end{aligned} \tag{11.25}$$

where we also used the second condition Eq. (11.6). Using the relation

$$\begin{aligned}\mathbf{e}_{(k)} &= \mathbf{x}_{(k)} - \mathbf{x} \\ &= \mathbf{x}_{(k-1)} + \alpha_{(k-1)}\mathbf{d}_{(k-1)} - \mathbf{x} \\ &= \mathbf{x}_{(k-2)} + \alpha_{(k-2)}\mathbf{d}_{(k-2)} + \alpha_{(k-1)}\mathbf{d}_{(k-1)} - \mathbf{x} \\ &= \cdots \\ &= \mathbf{x}_{(0)} - \mathbf{x} + \alpha_{(0)}\mathbf{d}_{(0)} + \cdots + \alpha_{(k-1)}\mathbf{d}_{(k-1)} \\ &= \mathbf{e}_{(0)} + \sum_{i=0}^{k-1}\alpha_{(i)}\mathbf{d}_{(i)}, \end{aligned} \tag{11.26}$$

$\delta_{(k)}$ can be simplified by

$$\delta_{(k)} = \frac{\mathbf{d}^\dagger_{(k)}A\mathbf{e}_{(k)}}{\mathbf{d}^\dagger_{(k)}A\mathbf{d}_{(k)}}. \tag{11.27}$$

Using $A\mathbf{x} = \mathbf{b}$, one finds

$$\mathbf{r}_{(k)} = \mathbf{b} - A\mathbf{x}_{(k)} = A(\mathbf{x} - \mathbf{x}_{(k)}) = -A\mathbf{e}_{(k)}. \tag{11.28}$$

Therefore $\delta_{(k)}$ can be written as

$$\begin{aligned}\delta_{(k)} &= -\frac{\mathbf{d}^\dagger_{(k)}\mathbf{r}_{(k)}}{\mathbf{d}^\dagger_{(k)}A\mathbf{d}_{(k)}} \\ &= -\alpha_{(k)}. \end{aligned} \tag{11.29}$$

From this result, the term $-\alpha_{(i)}$ can be replaced by $\delta_{(i)}$. Combining Eq. (11.26) and Eq. (11.23), $\mathbf{e}_{(k)}$ becomes

$$\begin{aligned}\mathbf{e}_{(k)} &= \mathbf{e}_{(0)} + \sum_{i=0}^{k-1} \alpha_{(i)}\mathbf{d}_{(i)} \\ &= \sum_{j=0}^{n-1} \delta_{(j)}\mathbf{d}_{(j)} + \sum_{i=0}^{k-1} \alpha_{(i)}\mathbf{d}_{(i)} \\ &= \sum_{j=k}^{n-1} \delta_{(j)}\mathbf{d}_{(j)} \end{aligned} \tag{11.30}$$

Finally, this result shows that after n iterations, the error $\mathbf{e}_{(n)}$, which defined by $(\mathbf{x}_{(n)} - \mathbf{x})$, becomes zero and $\mathbf{x}_{(n)}$ is the exact solution.

If there are only $m(< n)$ different eigenvalues due to duplicated eigenvalues, it only needs m iterations to get the solution. It is also converges to the solution quickly if the eigenvalues of the matrix A are gathered near some values. We do not prove it here but give a intuitive explanation. In section 11.3.1, we show that finding a solution of $A \cdot \mathbf{x} = \mathbf{b}$ is equivalent to finding a solution of $\nabla f(\mathbf{x}) = 0$, where A, $\mathbf{b}$ and $f(\mathbf{x})$ are defined in Eq. (11.35). Here, A is a hessian matrix so its eigenvalues represent the shape of the form of functional space. If some eigelvalues are equal or very close, the shape of functional space is isotropic in those directions. For simplicity, let us consider a case that the function is an elliptic paraboloid that the two eigenvalues are the same. Then one may reach to the minimum point only by one movement toward the gradient direction. Generally, the more isotropic functional space is, the easier to reach the minimum point. Therefore duplicated or gathered eigenvalues yield quick convergence.

11.2.4. Practical implementation

To solve a linear equation iteratively, one must check the initial point $\mathbf{x}_{(0)}$ and the stopping condition. If a guessed solution is given, it would be used as an initial point. If there is no guessed solutions, any initial point (e.g. $\mathbf{x}_{(0)} = 0$) will work. When it reaches the solution, the residual vector $\mathbf{r}_{(i)}$ becomes zero. Because the method uses the $\mathbf{r}^{\dagger}_{(i)}\mathbf{r}_{(i)}$ as a denominator to calculate $\beta_{(i+1)}$, Eq. (11.14), the iteration must be stopped when it gets $\mathbf{r}_{(i)} = 0$. If only a 'good enough' solution is needed, introducing an error tolerance ϵ will be a good choice. The iteration may be stopped the iteration when $|\mathbf{r}_{(i)}| < \epsilon|\mathbf{r}_{(0)}|$. Algorithm 1 shows the skeleton of implementation code.

Algorithm 1 Skeleton code of CG algorithm

$\mathbf{r} \leftarrow \mathbf{b} - A\mathbf{x}$
$\mathbf{d} \leftarrow \mathbf{r}$
$\delta_{new} \leftarrow \mathbf{r}^\dagger \mathbf{r}$
$\delta_0 \leftarrow \delta_{new}$
for $i = 1$ **to** N_{max} **do**
  $\alpha \leftarrow \frac{\delta_{new}}{\mathbf{d}^\dagger A \mathbf{d}}$
  $\mathbf{x} \leftarrow \mathbf{x} + \alpha \mathbf{d}$
  $\mathbf{r} \leftarrow \mathbf{b} - A\mathbf{x}$
  $\delta_{old} \leftarrow \delta_{new}$
  $\delta_{new} \leftarrow \mathbf{r}^\dagger \mathbf{r}$
  $\beta \leftarrow \frac{\delta_{new}}{\delta_{old}}$
  $\mathbf{d} \leftarrow \mathbf{r} + \beta \mathbf{d}$
  if $\delta_{new} \leq \epsilon^2 \delta_0$ **then**
    break for loop
  end if
end for

11.2.5. Variants

Conjugate gradient method only works for positive definite and hermitian matrix A. For a non-hermitian matrices, there is an algorithm called *biconjugate gradient method* (abbreviated as BiCG). It is a generalized CG method. In 1992, a variant of BiCG is proposed by H. A. van der Vorst. [136] It is called *biconjugate gradient stabilized method* (abbreviated as BiCGSTAB) and it has faster convergence than BiCG method.

Sometimes *preconditioning* accelerates the convergence of conjugate gradient method and its variants. A *preconditioner* M is a matrix whose inverse will be multiplied to both sides of Eq. (11.1):

$$M^{-1}A\mathbf{x} = M^{-1}\mathbf{b} \tag{11.31}$$
$$A'\mathbf{x} = \mathbf{b}'. \tag{11.32}$$

If the new matrix A' has better spectral properties (duplicated or gathered eigenvalues) than the original matrix A, then CG solves the new linear equation (Eq. (11.32)) faster than the original equation (Eq. (11.1)). If one takes M as A itself, the CG yields solution by one iteration. However, this is practically impossible because if one knows the inverse of A, one dose not need to use CG method to solve the linear equation. Therefore, one should take M as a matrix that is an approximation of A but easy to convert. For example, Jacobi preconditioner is a diagonal matrix whose diagonals are the same values to those of A.

11.3. Function minimization using CG

Conjugate gradient method can be applied to find a minimum point of general real functions. Let us take Taylor series approximation to a function in quadratic form:

$$f(\mathbf{x}_0 + \Delta\mathbf{x}) = f(\mathbf{x}_0) + \frac{\partial f(\mathbf{x}_0)}{\partial x_i}\Delta x_i + \frac{1}{2}\frac{\partial^2 f(\mathbf{x}_0)}{\partial x_i \partial x_j}\Delta x_i \Delta x_j + \cdots \quad (11.33)$$

$$\approx c - \mathbf{b}\cdot\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}\cdot A\cdot\Delta\mathbf{x} \quad (11.34)$$

where $c = f(\mathbf{x}_0)$, $\mathbf{b} = -\nabla f(\mathbf{x}_0)$ and $A_{ij} = \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_i \partial x_j}$. Then the method finds the solution of $\nabla f(\mathbf{x}) = 0$ ($A\cdot\mathbf{x} = \mathbf{b}$) as a temporary minimum (maximum) point. Then approximates in quadratic form at that point and finds the solution again. This iteration will take us to the minimum (maximum) point. Note that this method cannot distinguish the minimum and maximum. We will talk about this later.

