

HADRON STRUCTURE FROM LATTICE QUANTUM CHROMODYNAMICS USING TWISTED MASS FERMIONS

A Dissertation
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

by
Colin J. Lauer
May 2021

Examining Committee Members:

Martha Constantinou, Department of Physics, Advisory Chair

Andreas Metz, Department of Physics

Nikolaos Sparveris, Department of Physics

Krzysztof Cichy, External Reader, Adam Mickiewicz University

ABSTRACT

Hadron structure is an important field in particle physics because hadrons make up most of the matter in nature. The theory of the strong nuclear force, via which the partons of hadrons interact, is Quantum Chromodynamics (QCD) and cannot be solved analytically. Lattice QCD (LQCD) is an ideal formulation of QCD and is the only formulation starting from first principles. In this thesis, we use LQCD for two primary topics of study: 1) nucleon structure and 2) pion and kaon structure.

In the first study, we calculate the quark momentum fraction, helicity, and transversity for the nucleon. The calculations are performed on three ensembles at the physical point of the pion mass allowing us to study finite volume, discretization, strange and charm quark quenching, and excited-state systematic effects. Our calculations of the helicity and transversity are first predictions at the physical point.

In the second study, we investigate pion and kaon structure. We calculate the first three non-trivial Mellin moments of the meson parton distribution functions (PDFs). For the kaon, this is the first direct calculation of the second and third moments. We carefully choose which matrix elements we implement so that there is no mixing with lower derivative operators, avoiding systematic uncertainties which are not well understood. We also perform an extensive study of the excited-state contamination.

In a pioneering study, we show that the full x -dependence of the PDFs can be calculated from the first three Mellin moments. Such a calculation was previously thought to be unfeasible using moments calculated from LQCD. Our reconstruction of the PDFs allow us to comment on SU(3) flavor symmetry breaking and the high- x behavior of the pion PDF which are both interesting topics in hadron structure.

ACKNOWLEDGEMENTS

I would like to express my great appreciation to my advisor, Martha Constantinou, for her instruction and guidance throughout my research. I would also like to thank my collaborators, Constantia Alexandrou, Simone Bacchio, Ian Cloët, Kyriacos Hadjiyiannakou, and Giannis Koutsou, for their work on this project and for many fruitful discussions.

TABLE OF CONTENTS

ABSTRACT	i
ACKNOWLEDGEMENTS	ii
LIST OF FIGURES	xi
LIST OF TABLES	xiv
1 INTRODUCTION	1
1.1 Quantum Chromodynamics	2
1.2 Parton Distribution Functions	5
2 LATTICE QCD	8
2.1 The Feynman Path Integral	8
2.2 Discretized Fields	10
2.2.1 Grassmann Algebra	13
2.2.2 Quark Sources	18
2.2.3 Inversion Algorithms	20
2.3 Lattice Actions	23
2.3.1 Wilson Action	24
2.3.2 Twisted Mass Fermion Action	26
2.4 Gauge Field Configurations	29
2.4.1 Markov Chains	29
2.4.2 Dynamical Fermions	32
2.4.3 Gauge Smearing	38
2.5 Systematic uncertainties	38

3	CALCULATING HADRON STRUCTURE QUANTITIES ON THE LATTICE	41
3.1	Matrix Elements	41
3.2	Renormalization	45
3.3	Controlling Excited State Systematic Uncertainties	49
3.4	Decomposition	50
3.5	Jackknife Resampling	53
4	NUCLEON STRUCTURE	55
4.1	Calculation Details	55
4.2	Systematic Uncertainties	56
4.3	Comparison With Other Calculations	61
5	MESON STRUCTURE	64
5.1	Calculation Details	65
5.2	Effective Mass	68
5.3	Mellin Moments of PDFs	70
5.3.1	First Mellin Moment $\langle \mathbf{x} \rangle$	72
5.3.2	The Second Mellin Moment $\langle \mathbf{x}^2 \rangle$	75
5.3.3	The Third Mellin Moment $\langle \mathbf{x}^3 \rangle$	78
5.3.4	Comparison Between Moments	80
5.3.5	Comparison with Other Studies	82
5.4	Reconstruction of PDFs	89
5.5	Form Factors	97
6	CONCLUSION	102
	BIBLIOGRAPHY	104
	A DECOMPOSITION EQUATIONS	117
	APPENDIX	117
A.1	Nucleon Rest Frame	117
A.2	Meson General frame	118
A.2.1	General momentum transfer	118
A.2.2	Zero momentum transfer	119

A.3	Meson Rest frame	119
A.3.1	General momentum transfer	119
A.3.2	Zero momentum transfer	119

LIST OF FIGURES

1.1	The kinematics in the (x, Q^2) for which there are available hadronic cross-section data for global fits of the unpolarized, helicity, and transversity PDFs [1].	6
2.1	A graphical representation of a forward (left) and backward (right) oriented gauge link $U_\mu(n)$ and $U_{-\mu}(n)$	12
2.2	A graphical representation of a plaquette $U_{\mu\nu}$	13
3.1	Three-point function diagrams for a meson (left) and nucleon (right). The source is at space-time position 0, the current is inserted at $x = (t, \mathbf{x})$, and the sink is at $x_s = (T_{\text{sink}}, \mathbf{x}_s)$. Lines with forward arrows represent the quark propagators and lines with backward facing arrows represent anti-quark propagators.	42
3.2	Chirally extrapolated results for Z_{vD} (top), Z_{vDD} (middle), and Z_{vDDD} (bottom). Blue triangles correspond to RI' scheme, black circles to $\overline{\text{MS}}$ scheme and magenta diamonds to the subtracted results in the $\overline{\text{MS}}$ scheme. The data are plotted as a function of the initial RI' scale $(a\mu_0)^2$. The dashed line corresponds to the fit of Eq. (3.16), and a filled magenta diamond represent our final value for \mathcal{Z}_{vD} , \mathcal{Z}_{vDD} , and $\mathcal{Z}_{\text{vDDD}}$	48

- 4.1 Ensemble cB211.072.64 results for $\langle x \rangle_{u-d}$ (top row), $\langle x \rangle_{\Delta u-\Delta d}$ (middle row), $\langle x \rangle_{\delta u-\delta d}$ (bottom row). In the left column, we plot the ratios in Eq. (3.22) multiplied by the renormalization and kinematic factors according to Eqs. (A.1-A.4) for $T_{\text{sink}}/a = 8, 10, 12, 14, 16, 18, 20$ as blue circles, orange squares, green diamonds, red downward-pointing triangles, purple upward-pointing triangles, brown left-pointing triangles, and magenta right-pointing triangles, respectively. The blue, orange, and green curves are functions calculated from the two-state fit parameters for each value of T_{sink} . In the center panel, the plateau fits for each T_{sink} are plotted and the gray curve is the function obtained from the two-state fit parameters using $t_i = T_{\text{sink}}/2$ band. In the right panel, the black squares are the two-state fit values using the ratio from Eq. (3.21) plotted as function of the lowest T_{sink} used in the fit. The open square is the value selected as the result and the gray bands show its error range across all three panels. The green triangles are calculated using the summation method, which is explained in [2] . . . 58
- 4.2 Ensemble cA2.09.64 results with notation similar to Figure 4.1. The different source-separations, $T_{\text{sink}}/a = 14, 16, 18$, are shown as blue squares for, orange squares, and green diamonds, respectively. 59
- 4.3 Ensemble cA2.09.48 results with notation similar to 4.1. The different source-separations, $T_{\text{sink}}/a = 10, 12, 14, 16, 18$, are shown as blue squares for, orange squares, green diamonds, red downward-pointing triangles, and purple upward-pointing triangles respectively. 60
- 4.4 A comparison of $\langle x \rangle_{u-d}^N$ calculated by different groups and methods. Our results are shown as a red star(cB211.072.64), a green diamond (cA2.09.64), and blue square (cA2.09.48). The results from the lattice calculation of the the LHPC group [3] is plotted as an orange triangle. The open and uncolored shapes are phenomenological calculations of global fits [4, 5, 6, 7, 8, 9]. 61

4.5 A comparison of $\langle x \rangle_{u-d}^N$ calculated by different groups and methods. Our results are shown by a green triangle and its errors by a green band. The other lattice calculations are plotted as an orange triangle for the LHPC group and a blue square and red star for two other ETMC ensembles. The open and uncolored shapes are phenomenological calculations of global fits. 62

4.6 A comparison of $\langle x \rangle_{u-d}^N$ calculated by different groups and methods. Our results are shown by a green triangle and its errors by a green band. The other lattice calculations are plotted as an orange triangle for the LHPC group and a blue square and red star for two other ETMC ensembles. The open and uncolored shapes are phenomenological calculations of global fits. 63

5.1 Pion (top) and kaon (bottom) mass in the rest frame as a function of the the lowest value of t/a , t_{low}/a entering the fit. The red squares are the results calculated the plateau method and the green circles are those calculated from the two-state fit. We also show the values we choose as our final results of the plateau and two-state fits as a purple square and blue circle, respectively. 69

5.2 Pion (top) and kaon (bottom) mass in the boosted frame as a function of the the lowest value of t/a entering the fit. The notation is the same as in Fig. 5.1. 70

5.3 Pion (left) and kaon (right) m_{eff} in the rest frame. The fitted value from the plateau m_{plat} is shown with a red band, and from the two-state fit applied on m_{eff} with a green band. 71

5.4 Plots of the meson $\langle x \rangle$ results in the rest frame. The three rows are for the up part of the pion (top) and the up (center) and strange (bottom) parts of the kaon. In the left column, the ratios corresponding to $\langle x \rangle$ calculated according to Eq. A.19 are plotted for $T_{\text{sink}}/a = 12, 14, 16, 18, 20, 24$ as blue circles, red squares, green downward-pointing triangles, magenta left-pointing triangles, cyan right-pointing triangles, and orange downward-pointing triangles, respectively. In the right column, the plateau fits are plotted as a function of T_{sink} . The gray bands in the right column are calculated from the two-state parameters as a function of T_{sink} and at a constant $t = T_{\text{sink}}/2$. The purple bands going across all plots are the results from the two-state fit. 73

5.5 $\langle x \rangle$ results in the $\mathbf{p}^2 = \frac{12\pi^2}{L^2}$ boosted frame plotted with notation similar to the left column of Figure 5.4. The different source-separations, $T_{\text{sink}}/a = 14, 16, 18$, are shown as blue circles, red squares, and green triangles, respectively. 74

5.6 Comparison of $\langle x \rangle$ in the boosted (filled symbols) and rest frame (open symbols). From top to bottom we show $\langle x \rangle$ for the pion and kaon up and strange contribution. Results at $T_{\text{sink}}/a = 14, 16, 18$ are shown in the left, center and right panels, respectively. 76

5.7 $\langle x^2 \rangle$ plotted with notation similar to Figure 5.4. The different source-separations, $T_{\text{sink}}/a = 12, 14, 16, 18$, are shown as blue circles, red squares, green upward-pointing triangles, and magenta left-pointing triangles, respectively. 77

5.8 $\langle x^2 \rangle$ plotted with notation similar to Figure 5.4. The different source-separations, $T_{\text{sink}}/a = 12, 14, 16, 18$, are shown as blue circles, red squares, green upward-pointing triangles, and magenta left-pointing triangles, respectively. 79

5.9	Top: The x dependence of $xq_\pi^u(x)$ for the 2-parameter (blue band) and 3-parameter (pink band) fits. Bottom: Same as top panel for $xq_K^u(x)$ (left) and $xq_K^s(x)$ (right). Results are given in the $\overline{\text{MS}}$ scheme at 27 GeV ²	91
5.10	Dependence of $xq_\pi^u(x)$ (top), $xq_K^u(x)$ (lower left panel) and $xq_K^s(x)$ (lower right panel) on the source-sink time separation. The pink, green, yellow and purple bands correspond to $t_s/a = 14, 16, 18$ and the 2-state fit, respectively. The results are given in the $\overline{\text{MS}}$ scheme at a scale of 27 GeV ²	92
5.11	Top: The x dependence of $xq_\pi^u(x)$ using the 2-parameter with $\langle x^{n_{\text{max}}} \rangle = \langle x^2 \rangle, \langle x^3 \rangle, \langle x^4 \rangle$, shown with blue, pink and green bands. For the green band we use a constraint of $\langle x^4 \rangle_\pi^u$ [10]. Bottom: Same as top panel for $xq_K^u(x)$ (left panel) and $xq_K^s(x)$ (right panel) using the BLFQ-NJL [11] $\langle x^4 \rangle_K$ values as constraints for the pink band.	93
5.12	Pion PDF using the JAM data (blue band) and the reconstructed PDF using its moments with $n \leq 3$ (pink band). The reported scale is 27 GeV ²	94
5.13	Comparison of $xq_\pi^u(x)$ (blue band), $xq_K^u(x)$ (pink band) and $xq_K^s(x)$ (green band) at 27 GeV ² . The reconstruction uses our lattice data up to $\langle x^3 \rangle$ obtained with the 2-state fits analysis and a 2-parameter fit.	94

5.14 Top left panel: Comparison of $xq_{\pi}^u(x)$ with other lattice calculations, experimental data and global analysis, all in the $\overline{\text{MS}}$ at 27 GeV². Our results (blue band) use data up to $\langle x^3 \rangle$ obtained with the 2-state fits analysis and a 2-parameter fit. The E615 data [12] are plotted as gray points and the rescaled ASV data [13] as a solid cyan curve. The JAM global fit is shown with a red band, and the lattice results from pseudo-ITD [14] and current current correlators (LCS) [15] are shown with orange and green band, respectively. Top right panel: Comparison of our results for $xq_{\pi}^u(x)$ with DSE [16] (dotted orange curve), the updated DSE'18 [17] (dot-dashed red curve), BLFQ-NJL [11] (dashed green curve), and χCQ [18] (solid purple curve). Bottom panel: Same as top panel for $xq_K^u(x)$ (left) and $xq_K^s(x)$ (right). 96

5.15 Preliminary results of the scalar form factors for the up part pion (top), up part of the kaon (bottom left), and strange part of the kaon (bottom right) plotted as a function of Q^2 . The different source-sink separations are plotted as the colored symbols as in Figure 5.4. 98

5.16 Preliminary results of the vector form factors for the pion (top) and kaon (bottom) in the rest frame. In addition to the form factors at the various values of T_{sink} , the values calculated from the two-state fit is also plotted as purple stars. 99

5.17 Comparison of preliminary results of the vector form factors for the pion (top) and kaon (bottom) between the rest frame (open symbols) and boosted frame (closed symbols). 99