11.3.1. Minimization of quadratic functions

Let us consider that one finds a minimum point of quadratic functions of following form:

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}\cdot A\cdot\mathbf{x} - \mathbf{b}\cdot\mathbf{x} + c \quad (11.35)$$

where $\mathbf{x}$ and $\mathbf{b}$ are N dimensional vectors, c is a real constant and A is a $N\times N$ real symmetric, positive definite ($^{\forall}\mathbf{x} \neq 0, \quad \mathbf{x}\cdot A\cdot\mathbf{x} > 0$) matrix.

The function $f(x)$ is minimized by the solution of $A\cdot\mathbf{x} = \mathbf{b}$. This sentence can be proved as follow: For all $\mathbf{a}\in\mathbb{R}^N$, a function $f(\mathbf{x}+\mathbf{a})$ can be written as

$$f(\mathbf{x}+\mathbf{a}) = \frac{1}{2}(\mathbf{x}+\mathbf{a})\cdot A\cdot(\mathbf{x}+\mathbf{a}) - \mathbf{b}\cdot(\mathbf{x}+\mathbf{a}) + c$$

$$= f(\mathbf{x}) + \frac{1}{2}(\mathbf{a}\cdot A\cdot\mathbf{x} + \mathbf{x}\cdot A\cdot\mathbf{a}) - \mathbf{b}\cdot\mathbf{a} + \frac{1}{2}\mathbf{a}\cdot A\cdot\mathbf{a}. \quad (11.36)$$

Using the fact that

$$\mathbf{x}\cdot A\cdot\mathbf{a} = (\mathbf{x}\cdot A\cdot\mathbf{a})^T = \mathbf{a}\cdot A^T\cdot\mathbf{x} = \mathbf{a}\cdot A\cdot\mathbf{x}, \quad (11.37)$$

Eq. (11.36) can be simplified as

$$f(\mathbf{x}+\mathbf{a}) = f(\mathbf{x}) + \mathbf{a}\cdot A\cdot\mathbf{x} - \mathbf{b}\cdot\mathbf{a} + \frac{1}{2}\mathbf{a}\cdot A\cdot\mathbf{a}. \quad (11.38)$$

Because $A \cdot \mathbf{x} = \mathbf{b}$, Eq. (11.38) can be written by

$$f(\mathbf{x} + \mathbf{a}) = f(\mathbf{x}) + \frac{1}{2}\mathbf{a} \cdot A \cdot \mathbf{a}. \tag{11.39}$$

Since the matrix A is positive definite, $^{\forall}\mathbf{a} \neq 0$ increases $f(\mathbf{x})$. Furthermore, since $\nabla f(\mathbf{x}) = A \cdot \mathbf{x} - \mathbf{b}$, finding a solution of $A \cdot \mathbf{x} = \mathbf{b}$ is equal to finding a solution of $\nabla f(\mathbf{x}) = 0$.

11.3.2. Outline of minimization for general functions

At first the method approximates the function as a quadratic form at some point. Then finds a minimum point of the quadratic form. After that, it approximates the function as a quadratic form at the temporary minimum and finds a minimum point of new quadratic form again. The outline of this procedure is

$$\begin{aligned}
\mathbf{d}_{(0)} &= \mathbf{r}_{(0)} = \mathbf{b} - A \cdot \mathbf{x}_{(0)} = -\nabla f(\mathbf{x}_0) \\
\mathbf{x}_{(i+1)} &= \mathbf{x}_{(i)} + \alpha_{(i)}\mathbf{d}_{(i)} \\
\mathbf{r}_{(i+1)} &= -\nabla f(\mathbf{x}_{(i+1)}) \\
\mathbf{d}_{(i+1)} &= \mathbf{r}_{(i+1)} + \beta_{(i+1)}\mathbf{d}_{(i)}.
\end{aligned}$$

11.3.3. Calculation of $\beta_{(i+1)}$

A constant, $\beta_{(i+1)}$, is needed to find the solution of $A \cdot \mathbf{x} = \mathbf{b}$. For the constant, $\beta_{(i+1)}$ of conjugate gradient method Eq. (11.14). This is a Fletcher-Reeves formula [137]:

$$\beta^{FR}_{(i+1)} = \frac{\mathbf{r}_{(i+1)} \cdot \mathbf{r}_{(i+1)}}{\mathbf{r}_{(i)} \cdot \mathbf{r}_{(i)}}. \tag{11.40}$$

In general, however, the function is not quadratic and Fletcher-Reeves formula might make some problems. There is a improved version of $\beta_{(i+1)}$-finding method, Polak-Ribiere formula:

$$\beta^{PR}_{(i+1)} = \frac{\mathbf{r}_{(i+1)} \cdot (\mathbf{r}_{(i+1)} - \mathbf{r}_{(i)})}{\mathbf{r}_{(i)} \cdot \mathbf{r}_{(i)}}. \tag{11.41}$$

The additional term, $\mathbf{r}_{(i+1)} \cdot \mathbf{r}_{(i)}$, is zero for exact quadratic forms but is not zero in general functions. Because the quadratic form of a general function is an approximation, the method does not need to find an exact minimum point of quadratic form. Only the rough minimum of quadratic forms is good enough to find real minimum of original function. When it arrives at the rough minimum of the approximated quadratic form, $\mathbf{r}_{(i+1)}$ becomes $\mathbf{r}_{(i+1)} \simeq \mathbf{r}_{(i)}$. Therefore the additional term, $-\mathbf{r}_{(i)}$, makes $\beta^{PR}_{(i+1)} \simeq 0$ around the minimum of the quadratically approximated function, and restarts this procedure at the minimum point

by setting the $\mathbf{d}_{(i+1)}$ as a negative gradient direction at that minimum point.

When PR method is used, a problem cycling infinitely can rarely happen. The β should be positive for quadratic form but not for general functions, and the negative β causes this problem by making the direction vector $\mathbf{d}$ to backward. To solve this problem, let us define β as

$$\beta = \begin{cases} \beta^{PR} & \text{if} \quad \beta^{PR} > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{11.42}$$

This choice of β guarantees convergence.

11.3.4. Calculation of $\alpha_{(i)}$

In conjugate gradient method, the constant $\alpha_{(i)}$ is calculated by Eq. (11.12):

$$\alpha_{(i)} = \frac{\mathbf{r}_{(i)} \cdot \mathbf{r}_{(i)}}{\mathbf{d}_{(i)} \cdot A \cdot \mathbf{d}_{(i)}} = \frac{\mathbf{r}_{(i)} \cdot \mathbf{d}_{(i)}}{\mathbf{d}_{(i)} \cdot A \cdot \mathbf{d}_{(i)}}. \tag{11.43}$$

This result is equivalent to the result of finding $\alpha_{(i)}$ that minimizes $f(\mathbf{x}_{(i)} + \alpha_{(i)}\mathbf{d}_{(i)})$. The proof is follows:

For a quadratic function, $f(\mathbf{x} + \alpha\mathbf{d})$ can be expanded as

$$\begin{aligned} f(\mathbf{x} + \alpha\mathbf{d}) &= f(\mathbf{x}) + \alpha\left[\frac{d}{d\alpha} f(\mathbf{x} + \alpha\mathbf{d})\right]_{\alpha=0} + \frac{1}{2}\alpha^2\left[\frac{d^2}{d\alpha^2} f(\mathbf{x} + \alpha\mathbf{d})\right]_{\alpha=0} \\ &= f(\mathbf{x}) + \alpha\nabla f(\mathbf{x}) \cdot \mathbf{d} + \frac{1}{2}\alpha^2 \mathbf{d} \cdot A \cdot \mathbf{d} \end{aligned} \tag{11.44}$$

To find the minimum of $f(\mathbf{x}+\alpha\mathbf{d})$, let us find α that satisfies $\frac{d}{d\alpha} f(\mathbf{x}+\alpha\mathbf{d}) = 0$:

$$\begin{aligned} \frac{d}{d\alpha} f(\mathbf{x} + \alpha\mathbf{d}) &= \nabla f(\mathbf{x}) \cdot \mathbf{d} + \alpha \mathbf{d} \cdot A \cdot \mathbf{d} = 0 \\ \alpha &= -\frac{\nabla f(\mathbf{x}) \cdot \mathbf{d}}{\mathbf{d} \cdot A \cdot \mathbf{d}} = \frac{\mathbf{r} \cdot \mathbf{d}}{\mathbf{d} \cdot A \cdot \mathbf{d}} \end{aligned} \tag{11.45}$$

where we used $\mathbf{r} = -\nabla f(\mathbf{x})$.

For general functions, the method uses α that minimizes $f(\mathbf{x}+\alpha\mathbf{d})$. If the line minimizer does not use the second-derivative of the function, the CG function minimizer does not use the second-derivative of the function.

11.3.5. Limits

Conjugate gradient method for a minimum search algorithm has some limits. First, this method cannot find global minimum. Second, this method breaks down when it encounters saddle point. Third, this method cannot distinguish

minima from maxima. Therefore, the success of conjugate gradient method for a minimum search algorithm depends strongly on the initial point.

11.3.6. Practical implementation

It might be a good choice that stopping the iteration when $\mathbf{x}$ is close enough to the exact solution ($|\mathbf{r}_{(i)}| < \epsilon|\mathbf{r}_{(0)}|$) The iteration also can be stopped when the number of iterations is exceed some maximum number of iterations.

For quadratic forms, $\mathbf{d} \cdot \mathbf{r} = \mathbf{r} \cdot \mathbf{r}$ and the term $\mathbf{r} \cdot \mathbf{d}$ always is positive. If one restarts the iteration by setting $\mathbf{d} = \mathbf{r}$ when $\mathbf{r} \cdot \mathbf{d} \leq 0$, the convergence will be improved.

Algorithm 2 shows the skeleton code of CG function minimizer.

Algorithm 2 Skeleton code of CG function minimizer

```
1: k ← 0
2: r ← −∇f(x)
3: d ← r
4: δ_new ← r · r
5: δ_0 ← δ_new
6: for i = 1 to N_max do
7:    Find α that minimizes f(x + αd)
8:    x ← x + αd
9:    r_old ← r
10:   r ← −∇f(x)
11:   δ_old ← δ_new
12:   δ_mid ← r · r_old
13:   δ_new ← r · r
14:   β ← (δ_new − δ_mid) / δ_old
15:   d ← r + βd
16:   k ← k + 1
17:   if δ_new ≤ ε²δ_0 then
18:      break for loop
19:   end if
20:   if k = N_dim or r · d ≤ 0 then
21:      d ← r
22:      k ← 0
23:   end if
24: end for
```

11.4. Function minimization using Newton method

Newton method is a minimum finding method for a function using hessian. Like the CG method for minimization, Newton method repeatedly finds a minimum of quadratically approximated function.

11.4.1. Outline of Newton method

Let us suppose that the current position is $\mathbf{x}_i$ and we want the new position, $\mathbf{x}_{i+1}$, to be the minimum position. Locally, the function $f(\mathbf{x}_{i+1})$ can be approximated quadratically:

$$\begin{aligned} f(\mathbf{x}_i + \Delta\mathbf{x}) &= f(\mathbf{x}_i) + \frac{\partial f(\mathbf{x}_i)}{\partial x_a}\Delta x_a + \frac{1}{2}\frac{\partial^2 f(\mathbf{x}_i)}{\partial x_a \partial x_b}\Delta x_a \Delta x_b + \cdots \\ &\approx c - \mathbf{b}\cdot\Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}\cdot A\cdot\Delta\mathbf{x} \end{aligned} \tag{11.46}$$

where $c = f(\mathbf{x}_i)$, $\mathbf{b} = -\nabla f(\mathbf{x}_i)$ and $A_{ab} = \frac{\partial^2 f(\mathbf{x}_i)}{\partial x_a \partial x_b}$. If the hessian matrix A is positive definite, the approximated function, (11.46), has a minimum x_{i+1} that satisfies

$$\nabla f(\mathbf{x}_{i+1}) = A\cdot\Delta\mathbf{x} - \mathbf{b} = 0 \tag{11.47}$$

Therefore, the new position $\mathbf{x}_{i+1}$ can be found by

$$\begin{aligned} \mathbf{x}_{i+1} &= \mathbf{x}_i + \Delta\mathbf{x} \\ &= \mathbf{x}_i + A^{-1}\cdot\mathbf{b} \end{aligned} \tag{11.48}$$

In general, the function is not quadratic and the hessian matrix A might not be positive definite. In that case, the convergence can be improved by updating $\mathbf{x}_{i+1}$ as

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha A^{-1}\cdot\mathbf{b} \tag{11.49}$$

where α is determined to minimize $f(\mathbf{x}_i + \alpha A^{-1}\cdot\mathbf{b})$. This is the Newton method for minimization.

The properties of hessian matrix A does not guarantee the existence of its inverse matrix. Therefore the inversion must be done by special methods that can trace the eigenvalues, such as singular value decomposition[133]. Modified version of Cholesky factorization also a good choice, which makes the minimization method stable [138, 139].

Appendix A. Noether current

Let us consider a Lagrangian $\mathcal{L}(\Phi_i, \partial_\mu \Phi_i)$ that depends on n independent fields, Φ_i for $i = 1, 2, \cdots, n$. Local infinitesimal transformation of the fields is

$$\Phi_i(x) \quad \rightarrow \quad \Phi_i(x)' = \Phi_i(x) + \delta\Phi_i(x), \tag{A.1}$$

where $\delta\Phi_i(x)$ can be written in terms of infinitesimal parameters $\epsilon_a(x)$,

$$\delta\Phi_i(x) = \epsilon_a(x)\Delta\Phi_i^a(x), \tag{A.2}$$

with an index of the underlying symmetry group generators, a. The variation of the Lagrangian is

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\Phi_i}\delta\Phi_i + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi_i)}\partial_\mu\delta\Phi_i \tag{A.3}$$

$$= \epsilon_a(x)\left[\frac{\partial\mathcal{L}}{\partial\Phi_i}\Delta\Phi_i^a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi_i)}\partial_\mu\Delta\Phi_i^a\right] + \partial_\mu\epsilon_a(x)\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi_i)}\Delta\Phi_i^a\right]. \tag{A.4}$$

By defining current $J^{\mu,a}$ as

$$J^{\mu,a} \equiv \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi_i)}\Delta\Phi_i^a, \tag{A.5}$$

$\delta\mathcal{L}$ can be written in the following form:

$$\delta\mathcal{L} = \epsilon_a(x)\partial_\mu J^{\mu,a} + \partial_\mu\epsilon_a(x)J^{\mu,a}, \tag{A.6}$$

where we used the equation of motion,

$$\frac{\partial\mathcal{L}}{\partial\Phi_i} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi_i)} = 0. \tag{A.7}$$

to obtain

$$\partial_\mu J^{\mu,a} = \frac{\partial\mathcal{L}}{\partial\Phi_i}\Delta\Phi_i^a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi_i)}\partial_\mu\Delta\Phi_i^a. \tag{A.8}$$

From Eq. (A.6), the currents and divergence of the currents can be obtained by

$$J^{\mu,a} = \frac{\partial \delta \mathcal{L}}{\partial(\partial_\mu \epsilon_a)} \tag{A.9}$$

$$\partial_\mu J^{\mu,a} = \frac{\partial \delta \mathcal{L}}{\partial \epsilon_a}. \tag{A.10}$$

Appendix B. Gamma function

In this section, we summarize the gamma function, incomplete gamma functions and regularized gamma functions. Here we present integration formulae that hold for complex numbers z and s whose real part is positive. They can be extended to all complex numbers, except the negative integers and 0, by analytic continuation.

The gamma function is

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt. \tag{B.1}$$

The gamma function satisfies $\Gamma(z+1) = z\Gamma(z)$.

The upper incomplete gamma function is

$$\Gamma(s,x) = \int_x^\infty t^{s-1} e^{-t} dt, \tag{B.2}$$

and the lower incomplete gamma function is

$$\gamma(s,x) = \int_0^x t^{s-1} e^{-t} dt. \tag{B.3}$$

The regularized gamma functions are

$$P(s,x) = \frac{\gamma(s,x)}{\Gamma(s)}, \tag{B.4}$$

$$Q(s,x) = \frac{\Gamma(s,x)}{\Gamma(s)} = 1 - P(s,x). \tag{B.5}$$

Appendix C. A Derivation of the Probability Distribution Function of χ^2 distribution

Here we derive the probability distribution function of the χ^2 distribution. To make it easier, we start with one degrees of freedom case and two degrees of freedom case. Then we generalize it to the χ^2 distribution with k degrees of freedom.

C.1. χ^2 distribution with one degrees of freedom

Let us consider a random variable X that distributed according to the standard normal distribution, $X \sim \mathcal{N}(0,1)$. The cdf(cumulative distribution function) of X is

$$F_X(x) = P(X < x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt. \tag{C.1}$$

Note that it satisfies $1 - F_X(x) = F_X(-x)$ as the pdf(probability distribution function) of normal random variable is symmetric. Let us define a new random variable $Y \equiv X^2$. One can find the cdf of Y as follows:

$$\text{If } y < 0, \quad F_Y(y) = P(Y < y) = 0 \tag{C.2}$$

$$\begin{aligned} \text{If } y \geq 0, \quad F_Y(y) &= P(Y < y) = P(X^2 < y) = P(|x| < \sqrt{y}) \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) = F_X(\sqrt{y}) - \{1 - F_X(\sqrt{y})\} \\ &= 2F_X(\sqrt{y}) - 1. \end{aligned} \tag{C.3}$$

The pdf, $f_Y(y)$, for $y \geq 0$ can be derived from the cdf by taking derivatives as follows:

$$\begin{aligned} f_Y(y) = \frac{dF_Y(y)}{dy} &= 2 \cdot \frac{dF_X(\sqrt{y})}{dy} \\ &= 2 \cdot \frac{d}{dy} \left[\int_{\infty}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \right] \\ &= 2 \cdot \frac{1}{\sqrt{2\pi}} e^{\frac{y}{2}} \cdot \frac{d}{dy}\left(\sqrt{y}\right) \\ &= \frac{1}{2^{\frac{1}{2}}\Gamma(\frac{1}{2})} y^{\frac{1}{2}-1} e^{\frac{y}{2}} \quad \text{for } x \geq 0, \end{aligned} \tag{C.4}$$

where we used $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ and the following property

$$\frac{d}{dy}\int_a^{f(y)} g(x)dx = g\big(f(y)\big)\cdot\frac{df(y)}{dy} \tag{C.5}$$

for analytic functions $f(x), g(x)$ and for a constant a.