5.18 Preliminary results of the tensor form factors for the pion (top) and kaon (bottom) in the rest frame. 100

5.19 Comparison of preliminary results of the tensor form factors for the pion (top) and kaon (bottom) between the rest frame (open symbols) and boosted frame (closed symbols). 101

LIST OF TABLES

3.1	Interpolating fields for the particles under study. The fields u , d , and s are the spinors for up, down, and strange quarks. The transposition T in the nucleon interpolator acts on the Dirac indices and C is charge conjugation operator.	42
3.2	The form factors under study, along with their corresponding charge or moment, insertion currents, and insertion current gamma structures. The notation $\{\dots\}$ and $[\dots]$ means symmetrization over enclosed indices and traceless.	43
3.3	Parameters for the $N_f = 4$ ensembles used for the renormalization functions.	46
3.4	Pion mass dependence of the renormalization function Z_{vD} . The first column is the pion mass (in lattice units) for the ensemble, the second (third) is the renormalization function at scale $(a\mu_0)^2=2$ ($(a\mu_0)^2=4$). The number in the parenthesis is the statistical error.	47
3.5	Table similar to Table (3.4) but for Z_{vDD}	47

4.1	Parameters of the ensembles used for the nucleon $\langle x \rangle$ calculation. The process of determining the lattice spacing from the nucleon mass is outlined in [19]. The systematic errors in the lattice spacing of cA2.09.48 and cA2.09.64 comes from an interpolation to the physical pion mass using one-loop chiral perturbation theory because the pion mass was underestimated [19]. The errors in m_π are due to the values being calculated from the lattice spacing.	56
4.2	Statistics used for each nucleon correlator. For cA2.09.48, an asterisk is used to label which source-sink separations were calculated using only the Γ_0 projector.	57
4.3	Final results for the three nucleon moments and three ensembles. In all cases, the results of the two-state fit are chosen as the final results.	60
5.1	Parameters of the ensemble used to calculate meson quantities under study.	66
5.2	Statistics used for each meson correlator.	67
5.3	Final results of the ground state energies and the corresponding value of t_{low} used in the fit. Note that for $\mathbf{p}^2 = 0$, $E_0 = m_{\text{eff}}$	71
5.4	Renormalized data for $\langle x \rangle$ for various T_{sink} values and the 2-state fit ((a) $T_{\text{sink}} \in [12 - 24]$, (b) $T_{\text{sink}} \in [14 - 18]$). The numbers shown in the parenthesis are the statistical errors.	72
5.5	Plateau fit results for $\langle x \rangle$ in the $\mathbf{p}^2 = \frac{12\pi^2}{L^2}$ boosted frame for the three T_{sink} values and the 2-state fit using $T_{\text{sink}} \in [14 - 18]$). The number shown in the parenthesis are the statistical errors.	75
5.6	Renormalized data for $\langle x^2 \rangle$ at each value of T_{sink} and the two-state fit ($T_{\text{sink}} \in [12 - 18]$). The number shown in the parenthesis is statistical error.	78

5.7	Renormalized data for $\langle x^3 \rangle$ at each value of T_{sink} and the two-state fit ($T_{\text{sink}} \in [12 - 18]$). The number shown in the parenthesis is statistical error.	79
5.8	Comparison of lattice results and phenomenological data for $\langle x \rangle^\pi$ and $\langle x^2 \rangle^\pi$	84
5.9	Comparison of lattice results for $\langle x^3 \rangle_\pi^u$ in the $\overline{\text{MS}}$ scheme at 2 GeV. The evolution from the reported scale (“initial” scale) to 2 GeV is applied to NNLO. Statistical and systematic uncertainties have been added in quadrature where applicable.	86
5.10	Comparison of lattice results for $\langle x^3 \rangle_K^u$ and $\langle x^3 \rangle_K^s$ in the $\overline{\text{MS}}$ scheme at 2 GeV. The evolution from the reported scale (“initial” scale) to 2 GeV is applied to NNLO. Statistical and systematic uncertainties have been added in quadrature where applicable.	87
5.11	Comparison of $\langle x^3 \rangle_\pi$ with global fits and model calculations in the $\overline{\text{MS}}$ scheme at 2 GeV. The evolution from the reported scale (“initial” scale) to 2 GeV is applied to NNLO. Statistical and systematic uncertainties have been added in quadrature where applicable.	88
5.12	Comparison of $\langle x^3 \rangle_K^u$ and $\langle x^3 \rangle_K^s$ with global fits and model calculations in the $\overline{\text{MS}}$ scheme at 2 GeV. The evolution from the reported scale (“initial” scale) to 2 GeV is applied to NNLO. Statistical and systematic uncertainties have been added in quadrature where applicable.	89
5.13	The values for the fit parameters, α , β and γ for q_π^u , q_K^u and q_K^s at 5.2 GeV. The error in the parenthesis is statistical.	90
5.14	Values for the first six moments for the pion and kaon. The number shown in the first parenthesis is statistical error, while the number in the second parenthesis is systematic due to excited-states contamination.	95

CHAPTER 1

INTRODUCTION

Hadrons make up most of the visible matter in the universe so understanding their structure and how they behave is fundamentally important in Physics. The strong nuclear interactions between hadrons and their constituent particles, quarks and gluons, are described by the theory of Quantum Chromodynamics (QCD). QCD is a part of the Standard Model of particle physics, which describes how particles interact via three out of the four known fundamental forces, with gravity being the only force not included in the theory.

Despite the importance of studying hadron structure, there are still open questions in the field that are not fully understood, such as the origin of mass and quark distributions in hadrons. The nucleons, pion, and kaons are among the most important hadrons. Studying nucleon, pion, and kaon structure provides valuable information for the experimental program of the future Electron-Ion-Collider [20, 21] as well as for existing high-energy experiments, such as deep-inelastic lepton scattering and Drell-Yan in hadron-hadron collisions at facilities such as Jefferson Lab, RHIC, Fermilab, and the LHC. In this thesis, we study their structure via Mellin moments of: a) parton distribution functions (PDFs) and b) generalized parton distributions (GPDs).

We study the properties of the nucleon, pion, and kaon using the lattice formulation of Quantum Chromodynamics (LQCD) for our calculations of hadron structure quantities. LQCD is the only formulation which solves the QCD Lagrangian non-perturbatively from first principles. In LQCD, space-time is discretized in a Euclidean lattice and large-scale numerical simulations are used to calculate physical quantities. The input parameters of LQCD are the quark masses and coupling constant which are the same as in the continuum theory.

This thesis is organized as follows: in Chapter 1, we cover the basics of QCD and the hadron structure quantities parton distribution functions. In Chapter 2, we introduce the theory of LQCD and its fundamentals. In Chapter 3, we explain how to calculate the hadron structure quantities under study from LQCD. We present and discuss our results for the nucleon in Chapter 4, and for the mesons in Chapter 5.

1.1 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the theory which describes the interaction between fundamental particles via the strong nuclear force. The particles which interact strongly are the quarks, anti-quarks, and gluons. An interesting property of the strong nuclear force that makes it unique from other fundamental forces is the running of the strong coupling. This leads to asymptotic freedom, that is, at short distances or high energies, the strong coupling decreases and the quarks behave as though they are free and perturbation theory can be used to solve the QCD equations. At long distances or low energies, the strong coupling becomes high and quarks are strongly bound in the hadron. This phenomenon, known as confinement, is the reason that quarks are always in a hadronic bound state with at least one other quark or anti-quark.

QCD describes the strong force in terms of the partonic degrees of freedom, quarks and gluons [22]. Before writing the QCD Lagrangian, let us give the definition of its various ingredients. Quarks are fermions and are represented by color-spinors

$$\psi_f = (\psi_\mu^a(x))_f, \quad (1.1)$$

where $a = 1, 2, 3$ is the color index, μ is the Dirac index, and $f = u, d, s, c, t, b$ is the flavor index. Gluons are the gauge bosons of the theory and are represented by elements of the symmetry $SU(3)$, 3×3 tensors which are unitary and traceless. They can be written in terms of the eight $SU(3)$ generators t^a

$$A_\mu = A_\mu^a t^a. \quad (1.2)$$

The product of a quark and anti-quark $\bar{\psi}\psi$ is invariant under the local SU(3) transformation

$$\psi \rightarrow \psi' = \Omega(x)\psi, \quad \bar{\psi} \rightarrow \bar{\psi}' = \Omega^\dagger(x)\bar{\psi} \quad (1.3)$$

$$\Omega = \exp\left(i\alpha^a(x)\frac{t^a}{2}\right), \quad (1.4)$$

where $\alpha^a(x)$ is a continuous function of the space-time position. It is not clear how the derivative acting on ψ , defined by

$$\partial_\mu\psi = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} [\psi(x + \epsilon_\mu) - \psi(x)], \quad (1.5)$$

where ϵ_μ is a vector with length ϵ in the direction μ , transforms according to Eq. (1.4) because the two terms are at different space-time positions. So we introduce the gluon fields in a function called the connector which transforms as

$$U(y, x) \rightarrow U'(y, x) = \Omega(y)U(y, x)\Omega^\dagger(x), \quad (1.6)$$

and then define the covariant derivative as

$$D_\mu\psi = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} [\psi(x + \epsilon_\mu) - U(x + \epsilon_\mu)\psi(x)]. \quad (1.7)$$

From the infinitesimal form of the connector

$$U(x + \epsilon_\mu, x) = 1 - ig\epsilon_\mu A_\mu^a(x)\frac{t^a}{2} + \mathcal{O}(\epsilon^2), \quad (1.8)$$

where g is the strong coupling constant, we see that the covariant derivative can be written in terms of the gauge fields as

$$D_\mu = \partial_\mu - igA_\mu^a\frac{t^a}{2}. \quad (1.9)$$

We want to see how the gauge fields transform under SU(3) rotations so we write the transformation of the connector

$$1 - ig\epsilon_\mu A_\mu^a(x)\frac{t^a}{2} \rightarrow \Omega(x + \epsilon_\mu) \left(1 - ig\epsilon_\mu A_\mu^a(x)\frac{t^a}{2}\right) \Omega^\dagger(x), \quad (1.10)$$

and expand the product

$$\begin{aligned}
\Omega(x + \epsilon_\mu)\Omega^\dagger(x) &= [(1 + \epsilon^\mu \partial_\mu + \mathcal{O}(\epsilon^2))\Omega(x)]\Omega^\dagger(x) \\
&= 1 + \epsilon^\mu (\partial_\mu \Omega(x))\Omega^\dagger(x) + \mathcal{O}(\epsilon^2) \\
&= 1 + \epsilon^\mu \Omega(x)(-\partial_\mu \Omega^\dagger(x)) + \mathcal{O}(\epsilon^2).
\end{aligned} \tag{1.11}$$

Now we can see that the gauge fields transform as

$$\begin{aligned}
A_\mu^a(x) \frac{t^a}{2} &\rightarrow \Omega(x) \left(A_\mu^a(x) \frac{t^a}{2} + \frac{i}{g} \partial_\mu \right) \Omega^\dagger(x) \\
&= \left(1 + i\alpha^a(x) \frac{t^a}{2} + \mathcal{O}(\alpha^2) \right) \left(A_\mu^a(x) \frac{t^a}{2} + \frac{i}{g} \partial_\mu \right) \left(1 + i\alpha^a(x) \frac{t^a}{2} - \mathcal{O}(\alpha^2) \right) \\
&= A_\mu^a(x) \frac{t^a}{2} + \frac{1}{g} (\partial_\mu \alpha^a(x)) \frac{t^a}{2} + i \left[\alpha^a(x) \frac{t^a}{2}, A_\mu^b(x) \frac{t^b}{2} \right].
\end{aligned} \tag{1.12}$$

From here, we can also write down the transformation of the covariant derivative acting on the quark fields as

$$\begin{aligned}
D_\mu \psi &\rightarrow \left(\partial_\mu - ig A_\mu^a(x) \frac{t^a}{2} - i(\partial \alpha^a) \frac{t^a}{2} + g \left[\alpha^a \frac{t^a}{2}, A_\mu^b \frac{t^b}{2} \right] \right) \left(1 + i\alpha^c \frac{t^c}{2} \right) \psi \\
&= \left(\partial_\mu - ig A_\mu^a(x) \frac{t^a}{2} \right) \left(1 + i\alpha^c \frac{t^c}{2} \right) \psi + \mathcal{O}(\alpha^2) \\
&= \left(1 + i\alpha^a \frac{t^a}{2} \right) D_\mu \psi \\
&= \Omega(x) D_\mu \psi.
\end{aligned} \tag{1.13}$$

and see that $D_\mu \psi$ transforms the same as ψ . Lastly, we define the field strength tensor as

$$F_{\mu\nu}^a = -i[D_\mu, D_\nu] = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g f^{abc} A_\mu^b A_\nu^c, \tag{1.14}$$

where f^{abc} are the anti-symmetric structure constants of SU(3).

Now we have all of the ingredients to construct the gauge invariant QCD Lagrangian

$$\mathcal{L}_{QCD} = \sum_{N_f} \bar{\psi}_f (i \not{D}_\mu - m_f) \psi_f - \frac{1}{2} F_{\mu\nu}^a F_a^{\mu\nu}. \tag{1.15}$$

Like the fermions in Quantum Electrodynamics (QED), the quarks interact with the

gauge bosons through the covariant derivative. Conversely, unlike the force carriers in QED, the gluons experience self-interactions. The action associated with this Lagrangian can be separated into a fermionic part and a gauge part given by

$$S[\psi, \bar{\psi}, A] = S_F[\psi, \bar{\psi}, A] + S_G[A], \quad (1.16)$$

which will be useful later in our lattice calculations. The fermionic part can be written as

$$S_F[\psi, \bar{\psi}, A] = \sum_{N_f} \bar{\psi}_f D_f \psi_f, \quad (1.17)$$

where D is called the Dirac matrix. The QCD Lagrangian cannot be solved analytically because it is highly non-linear and there are infinite degrees of freedom. As we will see later, the Dirac matrix is important for defining the action in LQCD and its numerical calculations, as its discretization is not unique.

1.2 Parton Distribution Functions

Parton distribution functions (PDFs) are important quantities in hadron structure. They are the probability densities of observing a parton with longitudinal momentum fraction x of the total hadron momentum and are defined in the light-cone frame where the hadron has momentum p with components $p^\pm = (p^0 \pm p^3)/\sqrt{2}$, $p^\perp = 0$ as

$$q(x) = \frac{1}{4\pi} \int dy^- e^{-iy^- xp^+} \langle p | \bar{\psi}(0, y^-, \mathbf{0}_\perp \gamma^+ \mathcal{G} \psi(0, 0, \mathbf{0}) | p \rangle, \quad (1.18)$$

$$q = g, u, \bar{u}, d, \bar{d}, s, \bar{s}, \dots, \quad (1.19)$$

where \mathcal{G} is a gauge link which connects the quark field ψ to infinity and is required for Eq. (1.18) to be gauge invariant. For hadrons with spin, such as nucleons, an unpolarized (helicity averaged) and polarized (helicity dependent) PDF can both be defined as

$$q(x) \equiv q^{\rightarrow}(x) + q^{\leftarrow}(x), \quad (1.20)$$

$$\Delta q(x) \equiv q^{\rightarrow}(x) - q^{\leftarrow}(x), \quad (1.21)$$

where \rightarrow (\leftarrow) indicates a sum over parton spins along (opposite) the hadron helicity.