C.2. χ^2 distribution with two degrees of freedom

Let us consider a random variable Q defined as

$$Q \equiv X_1^2 + X_2^2, \tag{C.6}$$

where X_i are independent random variables distributed following the standard normal distribution. By definition, the Q is distributed according to the χ^2 distribution with degrees of freedom 2. For constants a and b, it satisfies

$$P(a \le Q < b) = P(a \le X_1^2 + X_2^2 < b). \tag{C.7}$$

If we define a variable q as $q \equiv x_1^2 + x_2^2$, the probability of Q satisfies

$$\begin{aligned} P(q \le Q < q + dq) &= f_Q(q)dq \\ &= \int_S p(x_1, x_2)dx_1dx_2, \end{aligned} \tag{C.8}$$

where $f_Q(q)$ is the pdf of Q and $p(x_1, x_2)$ is a joint pdf of X_1 and X_2, which is a product of two standard normal pdfs as they are mutually independent, and S denotes the surface of 2-dimensional ball that $q \le x_1^2 + x_2^2 < q + dq$. Here, the surface of 2-dimensional ball is a outline circle of a disk. The integral can be represented using the pdf of the standard normal pdf:

$$\begin{aligned} f_Q(q)dq &= \int_S \frac{1}{\sqrt{2\pi}^2} e^{-\frac{1}{2}(x_1^2+x_2^2)}dx_1dx_2 \\ &= \frac{1}{2\pi}e^{-\frac{1}{2}q}\int_S dx_1dx_2. \end{aligned} \tag{C.9}$$

Here, $\int_S dx_1dx_2$ can be performed as

$$\int_S dx_1dx_2 = 2\pi r\cdot dr = \pi dq, \tag{C.10}$$

where $r^2 \equiv q$ and $2rdr = dq$. As a result, one finds

$$f_Q(x) = \frac{1}{2}e^{-\frac{1}{2}x} \quad \text{for } x \ge 0. \tag{C.11}$$

C.3. χ^2 distribution with k degrees of freedom

Let us consider a random variable Q defined as

$$Q \equiv X_1^2 + X_2^2 + \cdots + X_k^2, \tag{C.12}$$

where X_i are independent random variables distributed following the standard normal distribution. By definition, the Q is distributed according to the χ^2 distribution with degrees of freedom k. If we define a variable q as $q \equiv x_1^2 + x_2^2 + \cdots + x_k^2$, the probability of Q satisfies

$$\begin{aligned} P(q \le Q < q + dq) &= f_Q(q)dq \\ &= \int_S p(x_1, x_2, \ldots, x_k) dx_1 dx_2 \cdots dx_k, \end{aligned} \tag{C.13}$$

where $f_Q(q)$ is the pdf of Q and $p(x_1, x_2, \ldots)$ is a joint pdf of X_i, which is a product of k standard normal pdfs as they are mutually independent, and S denotes the surface of k-dimensional ball that $q \le x_1^2 + x_2^2 < q + dq$. The integral can be represented using the pdf of the standard normal pdf:

$$\begin{aligned} f_Q(q)dq &= \int_S \frac{1}{\sqrt{2\pi}^k} e^{-\frac{1}{2}(x_1^2 + x_2^2 + \cdots + x_k^2)} dx_1 dx_2 \cdots dx_k \\ &= \frac{1}{\sqrt{2\pi}^k} e^{-\frac{1}{2}q} \int_S dx_1 dx_2, \end{aligned} \tag{C.14}$$

where we assumed that q is a constant in range of $q \le x_1^2 + x_2^2 < q + dq$ at the last equality. If we define r as $r^2 \equiv q$, S is a $(k-1)$ sphere with radius r. Since the area of $(k-1)$ sphere with radius r is

$$A = \frac{k \cdot r^{k-1} \pi^{\frac{\pi}{2}}}{\Gamma(\frac{k}{2} + 1)}, \tag{C.15}$$

$\int_S dx_1 dx_2$ can be performed easily whose consequence is

$$f_Q(x) = \frac{1}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})} x^{\frac{k}{2} - 1} e^{-\frac{1}{2}x} \quad \text{for } x \ge 0. \tag{C.16}$$

Since $P(Q < x) = 0$ for $x < 0$, the $f_Q(x) = 0$ if $x < 0$.

Appendix D. Error of Jackknife Estimation for Variance of Mean

The difference between the function value of the original samples, Eq. (8.15), and the jackknife estimate, Eq. (8.23), depends on the form of the function. If the function is the sample variance of mean, defined in Eq. (7.30), the difference becomes $\mathcal{O}(\frac{1}{N})$. The proof is following. The variance of mean of m^{th} set of jackknife samples is

$$C_{ii}^{J,m} = \frac{1}{(N-1)(N-2)} \sum_{i \neq m} \left(x_i - \overline{x}_m^J\right)^2, \tag{D.1}$$

where $\overline{x}_m^J$ is the average value of m^{th} set of jackknife samples,

$$\overline{x}_m^J = \frac{1}{N-1} \sum_{i \neq m} x_i = \frac{1}{N-1}(N\overline{x} - x_m). \tag{D.2}$$

Eq. (D.1) can be rewritten in terms of $\overline{x}$ instead of $\overline{x}_m^J$:

$$\begin{aligned} C_{ii}^{J,m} &= \frac{1}{(N-1)(N-2)} \sum_{i \neq m} \left[x_i - \frac{1}{N-1}(N\overline{x} - x_m) \right]^2 \\ &= \frac{1}{(N-1)(N-2)} \sum_{i \neq m} \left[x_i - \frac{1}{N-1}\Big((N-1)\overline{x} + \overline{x} - x_m \Big) \right]^2 \\ &= \frac{1}{(N-1)(N-2)} \sum_{i \neq m} \left[(x_i - \overline{x}) + \frac{1}{N-1}(x_m - \overline{x}) \right]^2 . \end{aligned} \tag{D.3}$$

Expanding the square, $(A + \frac{1}{N-1}B)^2 = A^2 + \frac{2}{N-1}AB + \mathcal{O}(1/N^2)$, it becomes

$$\begin{aligned} C_{ii}^{J,m} &= \frac{1}{(N-1)(N-2)} \sum_{i \neq m} (x_i - \overline{x})^2 \\ &\quad + \frac{2(x_m - \overline{x})}{(N-1)^2(N-2)} \sum_{i \neq m} (x_i - \overline{x}) + \mathcal{O}\left(\frac{1}{N^3}\right) \end{aligned} \tag{D.4}$$

The deleted sum can be split into two parts:

$$\sum_{i\neq m}(x_i-\overline{x}) = \sum_{i=1}^{N}(x_i-\overline{x})-(x_m-\overline{x}) = -(x_m-\overline{x}) \tag{D.5}$$

$$\sum_{i\neq m}(x_i-\overline{x})^2 = \sum_{i=1}^{N}(x_i-\overline{x})^2-(x_m-\overline{x})^2. \tag{D.6}$$

Armed with theses equations, Eq. (D.4) can be rewritten as

$$\begin{aligned} C_{ii}^{J,m} &= \frac{1}{(N-1)(N-2)}\sum_{i=1}^{N}(x_i-\overline{x})^2-\frac{1}{(N-1)(N-2)}(x_m-\overline{x})^2 \\ &\quad -\frac{2}{(N-1)^2(N-2)}(x_m-\overline{x})^2+\mathcal{O}\left(\frac{1}{N^3}\right) \\ &= \frac{1}{(N-1)(N-2)}\sum_{i=1}^{N}(x_i-\overline{x})^2 \\ &\quad -\frac{N+1}{(N-1)^2(N-2)}(x_m-\overline{x})^2+\mathcal{O}\left(\frac{1}{N^3}\right) \end{aligned} \tag{D.7}$$