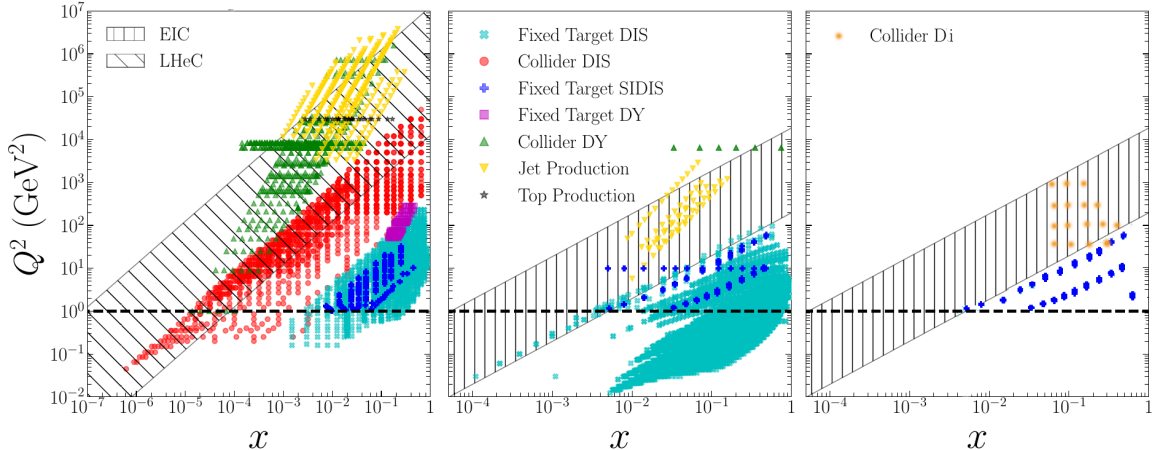


Figure 1.1: The kinematics in the (x, Q^2) for which there are available hadronic cross-section data for global fits of the unpolarized, helicity, and transversity PDFs [1].

The transverse PDF δq can also be defined which is the probability density of finding quarks with with an equal or opposite polarization as a transversely polarized hadron.

There are two main ways of calculating PDFs nonperturbatively [23]: global QCD analysis and LQCD. In global QCD analysis [24, 25, 26], factorization is used to separate experimental hard-scattering cross-sections into short- and long-range dynamics. The short-range or high energy dynamics depend on the process involved and can be calculated using perturbation theory while the long-range or low energy dynamics are the PDFs and are universal. The factorization scale which separates the two is called μ and depends on the experimental process. The information gained from perturbative QCD and global fits of experimental data are used to calculate PDF. High precision experiments in major facilities such as Jefferson Lab, HERA, and RHIC as well as advancements in perturbative QCD have led to global analysis calculations of PDFs with precisions of a few percent. The most well known nucleon PDF is the unpolarized, since it has the most available experimental data. The helicity only has a few hundred data sets [1, 27] and the transversity PDF is even less well known [28]. Fig. 1.1 shows the kinematics that the hadronic cross-section data which are included in global fits of each PDF have been experimentally measured.

Since the quark fields in PDFs are at light-like separations and since, as will be discussed later, LQCD uses a Euclidean space-time, the PDFs can not be calculated directly on the lattice. There are a number methods to calculate the x dependence of PDFs on the lattice but these are still under development. In this study we instead

study PDFs as has been done most often up to now in LQCD: by calculating the Mellin moments of PDFs

$$\langle x^n \rangle = \int_{-1}^1 x^n q(x). \quad (1.22)$$

PDFs can be calculated from their Mellin moments using an inverse Mellin transform but one needs to know a sufficient number of moments or else introduce model dependent information. The first few non-trivial moments are calculated in LQCD with the main focus being $n = 0$ and $n = 1$. This is because the moments $\langle x^4 \rangle$ and higher suffer from power-divergent mixing which is unavoidable and has to be eliminated in the renormalization procedure. Even though the PDFs cannot be reliably constructed from just the first three moments, these moments are useful to calculate as constraints to global fits and model calculations. Additionally, the Mellin moments contain physical information. For example, the first moment of the polarized PDF is the fraction of the hadron mass carried by that parton and the first moment of the unpolarized PDF is the parton momentum fraction. Furthermore, together with its zeroth moment, they provide information on the spin and mass decomposition [29, 30, 31].

The collinear PDFs only provide one-dimensional information on the structure of the hadrons under study. To understand the three-dimensional structure of hadrons, generalized parton distributions (GPDs) and transverse momentum-dependent distribution function (TMD PDFs) are also needed [32, 33, 34]. In this thesis, we will focus on PDFs and GPDs. The latter appear in processes such as deeply virtual Compton scattering and deeply virtual pion production and carry information about the distribution of the partons' transverse momentum. In the forward limit, that is zero momentum transfer, GPDs become PDFs. GPDs are also defined in the light-cone frame so calculating them on the lattice come with the same challenges as PDFs. The moments of GPDs, which we study here, are the generalized form factors. Generalized form factors depend on the momentum transfer squared and, in the forward limit, are the PDF Mellin moments.

CHAPTER 2

LATTICE QCD

Lattice Quantum Chromodynamics (LQCD) is the only formulation which solves the QCD starting from the Lagrangian. It uses Monte Carlo simulations and large scale computations. In this chapter, we introduce the motivations for using LQCD as well as its fundamental concepts.

2.1 The Feynman Path Integral

The expectation value of an observable can be calculated using the Feynman path integral [35]

$$\langle O \rangle = \frac{1}{Z} \int \mathcal{D}[\psi, \bar{\psi}] \mathcal{D}[A] e^{-S_E[\psi, \bar{\psi}, A]} O[\psi, \bar{\psi}, A], \quad (2.1)$$

where

$$Z = \int \mathcal{D}[\psi, \bar{\psi}] \mathcal{D}[A] e^{-S_E[\psi, \bar{\psi}, A]} \quad (2.2)$$

is the partition function and S_E is the Euclidean action. $\mathcal{D}[\psi, \bar{\psi}]$ and $\mathcal{D}[A]$ are products of integration measures for the quark, anti-quark and gauge fields at every space-time position, and $O[\psi, \bar{\psi}, A]$ is the functional which represents the observable. In the continuum, there are an infinite number of space-time position and therefore an infinite number of integration measures and paths to integrate over which causes ultraviolet divergences. In order to regularize the formulation so that the path integral can be solved numerically, the continuous space-time is discretized into a 4D lattice [36, 37]

$$\Lambda = \{n = (n_1, n_2, n_3, n_4) \mid n_1, n_2, n_3 = 0, 1, \dots, N_L - 1; n_4 = 0, 1, \dots, N_T - 1\}, \quad (2.3)$$

whose points are a finite distance apart, typically the same distance for all dimensions, which is called the lattice spacing and labeled as a . Now, the integration becomes a sum over all lattice points.

To understand the need for the Euclidean action instead of Minkowski, we write a correlation functions between two operators in the Heisenberg picture, in which operators evolve in time according to $\hat{O}(\tau) = e^{i\tau\hat{H}}\hat{O}e^{-i\tau\hat{H}}$,

$$\begin{aligned}\langle O_2(\tau)O_1(0) \rangle_{\mathcal{T}} &= \frac{1}{Z_{\mathcal{T}}} \text{tr} \left[e^{-i\mathcal{T}\hat{H}} \hat{O}_2(\tau) \hat{O}_1(0) \right] \\ &= \frac{1}{Z_{\mathcal{T}}} \text{tr} \left[e^{-i\mathcal{T}\hat{H}} e^{i\tau\hat{H}} \hat{O}_2 e^{-i\tau\hat{H}} \hat{O}_1 \right] \end{aligned} \quad (2.4)$$

$$\begin{aligned} &= \frac{1}{Z_{\mathcal{T}}} \text{tr} \left[e^{-i(\mathcal{T}-\tau)\hat{H}} \hat{O}_2 e^{-i\tau\hat{H}} \hat{O}_1 \right], \\ Z_{\mathcal{T}} &= \text{tr} \left[e^{-i\mathcal{T}\hat{H}} \right], \end{aligned} \quad (2.5)$$

where τ is in Minkowski time and \mathcal{T} is a formal maximal distance which will later be taken to the limit $\mathcal{T} \rightarrow \infty$. Writing the trace in the basis of eigenvectors $|n\rangle$ such that $\hat{H}|n\rangle = E_n|n\rangle$ and ordering n so that $E_0 \leq E_1 \leq E_2 \dots$, the correlator becomes

$$\begin{aligned}\langle O_2(\tau)O_1(0) \rangle_{\mathcal{T}} &= \frac{1}{Z_{\mathcal{T}}} \sum_{m,n} \langle m | e^{-i(\mathcal{T}-\tau)\hat{H}} \hat{O}_2 | n \rangle \langle n | e^{-i\tau\hat{H}} \hat{O}_1 | m \rangle \\ &= \frac{1}{Z_{\mathcal{T}}} \sum_{m,n} e^{-iE_m(\mathcal{T}-\tau)} \langle m | \hat{O}_2 | n \rangle e^{-iE_n\tau} \langle n | \hat{O}_1 | m \rangle, \end{aligned} \quad (2.6)$$

$$Z_{\mathcal{T}} = \sum_n e^{-iE_n\mathcal{T}}, \quad (2.7)$$

Since the exponentials are complex, they oscillate in complex space and infinite energies are needed in the sum. If we rotate from real to imaginary time ($i\tau = t$ and $i\mathcal{T} = T$), the correlator becomes

$$\langle O_2(t)O_1(0) \rangle_T = \frac{\sum_{m,n} e^{-\Delta E_m(T-t)} \langle m | \hat{O}_2 | n \rangle e^{-\Delta E_n t} \langle n | \hat{O}_1 | m \rangle}{1 + e^{-\Delta E_1 T} + e^{-\Delta E_2 T} + \dots}, \quad (2.8)$$

$$\Delta E_n = E_n - E_0, \quad (2.9)$$

and the exponential terms instead decays with increased energy and imaginary time. This rotation is what is called a Wick rotation and is equivalent to changing from Minkowski to Euclidean space. In the limit $T \rightarrow \infty$, only terms with $m = 0$ remain

and the denominator is one, so the correlator becomes

$$\lim_{T \rightarrow \infty} \langle O_2(t) O_1(0) \rangle_T = \sum_n \langle 0 | \hat{O}_2 | n \rangle \langle n | \hat{O}_1 | 0 \rangle e^{-E_n t}. \quad (2.10)$$

where we have simplified the notation by defining E_0 to be zero so that $\Delta E_n = E_n$. Now, at large enough t , only the ground state dominates.

The exponential factor in Eq. (2.10) is similar to the Boltzmann factor in statistical mechanics. This similarity extends to the rest of Eq. (2.1) which has the same structure as the formula to calculate the average of a function in the canonical ensemble. We can take advantage of the similarity to statistical mechanics by using Monte Carlo simulations to calculate the path integral instead of integrating over every possible path.

2.2 Discretized Fields

In the discretized space-time, the (anti-)quark fields $\psi(x)$ ($\bar{\psi}(x)$) are written as $\psi(n)$ ($\bar{\psi}(n)$), where we denote the position by n to emphasize the discretization, since $x \rightarrow na$. The fermions transform under a gauge transformation the same way as in the continuum

$$\psi(n) \rightarrow \psi'(n) = \Omega(n) \psi(n), \quad (2.11)$$

$$\bar{\psi}(n) \rightarrow \bar{\psi}'(n) = \bar{\psi}(n) \Omega(n)^\dagger. \quad (2.12)$$

Using a Taylor expansion around the unit vector $\hat{\mu}$

$$\psi(n \pm \hat{\mu}) = \psi(n) \pm a \partial_\mu \psi(n) + \frac{a^2}{2} \partial_\mu^2 \psi + \mathcal{O}(a^3), \quad (2.13)$$

we find that the partial derivative of the quarks can be discretized as

$$\partial_\mu \psi(n) \rightarrow \frac{\psi(n + \hat{\mu}) - \psi(n - \hat{\mu})}{2a} + \mathcal{O}(a^2). \quad (2.14)$$

If we look at how the action for free fermions is discretized

$$\begin{aligned}
S_F^0[\psi, \bar{\psi}] &= \int d^4x \bar{\psi}(x)(\gamma_\mu \partial_\mu + m)\psi(x) \\
&\rightarrow S_F^0[\psi, \bar{\psi}] = a^4 \sum_{n \in \Lambda} \bar{\psi}(n) \left(\sum_{\mu=1}^4 \gamma_\mu \frac{\psi(n + \hat{\mu}) - \psi(n - \hat{\mu})}{2a} \right) + m\psi(n), \tag{2.15}
\end{aligned}$$

we see that it is not gauge invariant because the derivative terms transform as

$$\bar{\psi}(n)\psi(n \pm \hat{\mu}) \rightarrow \bar{\psi}'(n)\psi'(n \pm \hat{\mu}) = \bar{\psi}(n)\Omega(n)^\dagger \Omega(n \pm \hat{\mu})\psi(n \pm \hat{\mu}). \tag{2.16}$$

To retain gauge invariance of the action, we need to introduce a field to go between the fermion fields which transforms as

$$U_\mu(n) \rightarrow U'_\mu(n) = \Omega(n)U_\mu(n)\Omega(n + \hat{\mu})^\dagger, \tag{2.17}$$

so that

$$\bar{\psi}(n)U_\mu(n)\psi(n \pm \hat{\mu}) = \bar{\psi}'(n)U'_\mu(n)\psi'(n \pm \hat{\mu}). \tag{2.18}$$

These fields, called gauge links, are part of the gauge group SU(3) and replace the gauge fields A on the lattice. In the continuum limit, they become the gauge transporter which connects a location x to y along a path C_{xy}

$$G(x, y) = P \exp \left(i \int_{C_{xy}} A \cdot dS \right). \tag{2.19}$$

The gauge links are located between the lattice points n and $n + \hat{\mu}$, as illustrated by the left diagram in Figure 2.1. They are also oriented, going from n to $n + \hat{\mu}$, and either the adjoint of a gauge link or a negative index can be used to represent a link oriented in the opposite direction, that is

$$U_{-\mu}(n) = U_\mu(n - \hat{\mu})^\dagger, \tag{2.20}$$

as can be seen in the right diagram in Figure 2.1.