Jackknife estimate of the variance of mean is average of $C_{ii}^{J,m}$ over m:

$$\begin{aligned} r_\star &= \frac{1}{N}\sum_{m=1}^{N}C_{ii}^{J,m} \\ &= \frac{1}{(N-1)(N-2)}\sum_{i=1}^{N}(x_i-\overline{x})^2 \\ &\quad -\frac{N+1}{N(N-1)^2(N-2)}\sum_{m=1}^{N}(x_m-\overline{x})^2+\mathcal{O}\left(\frac{1}{N^3}\right) \\ &= \frac{N}{N-2}\sigma_{\overline{x}}^2-\frac{N+1}{(N-1)(N-2)}\sigma_{\overline{x}}^2+\mathcal{O}\left(\frac{1}{N^3}\right) \\ &= \sigma_{\overline{x}}^2+\frac{1}{N}\sigma_{\overline{x}}^2+\mathcal{O}\left(\frac{1}{N^3}\right) \end{aligned} \tag{D.8}$$

$$\sigma_{\overline{x}}^2 = \frac{1}{N+1}\sum_{m=1}^{N}C_{ii}^{J,m}+\mathcal{O}\left(\frac{1}{N^3}\right) \tag{D.9}$$

Here, Taylor expansion was used to show the last equality in Eq. (D.8):

$$\frac{1}{N-a} = \frac{1}{N}\left(1+\frac{a}{N}\right)+\mathcal{O}\left(\frac{1}{N^3}\right). \tag{D.10}$$

Consequently, the error of jackknife estimate of variance of mean is $\mathcal{O}(1/N)$ of $\sigma_{\overline{x}}^2$. Since $\sigma_{\overline{x}}^2$ itself is $\mathcal{O}(1/N)$, in other words, the difference should defined by

$$\text{difference} \equiv 2 \left| \frac{r_\star - \sigma_{\overline{x}}^2}{r_\star + \sigma_{\overline{x}}^2} \right|, \tag{D.11}$$

and the difference is $\mathcal{O}(1/N)$.

Bibliography

[1] O.W. Greenberg, "Spin and Unitary Spin Independence in a Paraquark Model of Baryons and Mesons", *Phys.Rev.Lett.* **13** (1964) 598–602.

[2] M.Y. Han and Yoichiro Nambu, "Three Triplet Model with Double SU(3) Symmetry", *Phys.Rev.* **139** (1965) B1006–B1010.

[3] Sidney R. Coleman and David J. Gross, "Price of asymptotic freedom", *Phys.Rev.Lett.* **31** (1973) 851–854.

[4] H. David Politzer, "Reliable Perturbative Results for Strong Interactions?", *Phys.Rev.Lett.* **30** (1973) 1346–1349.

[5] D.J. Gross and Frank Wilczek, "Asymptotically Free Gauge Theories. 1", *Phys.Rev.* **D8** (1973) 3633–3652.

[6] A. Bazavov, D. Toussaint, C. Bernard, J. Laiho, C. DeTar, et al., "Nonperturbative QCD simulations with 2+1 flavors of improved staggered quarks", *Rev.Mod.Phys.* **82** (2010) 1349–1417, `arXiv:0903.3598`.

[7] Claude W. Bernard, Tom Burch, Kostas Orginos, Doug Toussaint, Thomas A. DeGrand, et al., "The Static quark potential in three flavor QCD", *Phys.Rev.* **D62** (2000) 034503, `arXiv:hep-lat/0002028`.

[8] **JLQCD collaboration**, H. Fukaya et al., "Determination of the chiral condensate from 2+1-flavor lattice QCD", *Phys.Rev.Lett.* **104** (2010) 122002, `arXiv:0911.5555`.

[9] **JLQCD and TWQCD collaborations**, H. Fukaya et al., "Determination of the chiral condensate from QCD Dirac spectrum on the lattice", *Phys.Rev.* **D83** (2011) 074501, `arXiv:1012.4052`.

[10] Thomas DeGrand, Zhaofeng Liu, and Stefan Schaefer, "Quark condensate in two-flavor QCD", *Phys.Rev.* **D74** (2006) 094504, `arXiv:hep-lat/0608019`.

[11] Kenneth G. Wilson, "Confinement of Quarks", *Phys.Rev.* **D10** (1974) 2445–2459.

[12] Holger Bech Nielsen and M. Ninomiya, "No Go Theorem for Regularizing Chiral Fermions", *Phys.Lett.* **B105** (1981) 219.

[13] Steven A. Gottlieb, W. Liu, D. Toussaint, R.L. Renken, and R.L. Sugar, "Hybrid Molecular Dynamics Algorithms for the Numerical Simulation of Quantum Chromodynamics", *Phys.Rev.* **D35** (1987) 2531–2542.

[14] S. Duane, A.D. Kennedy, B.J. Pendleton, and D. Roweth, "Hybrid Monte Carlo", *Phys.Lett.* **B195** (1987) 216–222.

[15] David J.E. Callaway and Aneesur Rahman, "THE MICROCANONICAL ENSEMBLE: A NEW FORMULATION OF LATTICE GAUGE THEORY", *Phys.Rev.Lett.* **49** (1982) 613.

[16] David J.E. Callaway and Aneesur Rahman, "LATTICE GAUGE THEORY IN MICROCANONICAL ENSEMBLE", *Phys.Rev.* **D28** (1983) 1506.

[17] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, "Equation of state calculations by fast computing machines", *J.Chem.Phys.* **21** (1953) 1087–1092.

[18] Taegil Bae, Yong-Chull Jang, Hyung-Jin Kim, Jangho Kim, Jongjeong Kim, Kwangwoo Kim, Boram Yoon, Weonjong Lee, Chulwoo Jung, and Stephen R. Sharpe, "First results for B_K on the ultrafine ($a = 0.045$ fm) ensemble", *PoS* **LATTICE2010** (2010) 296, `arXiv:1010.4781`.

[19] **MILC Collaboration**, A. Bazavov et al., "Properties of light pseudoscalars from lattice QCD with HISQ ensembles", *PoS* **LATTICE2011** (2011) 107, `arXiv:1111.4314`.

[20] **Fermilab Lattice Collaboration, MILC Collaboration**, A. Bazavov et al., "Pseudoscalar meson physics with four dynamical quarks", `arXiv:1210.8431`.

[21] S. Basak, S. Datta, M. Padmanath, P. Majumdar, and N. Mathur, "Charm and strange hadron spectra from overlap fermions on HISQ gauge configurations", `arXiv:1211.6277`.

[22] Federico Farchioni, Gregorio Herdoiza, Karl Jansen, Andreas Nube, Marcus Petschlies, et al., "Pseudoscalar decay constants from $N_f = 2 + 1 + 1$ twisted mass lattice QCD", *PoS* **LATTICE2010** (2010) 128, `arXiv:1012.0200`.

[23] R. Baron, Ph. Boucaud, J. Carbonell, A. Deuzeman, V. Drach, et al., "Light hadrons from lattice QCD with light (u,d), strange and charm dynamical quarks", *JHEP* **1006** (2010) 111, `arXiv:1004.5284`.

[24] Krzysztof Cichy, Vincent Drach, Elena Garcia Ramos, Karl Jansen, Chris Michael, et al., "Properties of pseudoscalar flavour-singlet mesons from 2+1+1 twisted mass lattice QCD", `arXiv:1211.4497`.

[25] C. Alexandrou, M. Constantinou, S. Dinter, V. Drach, K. Hadjiyiannakou, et al., "Sigma terms and strangeness content of the nucleon with $Nf = 2 + 1 + 1$ twisted mass fermions", arXiv:1211.4447.

[26] **PACS-CS Collaboration**, S. Aoki et al., "Physical Point Simulation in 2+1 Flavor Lattice QCD", *Phys.Rev.* **D81** (2010) 074503, arXiv:0911.2561.

[27] S. Durr, Z. Fodor, C. Hoelbling, S.D. Katz, S. Krieg, et al., "Lattice QCD at the physical point: Simulation and analysis details", *JHEP* **1108** (2011) 148, arXiv:1011.2711.

[28] Gilberto Colangelo, Stephan Durr, Andreas Juttner, Laurent Lellouch, Heinrich Leutwyler, et al., "Review of lattice results concerning low energy particle physics", *Eur.Phys.J.* **C71** (2011) 1695, arXiv:1011.4408.

[29] **PACS-CS collaboration**, S. Aoki et al., "Non-perturbative renormalization of quark mass in $N_f = 2 + 1$ QCD with the Schroedinger functional scheme", *JHEP* **1008** (2010) 101, arXiv:1006.1164.

[30] A. Bazavov, C. Bernard, C. DeTar, X. Du, W. Freeman, et al., "Staggered chiral perturbation theory in the two-flavor case and SU(2) analysis of the MILC data", *PoS* **LATTICE2010** (2010) 083, arXiv:1011.1792.

[31] C. McNeile, C.T.H. Davies, E. Follana, K. Hornbostel, and G.P. Lepage, "High-Precision c and b Masses, and QCD Coupling from Current-Current Correlators in Lattice and Continuum QCD", *Phys.Rev.* **D82** (2010) 034512, arXiv:1004.4285.