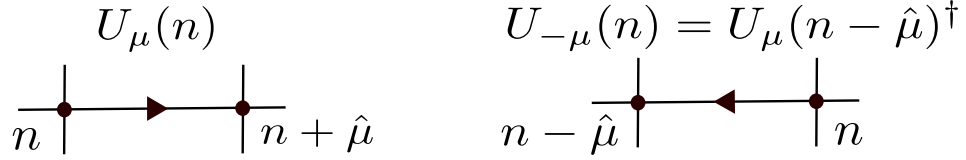


Figure 2.1: A graphical representation of a forward (left) and backward (right) oriented gauge link $U_\mu(n)$ and $U_{-\mu}(n)$.

One object of interest which contains gauge links is a product of k gauge links connecting two lattice points n_0 and n_1 along a path \mathcal{P} . This product is written as

$$\begin{aligned}
 P[U] &= U_{\mu_0}(n_0)U_{\mu_1}(n_0 + \hat{\mu}_0)\dots U_{\mu_{k-1}}(n_1 - \hat{\mu}_{k-1}) \\
 &\equiv \prod_{(n,\mu)\in\mathcal{P}} U_\mu(n).
 \end{aligned}
 \tag{2.21}$$

By transforming the gauge link in $P[U]$ according to Eq. (2.17), it is easy to see that the Ω 's between U 's cancel so that only $\Omega(n_0)$ and $\Omega(n_1)^\dagger$ remain on either end of the product. From this, we find that $P[U]$ transforms similarly to a single gauge link

$$P[U] \rightarrow P[U] = \Omega(n_0)P[U]\Omega(n_1)^\dagger,
 \tag{2.22}$$

so that $\bar{\psi}(n_0)P[U]\psi(n_1)$ is also gauge invariant.

Another gauge invariant object is the trace of a closed path of gauge links along a loop \mathcal{L} which is written as

$$L[U] = \text{tr} \left[\prod_{(n,\mu)\in\mathcal{L}} U_\mu(n) \right].
 \tag{2.23}$$

To see that $L[U]$ is gauge invariant, we again apply Eq. (2.17), cancel all Ω 's between gauge links so that we are left with

$$\begin{aligned}
 L[U'] &= \text{tr} \left[\Omega(n_0) \prod_{(n,\mu)\in\mathcal{L}} U_\mu(n) \Omega(n_0)^\dagger \right] \\
 &= \text{tr} \left[\prod_{(n,\mu)\in\mathcal{L}} U_\mu(n) \right] = L[U],
 \end{aligned}
 \tag{2.24}$$

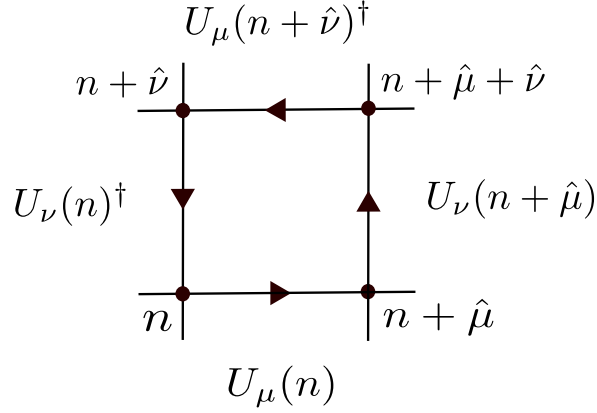


Figure 2.2: A graphical representation of a plaquette $U_{\mu\nu}$.

where in the second line, we applied the cyclic property of traces. The simplest non-trivial loop which can be calculated is a square made of four gauge links with area a^2 . This loop is called a plaquette and is written as

$$\begin{aligned}
 U_{\mu\nu} &= U_{\mu}(n)U_{\nu}(n + \hat{\mu})U_{-\mu}(n + \hat{\mu} + \hat{\nu})U_{-\nu}(n + \hat{\nu}) \\
 &= U_{\mu}(n)U_{\nu}(n + \hat{\mu})U_{\mu}(n + \hat{\nu})^\dagger U_{\nu}(n)^\dagger.
 \end{aligned}
 \tag{2.25}$$

A diagram of a plaquette is given in Figure 2.2.

2.2.1 Grassmann Algebra

Since quarks follow Fermi statistic, the expectation value changes sign if its indices are interchanged, that is

$$\psi^{(f)}(n)_\alpha \psi^{(f')}(n')_{\alpha'} = -\psi^{(f')}(n')_{\alpha'} \psi^{(f)}(n)_\alpha,
 \tag{2.26}$$

$$\bar{\psi}^{(f)}(n)_\alpha \bar{\psi}^{(f')}(n')_{\alpha'} = -\bar{\psi}^{(f')}(n')_{\alpha'} \bar{\psi}^{(f)}(n)_\alpha,
 \tag{2.27}$$

$$\bar{\psi}^{(f)}(n)_\alpha \bar{\psi}^{(f')}(n')_{\alpha'} = -\bar{\psi}^{(f')}(n')_{\alpha'} \bar{\psi}^{(f)}(n)_\alpha.
 \tag{2.28}$$

Numbers with this kind of behavior are called Grassmann numbers [38]. In general, Grassmann numbers obey

$$\eta_i \eta_j = -\eta_j \eta_i, \quad i = 1, 2, \dots, N.
 \tag{2.29}$$

which implies $\eta_i^2 = 0$. So polynomials of Grassmann numbers have a finite number of terms each with up to only one of each Grassmann number since higher order terms are zero. Any Grassmann polynomial can then be expressed as

$$A = a + \sum_i a_i \eta_i + \sum_{i < j} a_{ij} \eta_i \eta_j + \sum_{i < j < k} a_{ijk} \eta_i \eta_j \eta_k + \dots + a_{12\dots N} \eta_1 \eta_2 \dots \eta_N, \quad (2.30)$$

where the coefficients $a_i, a_{ij}, \dots, a_{12\dots N}$ are complex numbers. An algebra, called Grassmann algebra, can be constructed from the addition and multiplication of such polynomials and the Grassmann numbers are the generators.

The derivative rules of Grassmann algebra are

$$\frac{\partial}{\partial \eta_i} 1 = 0, \quad \frac{\partial}{\partial \eta_i} \eta_i = 1, \quad (2.31)$$

$$\frac{\partial}{\partial \eta_i} \frac{\partial}{\partial \eta_j} = - \frac{\partial}{\partial \eta_j} \frac{\partial}{\partial \eta_i}, \quad (2.32)$$

$$\frac{\partial}{\partial \eta_i} \eta_j = -\eta_j \frac{\partial}{\partial \eta_i}, \quad i \neq j. \quad (2.33)$$

As an example of Eqs. (2.32) and (2.33), we write A for $N = 2$

$$A = a + a_1 \eta_1 + a_2 \eta_2 + a_{12} \eta_1 \eta_2, \quad (2.34)$$

and take the partial derivative

$$\frac{\partial A}{\partial \eta_1} = a_1 + a_{12} \eta_2. \quad (2.35)$$

We can also switch the order of the two generators in the last term of A so that

$$A = a + a_1 \eta_1 + a_2 \eta_2 - a_{12} \eta_2 \eta_1, \quad (2.36)$$

and we see that, in order to agree with Eq. (2.35), we need Eq. (2.32) for $N = 2$

$$\frac{\partial}{\partial \eta_1} \eta_2 = -\eta_2 \frac{\partial}{\partial \eta_1}. \quad (2.37)$$

If we apply $\frac{\partial}{\partial \eta_2}$ to Eq. (2.37) we also see the need for Eq. (2.32).

We also need a definition of integration in the Grassmann algebra which has the same properties to the integration in \mathbb{R}^N . The integration over a subset $\Omega \subset \mathbb{R}^N$ of a

function f which is zero at the boundary $\partial\Omega$ is a linear functional of the function f and is given by

$$\int_{\Omega} d^N x f(x) = \int_{\Omega} dx_1 dx_2 \dots dx_N f(x_1, x_2, \dots, x_N). \quad (2.38)$$

In order to ensure that the Grassmann integration is a linear complex functional, it is required that

$$\int d^N \eta A \in \mathbb{C}, \quad \int d^N \eta (\lambda_1 A_1 + \lambda_2 A_2) = \lambda_1 \int d^N \eta A_1 + \lambda_2 \int d^N \eta A_2, \quad (2.39)$$

where λ_1 and λ_2 are complex numbers. Requiring that f vanishes at the boundary implies

$$\int_{\Omega} d^N x \frac{\partial}{\partial x_i} f(x_1, \dots, x_N) = 0, \quad (2.40)$$

and the equivalent statement for the Grassmann algebra is

$$\int d^N \eta \frac{\partial}{\partial \eta_i} A = 0. \quad (2.41)$$

This means that any integral of terms of a Grassmann polynomial which can be expressed as a derivative of other generators is zero. Therefore, only the integration of the highest order term of a polynomial is nonzero and the integral is proportional to the coefficient $a_{1\dots N}$. By requiring the normalization

$$\int d^N \eta \eta_1 \eta_2 \dots \eta_N = 1, \quad (2.42)$$

we find that the integration is

$$\int d^N \eta A = a_{12\dots N}. \quad (2.43)$$

Now, we write the integration measure as a product of measures for each η_i

$$d^N \eta = d\eta_N d\eta_{N-1} \dots d\eta_1, \quad (2.44)$$

which individually have the properties

$$\int d\eta_i \, 1 = 0, \quad \int d\eta_i \, \eta_i = 1, \quad (2.45)$$

$$d\eta_i d\eta_j = -d\eta_j d\eta_i. \quad (2.46)$$

If we consider a transformation of the generators

$$\eta'_i = \sum_{j=1}^N M_{ij} \eta_j, \quad (2.47)$$

which leaves the integration invariant, we find

$$\begin{aligned} \int d^N \eta \, \eta_1 \dots \eta_N &= \int d^N \eta' \, \eta'_1 \dots \eta'_N \\ &= \int d^N \eta' \, \sum_{i_1, \dots, i_N} M_{1i_1} \dots M_{Ni_N} \eta_{i_1} \dots \eta_{i_N} \\ &= \int d^N \eta' \, \sum_{i_1, \dots, i_N} M_{1i_1} \dots M_{Ni_N} \epsilon_{i_1 i_2 \dots i_N} \eta_1 \dots \eta_N \\ &= \det[M] \int d^N \eta' \, \eta_1 \dots \eta_N. \end{aligned} \quad (2.48)$$

In the third step, we reorder the generators and apply the proper sign with $\epsilon_{i_1 i_2 \dots i_N}$ and in the last step, we use the definition of the determinant. So, the integration measure transforms as

$$d^N \eta = \det[M] d^N \eta'. \quad (2.49)$$

We want to use the properties of Grassman variables to find how to calculate the fermionic part of the correlators. We start with the fermionic partition function [39,

$$\begin{aligned}
Z_F &= \int \prod_{i=1}^N d\eta_i d\bar{\eta}_i \exp \left(\sum_{j,k=1}^N \bar{\eta}_j M_{jk} \eta_k \right) \\
&= \det[M] \prod_{i=1}^N \int d\eta'_i d\bar{\eta}_i \exp \left(\sum_{j=1}^N \bar{\eta}_j \eta'_j \right) \\
&= \det[M] \prod_{i=1}^N \int d\eta'_i d\bar{\eta}_i \exp (\bar{\eta}_i \eta'_i) \\
&= \det[M] \prod_{i=1}^N \int d\eta'_i d\bar{\eta}_i (1 + \bar{\eta}_i \eta'_i) \\
&= \det[M] \prod_{i=1}^N \left(\int d\eta'_i d\bar{\eta}_i 1 + \int d\eta'_i d\bar{\eta}_i \bar{\eta}_i \eta'_i \right) \\
Z_F &= \det[M], \tag{2.51}
\end{aligned}$$

where in the second line, we use the transformations Eqs. (2.47) and (2.49), in the fourth we expand the exponential and use $\eta_i^2 = 0$, and in the sixth we use the integrals in Eq. (2.45). Next, we want to calculate the generating functional for fermions which is the integral

$$W[\theta, \bar{\theta}] = \int \prod_{i=1}^N d\eta_i d\bar{\eta}_i \exp \left(\sum_{j,k=1}^N \bar{\eta}_j M_{jk} \eta_k + \sum_{j=1}^N \bar{\theta}_j \eta_j + \sum_{j=1}^N \bar{\eta}_j \theta_j \right). \tag{2.52}$$

We start by rewriting the terms in the exponential

$$\begin{aligned}
\bar{\eta}_j M_{jk} \eta_k + \bar{\theta}_j \eta_j + \bar{\eta}_j \theta_j &= \bar{\eta}_j M_{jk} \eta_k + \bar{\theta}_j \eta_j + \bar{\eta}_j \theta_j + \bar{\theta}_j (M^{-1})_{jk} \theta_k - \bar{\theta}_j (M^{-1})_{jk} \theta_k \\
&= \left(\bar{\eta}_j + \bar{\theta}_i (M^{-1})_{ij} \right) M_{jk} (\eta_k + (M^{-1})_{kl} \theta_l) - \bar{\theta}_j (M^{-1})_{jk} \theta_k. \tag{2.53}
\end{aligned}$$

Then, using the change of variables

$$\eta'_k = \eta_k + (M^{-1})_{kl} \theta_l, \quad \bar{\eta}'_j = \bar{\eta}_j + \bar{\theta}_i (M^{-1})_{ij}, \tag{2.54}$$

$$d\eta'_k = d\eta_k, \quad d\bar{\eta}'_j = d\bar{\eta}_j, \tag{2.55}$$

we rewrite Eq. (2.52) as

$$\begin{aligned}
W[\theta, \bar{\theta}] &= \int \prod_{i=1}^N d\eta_i d\bar{\eta}_i \exp \left(\sum_{j,k=1}^N \bar{\eta}'_j M_{jk} \eta'_k - \sum_{n,m=1}^N \bar{\theta}_n (M^{-1})_{nm} \theta_m \right) \\
&= \exp \left(- \sum_{n,m=1}^N \bar{\theta}_n (M^{-1})_{nm} \theta_m \right) \int \prod_{i=1}^N d\eta_i d\bar{\eta}_i \exp \left(\sum_{j,k=1}^N \bar{\eta}'_j M_{jk} \eta'_k \right)
\end{aligned} \tag{2.56}$$

$$W[\theta, \bar{\theta}] = \det[M] \exp \left(- \sum_{n,m=1}^N \bar{\theta}_n (M^{-1})_{nm} \theta_m \right), \tag{2.57}$$

where in the last line we used the solution to Eq. (2.51). Lastly, we want to calculate expectation values of products of Grassman numbers, called n -point functions,

$$\langle \eta_{j_1} \bar{\eta}_{k_1} \dots \eta_{j_n} \bar{\eta}_{k_n} \rangle = \frac{1}{Z_F} \int \prod_{i=1}^N d\eta_i d\bar{\eta}_i \eta_{j_1} \bar{\eta}_{k_1} \dots \eta_{j_n} \bar{\eta}_{k_n} \exp \left(\sum_{n,m=1}^N \bar{\eta}_n M_{nm} \eta_m \right). \tag{2.58}$$

We can solve Eq. (2.58) by writing the right hand side in terms of the fermion generating function

$$\langle \eta_{j_1} \bar{\eta}_{k_1} \dots \eta_{j_n} \bar{\eta}_{k_n} \rangle = \frac{1}{Z_F} \frac{\partial}{\partial \theta_{k_1}} \frac{\partial}{\partial \bar{\theta}_{j_1}} \dots \frac{\partial}{\partial \theta_{k_n}} \frac{\partial}{\partial \bar{\theta}_{j_n}} W[\theta, \bar{\theta}] \Big|_{\theta, \bar{\theta}=0}, \tag{2.59}$$

and substituting the right hand side of Eq. (2.57) for $W[\theta, \bar{\theta}]$. This results in what is called Wick's theorem

$$\langle \eta_{j_1} \bar{\eta}_{k_1} \dots \eta_{j_n} \bar{\eta}_{k_n} \rangle = (-1)^n \sum_{P(1,2,\dots,n)} \text{sign}(P) (M^{-1})_{i_1 j_{P_1}} (M^{-1})_{i_2 j_{P_2}} \dots (M^{-1})_{i_n j_{P_n}}, \tag{2.60}$$

where $P(1, 2, \dots, n)$ are the permutations of the numbers $1, 2, \dots, n$ and $\text{sign}(P)$ is the sign of the permutation. Comparing Eq. (2.60) to Eq. (2.1), we can use Wick's theorem to calculate correlators from the inverse of our lattice action's Dirac matrix. The inverse Dirac matrix is the quark propagator.