[32] S. Durr, Z. Fodor, C. Hoelbling, S.D. Katz, S. Krieg, et al., "Lattice QCD at the physical point: light quark masses", *Phys.Lett.* **B701** (2011) 265–268, arXiv:1011.2403.

[33] **RBC Collaboration, UKQCD Collaboration**, Y. Aoki et al., "Continuum Limit Physics from 2+1 Flavor Domain Wall QCD", *Phys.Rev.* **D83** (2011) 074508, arXiv:1011.0892.

[34] T. Blum, R. Zhou, T. Doi, M. Hayakawa, T. Izubuchi, et al., "Electromagnetic mass splittings of the low lying hadrons and quark masses from 2+1 flavor lattice QCD+QED", *Phys.Rev.* **D82** (2010) 094508, arXiv:1006.1311.

[35] C.T.H. Davies, C. McNeile, K.Y. Wong, E. Follana, R. Horgan, et al., "Precise Charm to Strange Mass Ratio and Light Quark Masses from Full Lattice QCD", *Phys.Rev.Lett.* **104** (2010) 132003, arXiv:0910.3102.

[36] **MILC Collaboration**, A. Bazavov et al., "MILC results for light pseudoscalars", *PoS* **CD09** (2009) 007, arXiv:0910.2966.

[37] **RBC-UKQCD Collaboration**, C. Allton et al., "Physical Results from 2+1 Flavor Domain Wall QCD and SU(2) Chiral Perturbation Theory", *Phys.Rev.* **D78** (2008) 114509, arXiv:0804.0473.

[38] **RBC-UKQCD Collaboration**, P.A. Boyle et al., "K $\to \pi$ form factors with reduced model dependence", *Eur.Phys.J.* **C69** (2010) 159–167, arXiv:1004.0886.

[39] P.A. Boyle, A. Juttner, R.D. Kenway, C.T. Sachrajda, S. Sasaki, et al., "K(l3) semileptonic form-factor from 2+1 flavour lattice QCD", *Phys.Rev.Lett.* **100** (2008) 141601, arXiv:0710.5136.

[40] **ETM Collaboration**, V. Lubicz, F. Mescia, L. Orifici, S. Simula, and C. Tarantino, "Improved analysis of the scalar and vector form factors of kaon semileptonic decays with $N_f = 2$ twisted-mass fermions", *PoS* **LATTICE2010** (2010) 316, arXiv:1012.3573.

[41] **ETM Collaboration**, V. Lubicz, F. Mescia, S. Simula, and C. Tarantino, "K $\to$ pi l nu Semileptonic Form Factors from Two-Flavor Lattice QCD", *Phys.Rev.* **D80** (2009) 111502, arXiv:0906.4728.

[42] S. Durr, Z. Fodor, C. Hoelbling, S.D. Katz, S. Krieg, et al., "The ratio FK/Fpi in QCD", *Phys.Rev.* **D81** (2010) 054507, arXiv:1001.4692.

[43] **HPQCD Collaboration, UKQCD Collaboration**, E. Follana, C.T.H. Davies, G.P. Lepage, and J. Shigemitsu, "High Precision determination of the pi, K, D and D(s) decay constants from lattice QCD", *Phys.Rev.Lett.* **100** (2008) 062002, arXiv:0706.1726.

[44] Taegil Bae, Yong-Chull Jang, Chulwoo Jung, Hyung-Jin Kim, Jangho Kim, Jongjeong Kim, Kwangwoo Kim, Sunghee Kim, Weonjong Lee, Stephen R. Sharpe, and Boram Yoon, "Kaon B-parameter from improved staggered fermions in $N_f = 2 + 1$ QCD", *Phys.Rev.Lett.* **109** (2012) 041601, arXiv:1111.5698.

[45] Y. Aoki, R. Arthur, T. Blum, P.A. Boyle, D. Brommel, et al., "Continuum Limit of B_K from 2+1 Flavor Domain Wall QCD", *Phys.Rev.* **D84** (2011) 014503, arXiv:1012.4178.

[46] Taegil Bae, Yong-Chull Jang, Chulwoo Jung, Hyung-Jin Kim, Jangho Kim, Jongjeong Kim, Kwangwoo Kim, Weonjong Lee, Stephen R. Sharpe, and Boram Yoon, "B_K using HYP-smeared staggered fermions in $N_f = 2 + 1$ unquenched QCD", *Phys.Rev.* **D82** (2010) 114509, arXiv:1008.5179.

[47] C. Aubin, Jack Laiho, and Ruth S. Van de Water, "The Neutral kaon mixing parameter B(K) from unquenched mixed-action lattice QCD", *Phys.Rev.* **D81** (2010) 014507, arXiv:0905.3947.

[48] **RBC Collaboration, UKQCD Collaboration**, D.J. Antonio et al., "Neutral kaon mixing from 2+1 flavor domain wall QCD", *Phys.Rev.Lett.* **100** (2008) 032001, arXiv:hep-ph/0702042.

[49] C. Aubin and C. Bernard, "Pseudoscalar decay constants in staggered chiral perturbation theory", *Phys.Rev.* **D68** (2003) 074011, arXiv:hep-lat/0306026.

[50] **Particle Data Group**, J. Beringer et al., "Review of particle physics", *Phys.Rev.D* **86** (2012) 010001.

[51] **MILC Collaboration**, A. Bazavov et al., "Results from the MILC collaboration's SU(3) chiral perturbation theory analysis", *PoS* **LAT2009** (2009) 079, arXiv:0910.3618.

[52] Yong-Chull Jang and Weonjong Lee, "Current Status of Indirect CP Violation in Neutral Kaon System", *PoS* **LATTICE2012** (2012) 269, arXiv:1211.0792.

[53] Yong-Chull Jang, Chulwoo Jung, Weonjong Lee, and Boram Yoon, "Covariance fitting of highly correlated data in lattice QCD", arXiv:1101.2248.

[54] **SWME Collaboration**, Boram Yoon, Yong-Chull Jang, Weonjong Lee, and Chulwoo Jung, "Trouble shooting for covariance fitting in highly correlated data", *PoS* **LATTICE2011** (2011) 296, arXiv:1111.0119.

[55] K. Symanzik, "Continuum Limit and Improved Action in Lattice Theories. 1. Principles and phi**4 Theory", *Nucl.Phys.* **B226** (1983) 187.

[56] K. Symanzik, "Continuum Limit and Improved Action in Lattice Theories. 2. O(N) Nonlinear Sigma Model in Perturbation Theory", *Nucl.Phys.* **B226** (1983) 205.

[57] M. Luscher and P. Weisz, "On-Shell Improved Lattice Gauge Theories", *Commun.Math.Phys.* **97** (1985) 59.

[58] John B. Kogut and Leonard Susskind, "Hamiltonian Formulation of Wilson's Lattice Gauge Theories", *Phys.Rev.* **D11** (1975) 395.

[59] S. Naik, "A mean field theory study of lattice gauge theory with finite temperature and with finite fermion density", *Nucl. Phys.* **B334** (1990) 611–636.

[60] **MILC Collaboration**, Kostas Orginos, Doug Toussaint, and R.L. Sugar, "Variants of fattening and flavor symmetry restoration", *Phys.Rev.* **D60** (1999) 054503, arXiv:hep-lat/9903032.

[61] G. Peter Lepage, "Flavor symmetry restoration and Symanzik improvement for staggered quarks", *Phys.Rev.* **D59** (1999) 074502, arXiv:hep-lat/9809157.

[62] G. Peter Lepage and Paul B. Mackenzie, "On the viability of lattice perturbation theory", *Phys. Rev.* **D48** (1993) 2250–2264, hep-lat/9209022.

[63] Anna Hasenfratz and Francesco Knechtli, "Flavor symmetry and the static potential with hypercubic blocking", *Phys.Rev.* **D64** (2001) 034504, arXiv:hep-lat/0103029.

[64] Taegil Bae, David H. Adams, Chulwoo Jung, Hyung-Jin Kim, Jongjeong Kim, et al., "Taste symmetry breaking with HYP-smeared staggered fermions", *Phys.Rev.* **D77** (2008) 094508, arXiv:0801.3000.

[65] **HPQCD Collaboration, UKQCD Collaboration**, E. Follana et al., "Highly improved staggered quarks on the lattice, with applications to charm physics", *Phys.Rev.* **D75** (2007) 054502, arXiv:hep-lat/0610092.

[66] Stephen L. Adler, "Axial vector vertex in spinor electrodynamics", *Phys.Rev.* **177** (1969) 2426–2438.

[67] Stephen L. Adler and William A. Bardeen, "Absence of higher order corrections in the anomalous axial vector divergence equation", *Phys.Rev.* **182** (1969) 1517–1536.

[68] William A. Bardeen, "Anomalous Ward identities in spinor field theories", *Phys.Rev.* **184** (1969) 1848–1857.