2.2.2 Quark Sources

Calculating the full quark propagator is prohibitively memory-intensive because it carries the information for a source (m, α, a) at every lattice point to a sink (n, β, b) at every lattice point, which is why it is called an all-to-all propagator. There are

also twelve complex numbers associated at each space-time point because there are four Dirac indices and three flavor indices. So the full all-to-all propagator has $(3 \times 4 \times L^3 \times T)^2$ complex numbers. Additionally, the propagator is highly correlated in space-time because it is calculated on the same gauge configuration so, instead, we only calculate the propagators for a few point sources

$$D^{-1}(n|m)_{\beta\alpha_0} = \sum_{m,\alpha,a} D^{-1}(n|m)_{\beta\alpha} S_0^{(m_0,\alpha_0,a_0)}(m)_a^\alpha, \quad (2.61)$$

where $S_0^{(m_0,\alpha_0,a_0)}(m)_a^\alpha$ is the point source

$$S_0^{(m_0,\alpha_0,a_0)}(m)_a^\alpha = \delta(m - m_0)\delta_{\alpha\alpha_0}\delta_{aa_0}, \quad (2.62)$$

located at (m_0, α_0, a_0) . Calculating the propagator for a point source is equivalent to calculating a single column of the all-to-all propagator.

To improve the hadron propagation signal, we smear the point-source fermions in space. The smeared fermions at $n_0 = (\mathbf{n}_0, n_t)$ are marked with a lower index

$$\psi_k^{(f')}(\mathbf{n}_0, n_t)_{\alpha_0} \equiv \sum_{\mathbf{n}_2} S_k^{(n_0,\alpha_0,a_0)}(\mathbf{n}_2)_{\alpha_2}^* \psi^{(f')}(\mathbf{n}_2, n_t)_{\alpha_2}, \quad (2.63)$$

$$\bar{\psi}_i^{(f)}(\mathbf{n}_0, n_t)_{\alpha_0} \equiv \sum_{\mathbf{n}_1} S_i^{(n_0,\alpha_0,a_0)}(\mathbf{n}_1)_{\alpha_1} \bar{\psi}^{(f)}(\mathbf{n}_1, n_t)_{\alpha_1}, \quad (2.64)$$

where we use different indices i and k to emphasize that the smearing functions need not be the same. The smearing function is typically applied multiple times to a point source in order to achieve the desired amount of smearing. We use Gaussian smearing in our simulations which changes the quark fields as

$$S\psi(\mathbf{n}, n_t) = \frac{1}{1 + 6\alpha} \left(\psi(\mathbf{n}, n_t) + \alpha \sum_{j=\pm 1}^{\pm 3} U_j(\mathbf{n}, n_t) \psi(\mathbf{n} + \hat{j}, n_t) \right), \quad (2.65)$$

where α is a parameter which controls how much the quarks are smeared for each step. We choose α and the number of smearing steps so that the root mean squared of the smeared quarks is roughly equal to the radius of the hadron under study. A

smearing quark propagator is then

$$\begin{aligned}
G_{ki}(n_0|m_0)_{\beta_0\alpha_0} &\equiv \left\langle \psi_k^{(f)}(n_0)_{\beta_0} \bar{\psi}_i^{(f)}(m_0)_{\alpha_0} \right\rangle_F \\
&= \left\langle \sum_{\mathbf{n}, \mathbf{m}} S_k^{(n_0, \beta_0, b_0)}(\mathbf{n})_{\beta}^* \psi^{(f)}(n)_{\beta} S_i^{(m_0, \alpha_0, a_0)}(\mathbf{m})_{\alpha} \bar{\psi}(m)_{\alpha} \right\rangle_F \\
&= \sum_{\mathbf{n}, \mathbf{m}} S_k^{(n_0, \beta_0, b_0)}(\mathbf{n})_{\beta}^* (D_f^{-1})(n|m)_{\beta\alpha} S_i^{(m_0, \alpha_0, a_0)}(\mathbf{m})_{\alpha}.
\end{aligned} \tag{2.66}$$

In practice, the way we calculate the smeared propagator is to first calculate the propagator from a smeared source and then apply the smearing at the sink.

2.2.3 Inversion Algorithms

The process of calculating the quark propagator from a source point is to solve the system of equations

$$DG = S \tag{2.67}$$

by inverting the Dirac matrix

$$G = D^{-1}S, \tag{2.68}$$

where G is the quark propagator and S the source point. Inverting a matrix the size of D is computationally intensive and is the main challenge when performing LQCD calculations. The Dirac matrix is a sparse matrix, meaning that the number of matrix elements which are zero is proportional to the size of the matrix. The nonzero elements are located in repeating patterns in the matrix and this pattern depends on the the discretization and boundary conditions used. The fact that the Dirac matrix is sparse means that some inversion algorithms are better suited than others, even though the inverse of the Dirac matrix is not sparse.

One of these inversion methods is the conjugate gradient (CG) solver [41]. The CG solver calculates the vector x^* which minimizes the equation

$$Q(x) = \frac{1}{2}x^T A x - x^T b, \tag{2.69}$$

where A is a real symmetric, positive definite $N \times N$ matrix. The minimizing vector

x^* is also the solution to $Ax = b$ because

$$\partial Q(x)|_{x^*} = Ax - b = 0. \quad (2.70)$$

The CG solver finds x^* iteratively using the formula

$$x^{(i+1)} = x^{(i)} + \alpha_i p^{(i)}, \quad (2.71)$$

where α_i is a real parameter and p_i is a direction vector which is orthogonal to $Ap^{(j)}$ where $p^{(j)}$ is a vector from any previous iteration. These vectors form a vector space called a Krylov space

$$\mathbb{K}^{(i)} = \text{span}(p^{(0)}, \dots, p^{(i)}) \quad (2.72)$$

and $x^{(i)}$ minimizes Eq. (2.69) in $\mathbb{K}^{(i)}$. For an $N \times N$ matrix, $x^* \in \mathbb{K}^{(N)}$ so the maximum number of steps to find the solution is N , though often fewer are needed for a sufficiently accurate approximate solution. At each iteration, the residual $r^{(i)} = b - Ax$ is calculated and the approximate solution is considered sufficiently accurate when $\|r^{(i)}\|$ is less than a given parameter. The number of steps to convergence increases with a greater condition number $K(A)$ which is given by

$$K(A) = \|A\| \|A^{-1}\| = \frac{|\lambda_{\max}|}{|\lambda_{\min}|}, \quad (2.73)$$

where λ_{\max} (λ_{\min}) is the largest (smallest) eigenvalue of A . The only calculations needed for the CG solver are scalar products and vector-matrix multiplications and only the vectors need to be stored which is why this is a good method to use when inverting sparse matrices.

The Dirac matrix is not positive definite hermitian so CG cannot be used to directly calculate the quark propagators. One solution is to invert the matrix $H^2 = DD^\dagger$ instead. The propagator can then be calculated using

$$G = D^{-1}S = D^\dagger H^{-2}S = D^\dagger S. \quad (2.74)$$

This has the drawback that H^2 has a larger condition number than D because $K(H^2) = K(D)^2$. Another solution is to use the bi-conjugate gradient (Bi-GCR)

algorithm which uses two bi-orthogonal search directions and residual vectors. There is also the Bi-GCR Stabilized (Bi-GCRStab) method and its improved version called Bi-GCRStab(2).

A major disadvantage to CG solvers is that the time it takes to solve a system of equations increases as the quark mass decreases, and scales like $1/m^2$. This is what is known as critical slowdown and, as a consequence, the time it takes to invert a Dirac matrix with physical quark masses becomes prohibitively long. A solution to this is to use multigrid (MG) methods [42] which do not suffer from critical slow down.

MG algorithms aim to reduce the time it takes to invert a matrix by coarsening the degrees of freedom. The three main steps in a MG inversion are smoothing, restriction, and prolongation. The smoothing step reduces the short wavelength fluctuations and is usually done before the restriction step and after the prolongations step, called pre- and post-smoothing. In the restriction step an operator R is used to map the points in the fine grid to points in the coarse grid. Similarly, a prolongation operator P is used to map the coarse grid points back to the fine grid points. These two steps are applied to define the coarse Dirac matrix

$$D_c = RDP. \tag{2.75}$$

Take for an example the block aggregation method, in which the restrictor averages over a certain number of points in the fine grid to create one point on the coarse grid. The prolongator is then the operator which satisfies the relation between the identity matrices on the fine and coarse grids

$$I_c = RIP. \tag{2.76}$$

This process can be repeated to create multiple levels of progressively coarser grids. Once the Dirac matrix is calculated for the coarsest grid, it is inverted using a standard inversion routine such as a CG solver before being prolonged back to the fine grid. The time to initialize the coarse grids is longer than it takes to initialize algorithms like CG so MG should be used when the difference between the time it takes to invert on the fine and coarse grid is high enough to make up for the longer setup time of MG or when enough inversions can be performed on the coarse grid that the setup time is worth it.

2.3 Lattice Actions

We can split the calculation of observables we are interested in studying into a gauge and fermionic part. The calculation of the expectation value then becomes

$$\langle O \rangle = \langle \langle O \rangle_F \rangle_G, \quad (2.77)$$

where the fermionic part is

$$\langle A \rangle_F = \frac{1}{Z_F[U]} \int \mathcal{D}[\psi, \bar{\psi}] e^{-S_F[\psi, \bar{\psi}, U]} A[\psi, \bar{\psi}, U], \quad (2.78)$$

$$Z_F[U] = \int \mathcal{D}[\psi, \bar{\psi}] e^{-S_F[\psi, \bar{\psi}, U]}, \quad (2.79)$$

and the gauge part is

$$\langle B \rangle_G = \frac{1}{Z} \int \mathcal{D}[U] e^{-S_G[U]} B[U]. \quad (2.80)$$

The fermionic expectation value is commonly a product of quark and antiquark fields

$$\langle \psi^{(f_1)}(n_1)_{\alpha_1} \dots \psi^{(f_k)}(n_k)_{\alpha_k} \bar{\psi}^{(g_1)}(m_1)_{\beta_1} \dots \bar{\psi}^{(g_k)}(m_k)_{\beta_k} \rangle_F, \quad (2.81)$$

where α_i and β_i are the Dirac indices, a_i and b_i are the color indices, n_i and m_i are the space-time lattice coordinates, and f_i and g_i are the quark flavor indices.

There is some freedom to how the action is discretized, as long as it recovers the form of the continuum QCD action in the continuum limit $a \rightarrow 0$. There are a number of lattice discretizations for both the gluonic and fermionic part of the action. Beyond the simple Wilson discretization, there is the improved three-level Symanzik [43], tadpole improved Lüscher-Weisz (TILW) [43], doubly-blocked Wilson (DBW2) [44] and Iwasaki [45] gluon actions. The latter is used for all calculations presented in this thesis. For the fermion part of the action, there are the Wilson, clover, staggered [46], domain-wall [47], and twisted mass fermions. In this section, we introduce the simplest fermion discretization, the Wilson action, as well as the discretization used in the results obtained in this thesis, the twisted mass formulation.