[69] J. Goldstone, "Field Theories with Superconductor Solutions", *Nuovo Cim.* **19** (1961) 154–164.

[70] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg, "Broken Symmetries", *Phys.Rev.* **127** (1962) 965–970.

[71] Stefan Scherer, "Introduction to chiral perturbation theory", *Adv.Nucl.Phys.* **27** (2003) 277, arXiv:hep-ph/0210398, To be edited by J.W. Negele and E. Vogt.

[72] H. Georgi, "Weak Interactions and Modern Particle Theory", 1984.

[73] J. Gasser and H. Leutwyler, "Chiral Perturbation Theory: Expansions in the Mass of the Strange Quark", *Nucl.Phys.* **B250** (1985) 465.

[74] **HPQCD Collaboration, MILC Collaboration, UKQCD Collaboration**, C. Aubin et al., "First determination of the strange and light quark masses from full lattice QCD", *Phys.Rev.* **D70** (2004) 031504, arXiv:hep-lat/0405022.

[75] **MILC Collaboration**, C. Aubin et al., "Light pseudoscalar decay constants, quark masses, and low energy constants from three-flavor lattice QCD", *Phys.Rev.* **D70** (2004) 114501, arXiv:hep-lat/0407028.

[76] **Fermilab Lattice Collaboration, MILC Collaboration, HPQCD Collaboration**, C. Aubin et al., "Semileptonic decays of D mesons in three-flavor lattice QCD", *Phys.Rev.Lett.* **94** (2005) 011601, arXiv:hep-ph/0408306.

[77] C. Aubin, C. Bernard, Carleton E. DeTar, M. Di Pierro, Elizabeth Dawn Freeland, et al., "Charmed meson decay constants in three-flavor lattice QCD", *Phys.Rev.Lett.* **95** (2005) 122002, arXiv:hep-lat/0506030.

[78] **HPQCD Collaboration**, Alan Gray et al., "The B meson decay constant from unquenched lattice QCD", *Phys.Rev.Lett.* **95** (2005) 212001, arXiv:hep-lat/0507015.

[79] Christopher Aubin and Tom Blum, "Lowest order hadronic contribution to the muon g-2", *PoS* **LAT2005** (2006) 089, arXiv:hep-lat/0509064.

[80] Masataka Okamoto, "Full determination of the CKM matrix using recent results from lattice QCD", *PoS* **LAT2005** (2006) 013, arXiv:hep-lat/0510113.

[81] Emel Dalgic, Alan Gray, Matthew Wingate, Christine T.H. Davies, G.Peter Lepage, et al., "B meson semileptonic form-factors from unquenched lattice QCD", *Phys.Rev.* **D73** (2006) 074502, arXiv:hep-lat/0601021.

[82] C. Aubin and T. Blum, "Calculating the hadronic vacuum polarization and leading hadronic contribution to the muon anomalous magnetic moment with improved staggered quarks", *Phys.Rev.* **D75** (2007) 114502, arXiv:hep-lat/0608011.

[83] C. Bernard, Carleton E. DeTar, M. Di Pierro, A.X. El-Khadra, R.T. Evans, et al., "The $\bar{B} \to D^* \ell \bar{\nu}$ form factor at zero recoil from three-flavor lattice QCD: A Model independent determination of $|V_{cb}|$", *Phys.Rev.* **D79** (2009) 014506, arXiv:0808.2519.

[84] Jon A. Bailey, C. Bernard, Carleton E. DeTar, M. Di Pierro, A.X. El-Khadra, et al., "The $B \to \pi \ell \nu$ semileptonic form factor from

three-flavor lattice QCD: A Model-independent determination of $|V_{ub}|$", *Phys.Rev.* **D79** (2009) 054507, arXiv:0811.3640.

[85] **HPQCD Collaboration**, Elvira Gamiz, Christine T.H. Davies, G.Peter Lepage, Junko Shigemitsu, and Matthew Wingate, "Neutral B Meson Mixing in Unquenched Lattice QCD", *Phys.Rev.* **D80** (2009) 014503, arXiv:0902.1815.

[86] Jangho Kim, Chulwoo Jung, Hyung-Jin Kim, Weonjong Lee, and Stephen R. Sharpe, "Finite volume effects in B_K with improved staggered fermions", *Phys.Rev.* **D83** (2011) 117501, arXiv:1101.2685.

[87] **Fermilab Lattice and MILC Collaborations**, A. Bazavov et al., "B- and D-meson decay constants from three-flavor lattice QCD", *Phys.Rev.* **D85** (2012) 114506, arXiv:1112.3051.

[88] Jon A. Bailey, A. Bazavov, C. Bernard, C.M. Bouchard, C. DeTar, et al., "$B_s \to D_s/B \to D$ Semileptonic Form-Factor Ratios and Their Application to BR($B_s^0 \to \mu^+\mu^-$)", *Phys.Rev.* **D85** (2012) 114502, arXiv:1202.6346.

[89] A. Bazavov, C. Bernard, C.M. Bouchard, C. DeTar, M. Di Pierro, et al., "Neutral B-meson mixing from three-flavor lattice QCD: Determination of the SU(3)-breaking ratio ξ", *Phys.Rev.* **D86** (2012) 034503, arXiv:1205.7013.

[90] Weon-Jong Lee and Stephen R. Sharpe, "Partial flavor symmetry restoration for chiral staggered fermions", *Phys.Rev.* **D60** (1999) 114503, arXiv:hep-lat/9905023.

[91] C. Aubin and C. Bernard, "Pion and kaon masses in staggered chiral perturbation theory", *Phys.Rev.* **D68** (2003) 034014, arXiv:hep-lat/0304014.

[92] Stephen R. Sharpe and Ruth S. Van de Water, "Staggered chiral perturbation theory at next-to-leading order", *Phys. Rev.* **D71** (2005) 114505, hep-lat/0409018.

[93] Jon A. Bailey, Boram Yoon, and Weonjong Lee, "Taste non-Goldstone pion decay constants in staggered chiral perturbation theory", arXiv:1211.1106.

[94] **SWME Collaboration**, Jon A. Bailey, Hyung-Jin Kim, and Weonjong Lee, "Taste non-Goldstone, flavor-charged pseudo-Goldstone boson masses in staggered chiral perturbation theory", *Phys.Rev.* **D85** (2012) 094503, arXiv:1112.2108.

[95] **MILC Collaboration**, C. Bernard, "Chiral logs in the presence of staggered flavor symmetry breaking", *Phys.Rev.* **D65** (2002) 054031, arXiv:hep-lat/0111051.

[96] Claude W. Bernard and Maarten F.L. Golterman, "Partially quenched gauge theories and an application to staggered fermions", *Phys.Rev.* **D49** (1994) 486–494, arXiv:hep-lat/9306005.

[97] P.H. Damgaard and K. Splittorff, "Partially quenched chiral perturbation theory and the replica method", *Phys.Rev.* **D62** (2000) 054509, arXiv:hep-lat/0003017.

[98] Claude Bernard, Maarten Golterman, and Yigal Shamir, "Effective field theories for QCD with rooted staggered fermions", *Phys.Rev.* **D77** (2008) 074505, arXiv:0712.2560.

[99] Stephen R. Sharpe and Noam Shoresh, "Partially quenched chiral perturbation theory without Phi0", *Phys.Rev.* **D64** (2001) 114510, arXiv:hep-lat/0108003.

[100] Claude W. Bernard and Maarten F.L. Golterman, "Chiral perturbation theory for the quenched approximation of QCD", *Phys.Rev.* **D46** (1992) 853–857, arXiv:hep-lat/9204007.

[101] T. Inami and C.S. Lim, "Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes k(L) $\to$ mu anti-mu, K+ $\to$ pi+ Neutrino anti-neutrino and K0 $\leftrightarrow$ anti-K0", *Prog.Theor.Phys.* **65** (1981) 297.

[102] Jon A. Bailey, Hyung-Jin Kim, Weonjong Lee, and Stephen R. Sharpe, "Kaon mixing matrix elements from beyond-the-Standard-Model operators in staggered chiral perturbation theory", *Phys.Rev.* **D85** (2012) 074507, arXiv:1202.1570.

[103] Marco Ciuchini, V. Lubicz, L. Conti, A. Vladikas, A. Donini, et al., "Delta M(K) and epsilon(K) in SUSY at the next-to-leading order", *JHEP* **9810** (1998) 008, arXiv:hep-ph/9808328.

[104] Andrzej J. Buras, Mikolaj Misiak, and Jorg Urban, "Two loop QCD anomalous dimensions of flavor changing four quark operators within and beyond the standard model", *Nucl.Phys.* **B586** (2000) 397–426, arXiv:hep-ph/0005183.

[105] F. Gabbiani, E. Gabrielli, A. Masiero, and L. Silvestrini, "A Complete analysis of FCNC and CP constraints in general SUSY extensions of the standard model", *Nucl.Phys.* **B477** (1996) 321–352, arXiv:hep-ph/9604387.

[106] **RBC and UKQCD Collaborations**, P.A. Boyle, N. Garron, and R.J. Hudspith, "Neutral kaon mixing beyond the standard model with $n_f = 2 + 1$ chiral fermions", *Phys.Rev.* **D86** (2012) 054028, arXiv:1206.5737.

[107] V. Bertone, N. Carrasco, M. Ciuchini, P. Dimopoulos, R. Frezzotti, et al., "Kaon Mixing Beyond the SM from Nf=2 tmQCD and model independent constraints from the UTA", arXiv:1207.1287.

[108] Weon-jong Lee and Stephen R. Sharpe, "One loop matching coefficients for improved staggered bilinears", *Phys.Rev.* **D66** (2002) 114501, arXiv:hep-lat/0208018.

[109] **HPQCD Collaboration**, C.T.H. Davies, E. Follana, I.D. Kendall, G. Peter Lepage, and C. McNeile, "Precise determination of the lattice spacing in full lattice QCD", *Phys.Rev.* **D81** (2010) 034506, arXiv:0910.1229.

[110] Jongjeong Kim, Weonjong Lee, and Stephen R. Sharpe, "One-loop matching of improved four-fermion staggered operators with an improved gluon action", *Phys.Rev.* **D83** (2011) 094503, arXiv:1102.1774.

[111] Jon A. Bailey, Taegil Bae, Yong-Chull Jang, Hwancheol Jeong, Chulwoo Jung, Hyung-Jin Kim, Jangho Kim, Jongjeong Kim, Kwangwoo Kim, Seonghee Kim, Weonjong Lee, Jaehoon Leem, Stephen R. Sharpe, and Boram Yoon, "Beyond the Standard Model corrections to $K^0 - \bar{K}^0$ mixing", arXiv:1211.1101.

[112] **SWME Collaboration**, Kwangwoo Kim et al., "B_K from improved staggered fermions using SU(3) chiral perturbation theory", *PoS* **LATTICE2011** (2011) 313, arXiv:1110.2575.

[113] Boram Yoon, Taegil Bae, Yong-Chull Jang, Hyung-Jin Kim, Jangho Kim, et al., "B_K with improved staggered fermions: analysis using SU(2) staggered chiral perturbation theory", *PoS* **LATTICE2010** (2010) 319, arXiv:1010.4778.

[114] Jon A. Bailey, Hyung-Jin Kim, Weonjong Lee, and Stephen R. Sharpe, "Chiral extrapolation of matrix elements of BSM kaon operators", arXiv:1210.7754.

[115] G. P. Lepage et al., "Constrained curve fitting", *Nucl. Phys. Proc. Suppl.* **106** (2002) 12–20, hep-lat/0110175.

[116] David H. Adams and Weonjong Lee, "Renormalization group evolution for the DeltaS=1 effective Hamiltonian with N(f)=2+1", *Phys.Rev.* **D75** (2007) 074502, arXiv:hep-lat/0701014.

[117] Sheldon Ross, "A first course in probability", Prentice Hall, eighth ed., 2009.

[118] Colin Morningstar, "The Monte Carlo method in quantum field theory", `hep-lat/0702020`.

[119] Mark J. Schervish, "Theory of Statistics", Springer-Verlag, 1995.

[120] B. Efron, "Bootstrap methods: Another look at the jackknife", *Ann. Statist.* **7** (1979) 1–26.

[121] M.H. Quenouille, "Notes on bias in estimation", *Biometrika* **43** (1956) 353–360.

[122] J.W. Tukey, "Bias and confidence in not-quite large samples (abstract)", *The Annals of Mathematical Statistics* **29** (1958) 614.

[123] Theodore W. Anderson, "An Introduction to Multivariate Statistical Analysis", Wiley Interscience, third ed., 2003.

[124] G. Peter Lepage, "A New Algorithm for Adaptive Multidimensional Integration", *J. Comput. Phys.* **27** (1978) 192.

[125] (ed.) DeGrand, Thomas A. and (ed.) Toussaint, D., "From actions to answers. Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics, Boulder, USA, June 5-30, 1989", 1990.

[126] D.S. Sivia and J. Skilling, "Data Analaysis — A Bayesian Tutorial", Oxford University Press, second ed., 2006.

[127] Christopher Michael, "Fitting correlated data", *Phys.Rev.* **D49** (1994) 2616–2619, `arXiv:hep-lat/9310026`.

[128] D. Toussaint and W. Freeman, "Sample size effects in multivariate fitting of correlated data", `arXiv:0808.2211`.

[129] **Fermilab Lattice and MILC Collaborations**, Jon A. Bailey et al., "Semileptonic decays of K and D mesons in 2+1 flavor QCD", *PoS* **LATTICE2010** (2010) 306, `arXiv:1011.2423`.

[130] **Fermilab Lattice and MILC Collaborations**, Jon A. Bailey et al., "$B \to Dl\nu$ at zero recoil: an update", *PoS* **LATTICE2010** (2010) 311, `arXiv:1011.2166`.

[131] Tanmoy Bhattacharya and et al., "Non-perturbative Renormalization Constants using Ward Identities", *Phys. Lett. B* **461** (1999) 79, `arXiv:hep-lat/9904011`.

[132] William Press, Saul Teukosky, William Vetterling, and Brian Flannary, "Numerical recipes", Cambridge University Press, New York, third ed., 2007. chapter 7, section 4.

[133] William Press et al., "Numerical recipes", Cambridge University Press, third ed., 2007.

[134] J. A. Nelder and R. Mead, "A simplex method for function minimization", *The Computer Journal* **7** (1965) 308–313.

[135] Magunus R. Hestenes and Eduard Stiefel, "Methods of conjugate gradients for solving linear systems", *J. Res. Natl. Bur. Stand.* **49** December (1952) 409 – 436.

[136] H. A. Van der Vorst, "Bi-CGSTAB: A Fast and Smoothly Converging Variant of Bi-CG for the Solution of Nonsymmetric Linear Systems", *SIAM J. Sci. and Statist. Comput.* **13** (1965) 631–644.

[137] R. Fletcher and C. Reeves, "Function minimization by conjugate gradient", *J. Comput.* **7** (1964), no. 2, 149 – 154.

[138] Walter Murray Philip Gill and Margaret Wrightt, "Practical optimization", Academic Press, 1981.

[139] Jorge Nocedal and Stephen Wright, "Numerical optimization", Springer, New York, 1999.

국문초록

제 1 편에서는 스태거드 카이랄 섭동 이론을 이용하여 taste non-Goldstone 파이온의 붕괴 상수를 버금 차수까지 계산하였다. 이 계산은 taste Goldstone 경우를 일반화 한 것이다. 해석적 보정항에만 새로운 낮은 에너지 결합 상수들이 존재하며, 카이랄 로그항에는 새로운 결합 상수들이 나타나지 않았다. 우리는 quenched, fully dynamical 그리고 partially quenched 경우에 대하여 카이랄 SU(3) 와 SU(3) 이론의 결과들을 기술하였다. 본 연구 결과는 기존에 존재하는 파이온 붕괴 상수와 낮은 에너지 결합 상수들을 더욱 정교하게 계산하는데 사용될 수 있다.

제 2 편에서는 $N_f = 2+1$ 의 MILC asqtad 격자 위에서 HYP-smearing 을 이용하여 향상된 스태거드 페르미온을 이용하여 초 표준모형의 B 파라미터를 계산하였다. 세 개의 서로 다른 격자 간격을 ($a \approx 0.045,\ 0.06,\ 0.09$ fm) 이용하여 연속 공간에서의 값을 추정하였다. 연산자 맞춤은 한 개의 고리 차수에서 섭동적으로 계산된 결과를 이용하였으며, 결과는 $\overline{\text{MS}}$ 체계의 2 GeV 와 3 GeV 로 주행시켰다. 카이랄 외삽과 연속 극한으로의 외삽에는 SU(2) 스태거드 카이랄 섭동 이론을 이용하였다. 본 논문에는 통계적 오차만을 갖고 있는 예비 결과를 보고하였다.

제 3 편에서는 기초 확률 이론, 오차 분석 기술 및 최소 χ^2 맞춤 등의 자료 분석 방법에 대하여 자세하게 설명하였다. 또한, 대각 근사 방법, 특이값 분해 방법, 베이지안 방법을 이용하여 강한 상관 관계를 갖고 있는 데이터를 분석하는 방법을 설명하였다. 그리고 공분산 행렬을 수정하지 않고도 공분산 맞춤을 할 수 있는 새로운 방법인 eigenmode shift 방법을 제시하였다.

주요어: 양자색역학, 격자 QCD, 카이랄 섭동 이론, 파이온 붕괴 상수, 초 표준모형 B 파라미터

학번 : 2006-20339