2.3.1 Wilson Action

The simplest way to discretize the QCD fermion action is

$$S_F[\psi, \bar{\psi}, U] = a^4 \sum_{n \in \Lambda} \bar{\psi}(n) \left(\sum_{\mu=1}^4 \gamma_\mu \frac{U_\mu(n) \psi(n + \hat{\mu}) - U_{-\mu}(n) \psi(n - \hat{\mu})}{2a} + m \psi(n) \right), \quad (2.82)$$

which is called the naive discretization. The Dirac matrix for this discretization is

$$D(n|m) = \sum_{\mu=1}^4 \gamma_\mu \frac{U_\mu(N) \delta_{n+\hat{\mu},m} - U_{-\mu}(n) \delta_{n-\hat{\mu},m}}{2a} + m, \quad (2.83)$$

or, using the notation $\gamma_{-\mu} = -\gamma_\mu$, it can be written more compactly as

$$D(n|m) = \frac{1}{2a} \sum_{\mu=\pm 1}^{\pm 4} \gamma_\mu U_\mu(n) \delta_{n+\mu,m} + m. \quad (2.84)$$

This discretization has an issue, however, which we can see by taking the Fourier transform of the Dirac matrix for the free fermion action, i.e. $U_\mu(n) = \mathbb{1}$

$$\begin{aligned} \tilde{D}(p|q) &= \frac{1}{\sqrt{|\Lambda|}} \sum_{n \in \Lambda} e^{-i(p-q) \cdot na} D(n|m), \\ &= \frac{1}{\sqrt{|\Lambda|}} \sum_{n, m \in \Lambda} e^{-ip \cdot na} \left(\sum_{\mu=1}^4 \gamma_\mu \frac{e^{iq_\mu a} - e^{-iq_\mu a}}{2a} + m \right), \\ &= \delta(p - q) \tilde{D}(p), \end{aligned} \quad (2.85)$$

$$\tilde{D}(p) = m + \frac{i}{a} \sum_{\mu=1}^4 \gamma_\mu \sin(p_\mu a). \quad (2.86)$$

The quark propagator is then

$$\tilde{D}^{-1}(p) = \frac{m - ia^{-1} \sum_{\mu} \gamma_\mu \sin(p_\mu a)}{m^2 + a^{-2} \sum_{\mu} \sin^2(p_\mu a)}, \quad (2.87)$$

and evaluated at $m = 0$ is

$$\tilde{D}^{-1}(p) \Big|_{m=0} = \frac{-ia^{-1} \sum_{\mu} \gamma_\mu \sin(p_\mu a)}{a^{-2} \sum_{\mu} \sin^2(p_\mu a)}. \quad (2.88)$$

Taking the continuum limit, this becomes

$$\tilde{D}^{-1}(p)\Big|_{m=0} \xrightarrow{a \rightarrow 0} \frac{-i \sum_{\mu} p_{\mu}}{p^2}, \quad (2.89)$$

so there is a pole at $p = (0, 0, 0, 0)$ in the continuum. On the lattice however, there are additional, nonphysical poles at $p = (\pi/a, 0, 0, 0)$, $p = (0, \pi/a, 0, 0)$, ..., $p = (\pi/a, \pi/a, \pi/a, \pi/a)$ due to the \sin^2 function in the denominator. The nonphysical particles at these poles are known as doublers. We can remove the doublers by introducing a so-called Wilson term [48] to the Dirac matrix

$$\tilde{D}(p) = m + \frac{i}{a} \sum_{\mu=1}^4 \gamma_{\mu} \sin(p_{\mu}a) + \tilde{W}, \quad (2.90)$$

$$\tilde{W} = \frac{1}{a} \sum_{\mu=1}^4 (1 - \cos(p_{\mu}a)). \quad (2.91)$$

This term is zero at $p = (0, 0, 0, 0)$, so it preserves the physical pole but, by calculating the inverse of Eq. (2.90), we find that it cancels out the doublers. Fourier transforming the Wilson term results in

$$W = -\frac{a}{2} \sum_{\mu=1}^4 \frac{U_{\mu}(n)_{ab} \delta_{n+\hat{\mu},m} - 2\delta_{ab} \delta_{n,m} + U_{-\mu}(n) \delta_{n-\hat{\mu},m}}{a^2}. \quad (2.92)$$

The sum in Eq. (2.92) is the discretization of the Laplace operator $\partial_{\mu} \partial_{\mu}$ and since it is multiplied by $-a/2$, the Wilson term goes to zero in the continuum limit. By adding the Wilson term to Eq. (2.83), we get what is called the Wilson action

$$D_W(n|m) = \left(m + \frac{4}{a}\right) - \frac{1}{2a} \sum_{\mu=\pm 1}^{\pm 4} (\mathbb{1} - \gamma_{\mu}) U_{\mu}(n) \delta_{n+\hat{\mu},m}. \quad (2.93)$$

Since the introduction of the Wilson action, there have been a lot of improvements to bring lattice results closer to the continuum limit by removing $\mathcal{O}(a)$ terms. Improved formulations are the clover-improved Wilson action, Twisted-mass fermions, as well as domain-wall and staggered fermions.

2.3.2 Twisted Mass Fermion Action

The twisted mass fermion action [49, 50] is another discretized fermion action and is the lattice action we use for our calculations. It is equivalent to the Euclidean QCD fermion action and, in the continuum, is written as

$$S_F^{\text{tm}}[\chi, \bar{\chi}, A] = \int d^4x \bar{\chi}(\not{D} + m + i\mu\gamma_5\tau_3)\chi, \quad (2.94)$$

where χ and $\bar{\chi}$ are mass-degenerate flavor doublets, τ_3 is the third Pauli matrix in flavor space, and μ is what is called the twisted mass. Let us consider the Euclidean QCD fermion action in order to understand its equivalence with the twisted mass action:

$$S_F[\psi, \bar{\psi}, A] = \int d^4x \bar{\psi}(\not{D} + M)\psi. \quad (2.95)$$

The form of the action is invariant under a transformation into the *twisted basis* via the relationship

$$\chi = e^{i\alpha\gamma_5\tau_3/2}\psi, \quad \bar{\chi} = \bar{\psi}e^{i\alpha\gamma_5\tau_3/2}, \quad (2.96)$$

and the mass transforms as

$$M' = Me^{i\alpha\gamma_5\tau_3}. \quad (2.97)$$

By expanding the exponential, collecting the even and odd terms, and using the fact that $\gamma_5^2 = \mathbb{1}$ and $\tau_3^2 = \mathbb{1}$, we can rewrite Eq. (2.97) as

$$M' = M(\cos \alpha + i\gamma_5\tau_3 \sin \alpha). \quad (2.98)$$

With the definitions

$$m = M \cos \alpha, \quad (2.99)$$

$$\mu = M \sin \alpha, \quad (2.100)$$

$$M = \sqrt{m^2 + \mu^2}, \quad (2.101)$$

$$\alpha = \arctan(\mu/m), \quad (2.102)$$

we see that Eq. (2.94) is simply Eq. (2.95) written in a different basis.

The major advantage to using twisted mass fermions is that, when $\alpha = \pi/2$ and, therefore, $m = 0$, which is called *maximal twist*, quantities calculated with the twisted mass action are automatically $\mathcal{O}(a)$ improved [51] without any other modifications to

the action. This can be seen by using the Symanzik improvement [52, 53] in which an effective action is obtained by expanding the action in terms of a , that is,

$$S_{\text{eff}} = S_0 + aS_1 + a^2S_2 + \dots, \quad (2.103)$$

and by requiring that each term is invariant under the same symmetries as the lattice action. S_0 is the continuum action and

$$S_k = \int d^4y \mathcal{L}_k(y) \quad (2.104)$$

where the Lagrangians $\mathcal{L}_k(y)$ are $4 + k$ -dimension linear combinations of local fields. We write correlation functions as

$$G_x(x_1, \dots, x_n) = \langle \phi(x_1) \dots \phi(x_n) \rangle \equiv \langle \Phi \rangle, \quad (2.105)$$

where $\phi(x)$ is a generic combination of gauge invariant fields made up of quark and gluon fields. These field are also expanded as

$$\phi_{\text{eff}} = \phi_0 + a\phi_1 + a^2\phi_2 + \dots, \quad (2.106)$$

and the correlation function is expanded as

$$\langle \Phi \rangle = \langle \Phi_0 \rangle_0 - a \int d^4y \langle \Phi_0 \mathcal{L}_1(y) \rangle_0 + a \langle \Phi_1 \rangle_0 + \mathcal{O}(a^2). \quad (2.107)$$

The notation $\langle \dots \rangle_0$ means that the expectation values is taken with the action as S_0 and

$$\langle \Phi_0 \rangle_0 \equiv \langle \phi_0(x_1) \dots \phi_0(x_n) \rangle_0, \quad (2.108)$$

$$\langle \Phi_1 \rangle_0 \equiv \sum_{k=1}^n \langle \phi_0(x_1) \dots \phi_1(x_k) \dots \phi_0(x_n) \rangle_0. \quad (2.109)$$

At maximal twist, that is, $m = 0$, the continuum twisted mass action in Eq. (2.94) becomes

$$S_F = \int d^4\bar{\chi} (\not{D} + i\mu\gamma_5\tau_3)\chi, \quad (2.110)$$

and is invariant under the discrete chiral symmetry transformation $\mathcal{R}_5^{1,2}$

$$\chi(x) \xrightarrow{\mathcal{R}_5^{1,2}} \chi'(x) = i\gamma_5\tau_{1,2}\chi(x), \quad (2.111)$$

$$\bar{\chi}(x) \xrightarrow{\mathcal{R}_5^{1,2}} \bar{\chi}'(x) = \bar{\chi}(x)i\gamma_5\tau_{1,2}. \quad (2.112)$$

In this case, \mathcal{L}_1 in Eq.(2.107) is

$$\mathcal{L}_1 = i\bar{\chi}\sigma_{\mu\nu}F_{\mu\nu}\chi + \mu^2\bar{\chi}\chi, \quad (2.113)$$

where $\mathcal{F}^{\mu\nu}$ is the discretized field strength tensor and μ is the twisted mass. If the correlation function is even under the $\mathcal{R}_5^{1,2}$ transformation in Eq. (2.112), then $\langle\Phi_0\mathcal{L}_1(y)\rangle_0$ in the second term of the right hand side of Eq. (2.107) is odd. But the terms in Eq. (2.107) need to have the same symmetries as S_0 which is invariant under $\mathcal{R}_5^{1,2}$, therefore,

$$\langle\Phi_0\mathcal{L}_1\rangle_0 = \langle\Phi_0\mathcal{L}_1'\rangle_0 = -\langle\Phi_0\mathcal{L}_1\rangle_0 = 0. \quad (2.114)$$

The operator Φ_1 is not chirally symmetric, however, it is invariant under the transformation $\mathcal{R}_5^{1,2} \times \mathcal{D}$, where the transformation \mathcal{D} is given by

$$\chi(x) \xrightarrow{\mathcal{D}} \chi'(x) = e^{3i\pi/2}\chi(-x), \quad (2.115)$$

$$\bar{\chi}(x) \xrightarrow{\mathcal{D}} \bar{\chi}'(x) = \bar{\chi}(-x)e^{3i\pi/2}, \quad (2.116)$$

and essentially counts the dimensions of the operator. This means that the operator Φ_1 has the opposite chirality as Φ_0 and, since the continuum action is chirally symmetric, the third term in Eq. (2.107) is also zero. So both $\mathcal{O}(a)$ terms vanish at maximal twist and the correlation function is automatically $\mathcal{O}(a)$ improved.

In our simulations, we use the twisted-clover action which has the added clover term from the Symanzik improvement, so that the action becomes

$$S_F[\chi, \bar{\chi}, U] = a^4 \sum_{x \in \Lambda} \bar{\chi}(x) \left(D_W(U) + \frac{i}{4} c_{SW} \sigma^{\mu\nu} \mathcal{F}^{\mu\nu}(U) + m + i\mu\gamma_5\tau_3 \right) \chi(x), \quad (2.117)$$

where c_{SW} is the Sheikholeslami-Wohlert improvement coefficient [54]. The main motivation is that in the standard clover improved action, the clover term is included to achieve $\mathcal{O}(a)$ improvement. The addition of the clover improvement to the twisted mass formulation has been proven beneficial for ensembles at or near the physical

point, it reduces isospin symmetry breaking effects [55]. The addition of the clover term, is relevant for all the ensembles that are part of this thesis.

2.4 Gauge Field Configurations

In the canonical ensemble of statistical mechanics, the average of a function $f(x)$ is calculated using the formula

$$\langle f \rangle = \frac{1}{Z} \int dx P(x) f(x) = \frac{\int dx e^{-E(x)\beta} f(x)}{\int dx f(x)}, \quad (2.118)$$

where $E(x)$ is the energy distribution and $\beta = 1/k_B T$. Comparing this to Eq. (2.1), the similarities are obvious. Because of this similarity and since we cannot evaluate the integral in Eq. (2.1) over every possible gauge field, even after discretization, we use a Monte Carlo approach to calculating the expectation value. That is, we calculate the quantity of interest on N gauge field configurations U_n which have a probability distribution proportional to $e^{-S_G[U_n]}$ and average the results over the gauge configurations or, stated mathematically,

$$\langle O \rangle = \frac{1}{N} \sum_{\substack{U_n \text{ with} \\ \text{probability} \\ \propto e^{-S_G[U_n]}}} O[U_n]. \quad (2.119)$$

This is equivalent to simulating the quantum fluctuations in the vacuum.

2.4.1 Markov Chains

To get the necessary probability distribution, we need to generate a *Markov chain* [56] of gauge configurations

$$U_0 \rightarrow U_1 \rightarrow \dots \rightarrow U_n \rightarrow \dots \quad (2.120)$$

which means that, starting with an arbitrary gauge configuration U_0 , we produce a new gauge configuration from the last by applying some process. Each update to a new configuration, or Monte Carlo step, can be thought of as a step in a nonphysical dimension, sometimes called computer time. This process is done until the equilibrium distribution $P(U)$ is reached. The update process needs to satisfy some conditions.

We write the probability of getting to a configuration U_n from U_{n-1} as

$$P(U_n = U' | U_{n-1} = U) = T(U'|U). \quad (2.121)$$

Since it is a probability,

$$0 \leq T(U'|U) \leq 1. \quad (2.122)$$

The normalization condition is that the total probability of updating from some configuration to any other configurations (including itself) is 1

$$\sum_{U'} t(U'|U) = 1. \quad (2.123)$$

Another condition is

$$\sum_U T(U'|U)P(U) = \sum_U T(U|U')P(U'), \quad (2.124)$$

meaning that updating into U' has the same probability as updating out of U' . This condition ensures that no extra probability is entering or leaving our Markov chain and is called the balance condition. One solution to the balance condition, called the detailed balance condition, is

$$T(U'|U)P(U) = T(U|U')P(U'), \quad (2.125)$$

which requires that each term of the sum in the balance condition holds true.

One updating method is the Metropolis algorithm [57]. A general overview of the Metropolis algorithm is that we use some *a priori* probability $T_0(U'|U)$ to select a new configuration U' from U and accept the change with probability

$$T_A(U'|U) = \min \left(1, \frac{T_0(U|U') \exp(-S[U'])}{T_0(U'|U) \exp(-S[U])} \right). \quad (2.126)$$

These steps are then repeated. The total acceptance probability is $T = T_0 T_A$ and

$$\begin{aligned}
T(U'|U)P[U] &= T(U'|U) \exp(-S[U]) \\
&= T_0(U'|U) \min \left(1, \frac{T_0(U|U') \exp(-S[U'])}{T_0(U'|U) \exp(-S[U])} \right) \exp(-S[U]) \\
&= \min (T_0(U'|U) \exp(-S[U]), T_0(U|U') \exp(-S[U'])) \\
&= T(U|U') \exp(-S[U']) \\
&= T(U|U')P[U'],
\end{aligned} \tag{2.127}$$

so the Metropolis algorithm obeys the detailed balance condition. If we choose a symmetric selection probability such that $T_0(U|U') = T_0(U'|U)$, the acceptance probability in Eq. (2.126) simplifies to

$$T_A(U'|U) = \min(1, \exp(-\Delta S)), \tag{2.128}$$

where $\Delta S = S[U'] - S[U]$ is the difference between the action of each configuration.

For an $SU(N)$ lattice gauge action, a single gauge link is updated at each update step. For the Wilson gauge action, the local action which is affected by one gauge link change is

$$\begin{aligned}
S[U_\mu(n)']_{\text{loc}} &= \frac{\beta}{N} \sum_{i=1}^6 \text{Re tr}[\mathbb{1} - U_\mu(n)'P_i] \\
&= \frac{\beta}{N} \text{Re tr}[\mathbb{1} - U_\mu(n)'A],
\end{aligned} \tag{2.129}$$

$$\begin{aligned}
A &= \sum_{i=1}^6 P_i = \sum_{\nu \neq \mu} (U_\nu(n + \hat{\mu})U_{-\mu}(n + \hat{\mu} + \hat{\nu})U_{-\nu}(n + \hat{\nu}) \\
&\quad + U_{-\nu}(n + \hat{\mu})U_{-\mu}(n + \hat{\mu} - \hat{\nu})U_\nu(n - \hat{\nu})),
\end{aligned} \tag{2.130}$$

where P_i , called a staple, is the product of the other three gauge links in each plaquette including $U_\mu(n)$. In this case, the change in the action is

$$\begin{aligned}
\Delta S &= S[U_\mu(n)']_{\text{loc}} - S[U_\mu(n)]_{\text{loc}} \\
&= -\frac{\beta}{N} \text{Re tr}[(U_\mu(n)' - U_\mu(n))A].
\end{aligned} \tag{2.131}$$

A common method for selecting $U_\mu(n)'$ is by the matrix multiplication

$$U_\mu(n)' = XU_\mu(n), \quad (2.132)$$

where X is a random element in $SU(N)$ and is around $\mathbb{1}$. An example of how to generate X for $SU(2)$ is to pick a random four-vector $r = (r_0, \mathbf{r})$ where $r_i \in (-1/2, 1/2)$ for $i = 0, 1, 2, 3$ and compute

$$x_0 = \text{sign}(r_0)\sqrt{1 - \epsilon^2}, \quad (2.133)$$

$$\mathbf{x} = \epsilon\mathbf{r}/|\mathbf{r}|, \quad (2.134)$$

where ϵ is the parameter that controls how far away X is spread around $\mathbb{1}$. Then X is calculated by

$$X = x_0\mathbb{1} + i\mathbf{x} \cdot \boldsymbol{\sigma}, \quad (2.135)$$

where $\boldsymbol{\sigma}$ is a vector of the Pauli matrices σ_i for $i = 1, 2, 3$. A similar process can be done for $SU(3)$ where we calculate three $SU(2)$ elements r , t , and s following the same procedure and then construct from these the matrices

$$R = \begin{pmatrix} r_{11} & r_{12} & 0 \\ r_{21} & r_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad S = \begin{pmatrix} s_{11} & 0 & s_{12} \\ 0 & 1 & 0 \\ s_{21} & 0 & s_{22} \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & t_{11} & t_{12} \\ 0 & t_{21} & t_{22} \end{pmatrix}. \quad (2.136)$$

From these, we calculate X by matrix multiplication

$$X = RST. \quad (2.137)$$

Now, the steps in the Metropolis algorithm are to update $U_\mu(n)$ to $U_\mu(n)'$ according to Eq. (2.132), accept $U_\mu(n)'$ if a random number r between 0 and 1 is less than or equal to $\exp(-\Delta S)$, and then repeat the process.

2.4.2 Dynamical Fermions

So far, the gauge configuration generation algorithms which have been discussed do not use any information about the fermions. Quark flavors which are not included in simulations are said to be quenched. In nature, the quantum fluctuations in the vacuum are caused by both quarks and gluons, so by including quarks in the produc-

tion of gauge configurations, called dynamical quarks, the simulations will be more similar to reality.

From Eqs. (2.51) and (2.57), we could include the up and down quark anti-quark loops when averaging over gauge fields by using

$$\langle O[\psi, \bar{\psi}] \rangle = \frac{\langle \det[D_u] \det[D_d] O[U] \rangle_G}{\langle \det[D_u] \det[D_d] \rangle_G}. \quad (2.138)$$

This calculation is noisy, however, because the determinant can vary a lot between gauge configurations. Instead, we include dynamical fermions by using the determinant as a weight in the probability of generating gauge configurations, that is,

$$P[U] \propto \frac{1}{Z} e^{-S_G[U]} \det[D_u] \det[D_d], \quad (2.139)$$

as long as the determinant is real and positive. The Dirac matrix is positive γ_5 -hermitian, that is, $\gamma_5 D \gamma_5 = D^\dagger$, then its determinant is positive, shown by

$$\det[D]^* = \det[D^\dagger] = \det[\gamma_5 D \gamma_5] = \det[D]. \quad (2.140)$$

Additionally, if the quark flavors are degenerate, then the product of their Dirac matrix determinants are positive, shown by,

$$\det[D_u] \det[D_d] = \det[D] \det[D] = \det[D] \det[D^\dagger] = \det[DD^\dagger] \geq 0. \quad (2.141)$$

The twisted-clover action is γ_5 -hermitian so, as long as it has degenerate quark pairs, its determinant is real and positive.

The Hybrid Monte Carlo algorithm [58, 59] is a Metropolis algorithm which includes dynamical fermions and uses concepts analogous to Hamiltonian mechanics [60, 61]. It uses these concepts so that auto correlations between configurations are low but the acceptance rate is not too low which becomes expensive. As an example, consider a real scalar Q with conjugate momentum P generated with a Gaussian probability $P_G(G) \propto \exp(-P^2)$. We define the action to be

$$S[Q, P] \equiv -\frac{1}{2} P^2 - S[Q], \quad (2.142)$$

which results in the Hamiltonian

$$H[Q, P] = \frac{1}{2}P^2 + S[Q] \quad (2.143)$$

and the computer time derivative

$$\dot{Q} = \frac{d}{d\tau}Q, \quad \dot{P} = \frac{d}{d\tau}P. \quad (2.144)$$

Then, the classical equations of motion are what are called molecular dynamics equations

$$\dot{P} = -\frac{\partial H}{\partial Q} = -\frac{\partial S}{\partial Q}, \quad (2.145)$$

$$\dot{Q} = \frac{\partial H}{\partial P} = P. \quad (2.146)$$

Because this is a numeric calculation changes in computer time are discrete steps, that is,

$$\epsilon \equiv \Delta\tau, \quad (2.147)$$

which lead to unavoidable errors. A number of steps in a row is called a trajectory, for which we use the notation

$$(P(\tau + n\epsilon), Q(\tau + n\epsilon)) \equiv (P', Q'). \quad (2.148)$$

An algorithm with lower error due to discrete time steps than a simple linear integration and meets the detailed balance equation Eq. (2.125) is leapfrog integration. This integration gets its name because the P is updated by a half step which is then used to update Q by a full step. From Eqs. (2.145) and (2.146), Q and P change by these updates to

$$Q(0) \longrightarrow Q(\epsilon) = Q(0) + P(\epsilon/2)\epsilon, \quad (2.149)$$

$$P(0) \rightarrow P(\epsilon/2) = P(0) - \left. \frac{\partial S}{\partial Q} \right|_{Q(0)} \frac{\epsilon}{2} \rightarrow P(\epsilon) = P(\epsilon/2) - \left. \frac{\partial S}{\partial Q} \right|_{Q(\epsilon)} \frac{\epsilon}{2}. \quad (2.150)$$

For ease of notation, we define Q and P after multiple steps as

$$Q(n\epsilon) \equiv Q_n, \quad P(n\epsilon) \equiv P_n, \quad (2.151)$$

and then we can write the change in Q and P after one full step as

$$Q_1 = Q_0 + P_0\epsilon - \frac{1}{2} \left. \frac{\partial S}{\partial Q} \right|_{Q_0} \epsilon^2, \quad (2.152)$$

$$P_1 = P_0 + \frac{1}{2} \left(\left. \frac{\partial S}{\partial Q} \right|_{Q_0} + \left. \frac{\partial S}{\partial Q} \right|_{Q_1} \right) \epsilon. \quad (2.153)$$

In order for the detailed balance equation to be met, there are two conditions. The first is that the integration measure $\mathcal{D}[P]\mathcal{D}[Q]$ is invariant and the second is that the trajectory is reversible

$$T_{\text{md}}(P', Q'|P, Q) = T_{\text{md}}(-P, Q|-P', Q'). \quad (2.154)$$

To show that the integration measure is invariant, we expand the Jacobian using the chain rule

$$\det \left[\frac{\partial(P_1, Q_1)}{\partial(P_0, Q_0)} \right] = \det \left[\frac{\partial(P_1, Q_1)}{\partial(P_{1/2}, Q_1)} \frac{\partial(P_{1/2}, Q_1)}{\partial(P_{1/2}, Q_0)} \frac{\partial(P_{1/2}, Q_0)}{\partial(P_0, Q_0)} \right]. \quad (2.155)$$

The three Jacobians on the right hand side are triangle matrices so their determinant is 1, for example, the determinant of the last Jacobian is

$$\det \left[\frac{\partial(P_{1/2}, Q_0)}{\partial(P_0, Q_0)} \right] = \det \begin{bmatrix} 1 & \dots \\ 0 & 1 \end{bmatrix} = 1 \quad (2.156)$$

and similarly for the other two. This proves that the integration measure is invariant to leapfrog integration steps. To show that leapfrog integration steps are reversible, we use Eqs. (2.152) and (2.153) to calculate Q and P starting at computer time ϵ and taking one step back

$$\begin{aligned} Q(\epsilon - \epsilon) &= Q_1 - P_1\epsilon - \frac{1}{2} \left. \frac{\partial S}{\partial Q} \right|_{Q_1} \epsilon^2, \\ &= Q_1 - P_0\epsilon + \frac{1}{2} \left(\left. \frac{\partial S}{\partial Q} \right|_{Q_0} + \left. \frac{\partial S}{\partial Q} \right|_{Q_1} \right) \epsilon^2 - \frac{1}{2} \left. \frac{\partial S}{\partial Q} \right|_{Q_1} \epsilon^2, \\ &= Q_1 - P_0\epsilon + \frac{1}{2} \left. \frac{\partial S}{\partial Q} \right|_{Q_0} \epsilon^2 = Q_0, \end{aligned} \quad (2.157)$$

$$P(\epsilon - \epsilon) = P_1 + \frac{1}{2} \left(\left. \frac{\partial S}{\partial Q} \right|_{Q_1} + \left. \frac{\partial S}{\partial Q} \right|_{Q_0} \right) \epsilon = P_0. \quad (2.158)$$

Now that we know how we will update our configurations, we accept an update with probability

$$T_A(P', Q'|P, Q) = \min \left(1, \frac{\exp(-H[P', Q'])}{\exp(-H[P, Q])} \right), \quad (2.159)$$

so that the full probability of producing a configuration Q' from a configuration Q is

$$T(Q'|Q) = \int \mathcal{D}[P]\mathcal{D}[P']T_A(P', Q'|P, Q)T_{\text{md}}(P', Q'|P, Q)e^{-P^2/2}. \quad (2.160)$$

We prove that Eq. (2.160) satisfies the detailed balance equation and is thus a valid Metropolis algorithm by writing the acceptance probability as

$$\begin{aligned} T_A(P', Q'|P, Q) &= \min \left(1, \frac{\exp(-P'^2/2 - S[Q'])}{\exp(-P^2/2 - S[Q])} \right) \\ &= \exp(-P^2/2 - S[Q'] + P^2/2 + S[Q]) \min \left(\frac{\exp(-P^2/2 - S[Q])}{\exp(-P'^2/2 - S[Q'])}, 1 \right) \\ &= \exp(-P'^2/2 - S[Q'] + P^2/2 + S[Q])T_A(P, Q|P', Q') \\ &= \exp(-P'^2/2 - S[Q'] + P^2/2 + S[Q])T_A(-P, Q|-P', Q'), \end{aligned} \quad (2.161)$$

and the full probability as

$$\begin{aligned} T(Q'|Q) &= \int \mathcal{D}[P]\mathcal{D}[P']T_A(-P, Q|-P', Q')T_{\text{md}}(-P, Q|-P'|Q')e^{-S[Q'] + S[Q] - P'^2/2} \\ &= \int \mathcal{D}[P]\mathcal{D}[P']T_A(P, Q|P', Q')T_{\text{md}}(P, Q|P'|Q')e^{-S[Q'] + S[Q] - P'^2/2}. \end{aligned} \quad (2.162)$$

From here, we can easily see that

$$e^{-S[Q]}T(Q'|Q) = e^{-S[Q']}T(Q|Q'), \quad (2.163)$$

that is, that the detailed balance equation is satisfied.

To understand how we use the molecular dynamics equations and leap-frog integration in our lattice simulations, first consider a vector of N complex variables $\phi = \phi_R + i\phi_I$ which are called pseudofermions. The pseudofermions have the same degrees of freedom as fermions, are eigenvectors of some matrix A , and their eigen-

values have a positive real part. The generating equation for this field is

$$\begin{aligned} W[\chi, \chi^\dagger] &\equiv \int_{\mathbb{R}^{2N}} \prod_{i=1}^N (d\phi_{R,i}, d\phi_{I,i}) \exp \left(- \sum_{i,j=1}^N \phi_i^\dagger A_{ij} \phi_j + \sum_{i=1}^N \phi_i^\dagger \chi_i + \sum_{i=1}^N \chi_i^\dagger \phi_i \right) \\ &= \frac{\pi^N}{|\det[A]|} \exp \left(\sum_{i,j=1}^N \chi_i^\dagger (A^{-1})_{ij} \chi_j \right), \end{aligned} \quad (2.164)$$

which is similar to the generating function Eq. (2.52) for fermions except the determinant is in the denominator and there is an irrelevant factor π^N . From Eq. (2.164), get can relate a Dirac matrix with a pair of flavor degenerate fermions to the pseudo fermions by [62]

$$\det[DD^\dagger] = \pi^{-N} \int \mathcal{D}[\phi_R] \mathcal{D}[\phi_I] e^{\phi^\dagger (DD^\dagger)^{-1} \phi}. \quad (2.165)$$

so that we can rewrite the path integral over fermions as

$$\int \mathcal{D}[\psi] \mathcal{D}[\bar{\psi}] d^{-\bar{\psi}_u D\psi_u - \bar{\psi}_d D\psi_d} = \pi^{-N} \int \mathcal{D}[\phi_R] \mathcal{D}[\phi_I] e^{-\phi^\dagger (DD^\dagger)^{-1} \phi}. \quad (2.166)$$

Now, instead of generating fermions, we generate pseudo-fermions from $\phi = D\chi$, where χ has a Gaussian generation probability $T(\chi) \propto e^{-\chi^\dagger \chi}$. From this and our gauge action, we calculate the ‘‘force’’ used to update to the next computer time step

$$F[U, \phi] = \sum_{i=1}^8 T_i \nabla^{(i)} (S_G[U] + \phi^\dagger (DD^\dagger)^{-1} \phi) \in \text{SU}(3). \quad (2.167)$$

We also need to generate our initial conjugate momentum P_0 which has the same degrees of freedom as the gauge fields and Gaussian generation probability $T(P_0) \propto \exp(-\text{tr}[P_0^2])$. Then, we update U and P using leapfrog steps

$$U_k = \exp \left(i\epsilon P_{k-\frac{1}{2}} \right) U_{k-1}, \quad P_{k+\frac{1}{2}} = P_{k-\frac{1}{2}} - \epsilon F[U, \phi]|_{U_k}, \quad (2.168)$$

For the first and last steps, P is instead updated with

$$P_{\frac{1}{2}} = P_0 - \frac{\epsilon}{2} F[U, \phi]|_{U_0}, \quad P_n = P_{n-\frac{1}{2}} - \frac{\epsilon}{2} F[U, \phi]|_{U_{n-\frac{1}{2}}}, \quad (2.169)$$

respectively. The trajectory is accepted if the condition for a random number r

$$r \in [0, 1) < \exp(\text{tr}[P'^2] + S_G[U'] + \phi^\dagger ((DD^\dagger)^{-1} - (D'D'^\dagger)^{-1}) \phi) \quad (2.170)$$

is met.

2.4.3 Gauge Smearing

In LQCD simulations, we are interested in the long distance behaviour of the correlation functions. This can have significant statistical uncertainties, however, if there are a lot of short range fluctuations in the the gauge field. These fluctuations can be reduced by smearing the gauge configurations.

When smearing, we first calculate the average over other nearby links. There are multiple methods of smearing which include different nearby gauge links in the average. We use APE smearing [63], where the six staples connected to a gauge link are averaged over. The formula used is

$$V_\mu(n) = (1 - \alpha)U_\mu + \frac{\alpha}{6} \sum_{\nu \neq \mu} C_{\mu\nu}, \quad (2.171)$$

where $C_{\mu\nu}(n)$ are the staples in the right hand side of Eq. (2.130) and the α is a real parameter which is chosen based on the gauge coupling. The average over gauge links is not necessarily still a member of $SU(3)$ so we need to project it back to $SU(3)$ by maximizing

$$\text{Re tr}[XV_\mu(n)^\dagger], X \in SU(3), \quad (2.172)$$

and then we use X as the new link variable $U'_\mu(n)$.

2.5 Systematic uncertainties

There are a number of systematic uncertainties inherent to LQCD. These systematic effects arise from finite volume, discretization, excited-state contamination, and non-physical quark masses. Here, we cover the basics of each source of systematic effect.

Finite volume errors are present in LQCD results because simulations are performed on finite lattices due to numerical limitations. To study these effects and estimate the related uncertainty, physical quantities must be calculated on multiple lattice ensembles with varying lattice sizes so that the volume dependence can be

determined. These results are then fitted to an ansatz proportional to powers of $1/L$ or some other volume dependence based on a model so that the infinite volume limit $V \rightarrow \infty$ can be taken. This can become expensive because, to properly study volume effects, the volume is varied and the lattice parameter a stays the same so the simulations require more and more lattice points.

The next systematic effect is present in LQCD because lattices are discretized with a finite lattice parameter. Similarly to the finite volume effects, discretization errors can be studied by performing calculations on multiple ensembles only now the lattice parameter is varied rather than the volume. After fitting these calculations, the results are extrapolated to the continuum limit $a \rightarrow 0$.

The quark masses are input parameters in the gauge configurations so they are tuned for an ensemble by comparing the lattice calculation of a hadron mass—typically a pion—to its physical value. Lattice calculations are less expensive with larger quark masses and it is only in the past 5–7 years that computers and algorithms have advanced to the point where simulations are feasible on ensembles with physical pion masses, known as the *physical point*. Even still, higher than physical pion masses are often used to make simulations less expensive which is useful when exploring new quantities. This does, however, cause another systematic uncertainty which can again be studied by calculating lattices results on ensembles with different pion masses. The function for these fits is determined using chiral perturbation theory (χ PT) and the extrapolation is taken to the physical point. χ PT is very successful for the mesonic sector, while the baryonic sector is less-studied.

When studying hadron structure, we primarily want to calculate ground state contributions to the quantities of interest, however, the hadron interpolators in the matrix elements only carry information about the quantum numbers and flavor composition so excited states can also be created, causing another source of uncertainty. Unlike the previous systematic uncertainties, the excited state contamination can be studied on a single ensemble. At large enough time separations between the source and sink of the hadron, called the source-sink separation and labelled T_{sink} , the ground state dominates. To ensure the hadron has decayed sufficiently to its ground state, and to estimate the errors of the excited state contamination, T_{sink} is varied and the T_{sink} dependence of the lattice data is studied. More details on this study are given below in Sec. 3.3.

In our calculations for the pion and kaon structure, we use one ensemble of gauge configurations reproducing a pion mass of about 260 MeV. Therefore, we focus on the elimination of excited-state effects. For the calculation of the proton structure, we study three ensembles, all at the physical point and at different volumes and lattices spacing. This allows us to comment on volume and discretization effects, besides excited-state contamination.

CHAPTER 3

CALCULATING HADRON STRUCTURE QUANTITIES ON THE LATTICE

In this chapter, we discuss the details of how to calculate physical quantities using LQCD simulations. We explain which matrix elements are calculated, our procedure to renormalize them, and how they are related to the physical values under study. We also describe the analysis techniques we use to control systematic uncertainties and auto-correlation effects coming from the gauge configurations.

3.1 Matrix Elements

In Section 2.2, we discussed how to calculate N-point functions using Wick's theorem, Eq. (2.60). In order to calculate the quantities under study, the two- and three-point functions are needed. The two-point functions are

$$C_H(x_s) = \langle H(x_s) | H(0) \rangle = \langle \Omega | O_H(x_s) \bar{O}_H(0) | \Omega \rangle , \quad (3.1)$$

where O_H is a hadron interpolation function. The interpolators for the hadrons under study are shown in Table 3.1. As an example, we take the connected diagram, which in a simplistic interpretation gives the valence quark contribution. The connected diagram is the only contribution to the isovector flavo combination, $u - d$, in the nucleon. It is also the major contribution to the mesonic quantities under study.

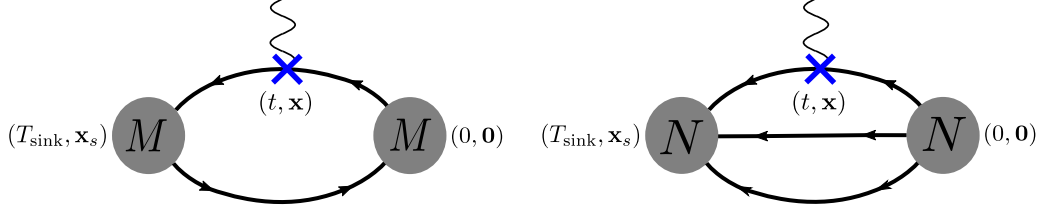


Figure 3.1: Three-point function diagrams for a meson (left) and nucleon (right). The source is at space-time position 0, the current is inserted at $x = (t, \mathbf{x})$, and the sink is at $x_s = (T_{\text{sink}}, \mathbf{x}_s)$. Lines with forward arrows represent the quark propagators and lines with backward facing arrows represent anti-quark propagators.

The study of the disconnected quarks and gluon diagrams is beyond the scope of this thesis. The three-point functions, shown in Figure 3.1, are

$$C_H^\Gamma(x_s, x) = \langle H(x_s) | J^\Gamma(x) | H(0) \rangle = \langle \Omega | O_H(x_s) J^\Gamma(x) \bar{O}_H(0) | \Omega \rangle, \quad (3.2)$$

where J^Γ is the insertion current with structure Γ

$$J^\Gamma(x) = \sum_{N_f} \bar{q}_f(x) \Gamma q_f(x) + \bar{g}(x) \Gamma g(x), \quad (3.3)$$

where q_f is a quark field with flavor f and g is a gluon field. The gamma structure is chosen depending on what physical quantities are being studied and the structures for the quantities in this study are given in Table 3.2.

H	O_H
π^+	$\bar{d}\gamma_5 u$
K^+	$\bar{s}\gamma_5 u$
N	$u(u^T C d)$

Table 3.1: Interpolating fields for the particles under study. The fields u , d , and s are the spinors for up, down, and strange quarks. The transposition T in the nucleon interpolator acts on the Dirac indices and C is charge conjugation operator.

As an example, we show the three-point function calculation for the specific case of a π^+ particle. For a π^+ , the interpolation functions are

$$O_{\pi^+}(x) = \bar{d}(x)\gamma_5 u(x), \quad \bar{O}_{\pi^+}(x) = -\bar{u}(x)\gamma_5 d(x), \quad (3.4)$$

Nucleons (spin 1/2)			
Form factor	Charge/moment	J^Γ	Γ
A_{20}, B_{20}, C_{20}	$\langle x \rangle_q$	$\mathcal{O}_V^{\mu\nu}$	$\gamma^{\{\mu} D^{\nu\}}$
$\tilde{A}_{20}, \tilde{B}_{20}$	$\langle x \rangle_{\Delta q}$	$\mathcal{O}_A^{\mu\nu}$	$\gamma_5 \gamma^{\{\mu} D^{\nu\}}$
$A_{T20}, B_{T20}, C_{T20}, D_{T20}$	$\langle x \rangle_{\delta q}$	$\mathcal{O}_T^{\mu\nu\rho}$	$\sigma^{[\mu\{\nu} D^{\rho\}}$
Mesons (spin 0)			
Form factor	Charge/moment	J^Γ	Γ
F_M	Q_M	\mathcal{O}_V^μ	γ_μ
A_{20}, B_{20}	$\langle x \rangle_q$	$\mathcal{O}_V^{\mu\nu}$	$\gamma^{\{\mu} D^{\nu\}}$
A_{30}, B_{30}	$\langle x^2 \rangle$	$\mathcal{O}_V^{\mu\nu\rho}$	$\gamma^{\{\mu} D^\nu D^\rho\}$
A_{40}, B_{40}	$\langle x^3 \rangle$	$\mathcal{O}_V^{\mu\nu\rho\sigma}$	$\gamma^{\{\mu} D^\nu D^\rho D^\sigma\}$

Table 3.2: The form factors under study, along with their corresponding charge or moment, insertion currents, and insertion current gamma structures. The notation $\{\dots\}$ and $[\dots]$ means symmetrization over enclosed indices and traceless.

so the three-point functions, including only the up quark term of the insertion current, are

$$\begin{aligned}
C_{\pi^+,u}^\Gamma(x_s, x) &= \langle \Omega | \bar{d}(x_s) \gamma_5 u(x_s) \bar{u}(x) \Gamma u(x) (-\bar{u}(0) \gamma_5 d(0)) | \Omega \rangle \\
&= - \langle \Omega | \bar{d}(x_s)_\alpha^a (\gamma_5)_{\alpha\beta} u(x_s)_\beta^a \bar{u}(x)_\gamma^b \Gamma_\gamma \delta u(x)_\delta^b \bar{u}(0)_\epsilon^c (\gamma_5)_{\epsilon\zeta} d(0)_\zeta^c | \Omega \rangle ,
\end{aligned} \tag{3.5}$$

where the Greek and Latin indices are the Dirac and color indices, respectively. Our goal now is what is known as fermion contraction, which is to group quarks and anti-quarks of the same flavors together, since their expectation values are the quark propagators, according to Wick's theorem. Since there are two u and two \bar{u} fields, there are also two different combinations of contractions for this term. One combination is the connected term

$$\begin{aligned}
C_{\pi^+,u}^{\Gamma,\text{conn.}}(x_s, x) &= \langle \Omega | (\gamma_5)_{\epsilon\zeta} d(0)_\zeta^c \bar{d}(x_s)_\alpha^a (\gamma_5)_{\alpha\beta} u(x_s)_\beta^a \bar{u}(x)_\gamma^b \Gamma_\gamma \delta u(x)_\delta^b \bar{u}(0)_\epsilon^c | \Omega \rangle \\
&= \text{tr} [\gamma_5 G_d(0, x_s) \gamma_5 G_u(x_s, x) \Gamma G_u(x, 0)] ,
\end{aligned} \tag{3.6}$$

where $G_q(x, y)$ is a quark propagator origination at y and ending at x . The other quark anti-quark combination is the disconnected term

$$\begin{aligned}
C_{\pi^+,u}^{\Gamma,\text{disconn.}}(x_s, x) &= \langle \Omega | d(0)_\zeta^c \bar{d}(x_s)_\alpha^a (\gamma_5)_{\alpha\beta} u(x_s)_\beta^a \bar{u}(0)_\epsilon^c (\gamma_5)_{\epsilon\zeta} u(x)_\delta^b \bar{u}(x)_\gamma^b \Gamma_\gamma \delta | \Omega \rangle \\
&= \text{tr} [G_d(0, x_s) \gamma_5 G_u(x_s, 0) \gamma_5] \text{tr} [G_u(x, x) \Gamma] .
\end{aligned} \tag{3.7}$$

This can similarly be shown for the other flavor and the gluon terms of the insertion

current, as well as for other hadrons. There is only the disconnected part, however, for the gluon part and the parts for quarks which are not a valence quark of the hadron.

To study the momentum dependence of the matrix elements, we Fourier transform the three-point functions. As an example, we continue with the connected part of the π^+ three-point functions:

$$C_{\pi^+,u}^{\Gamma,\text{conn.}}(T_{\text{sink}}, t; \mathbf{p}', \mathbf{q}) = -\text{tr} \left[\sum_{\mathbf{x}; x_0=t} e^{-i\mathbf{q}\cdot\mathbf{x}} \Gamma G_u(x, 0) \gamma_5 \sum_{\mathbf{x}_s; x_{0s}=T_{\text{sink}}} e^{-i\mathbf{p}'\cdot\mathbf{x}_s} G_d(0, x_s) \gamma_5 G_u(x_s, x) \right], \quad (3.8)$$

where \mathbf{p}' is the final pion momentum and \mathbf{q} is the momentum transferred by the insertion current. We fix the sink at time T_{sink} and momentum \mathbf{p}' by calculating the sequential propagator

$$S(0, x; \mathbf{p}') = \sum_{\mathbf{x}_s; x_{0s}=T_{\text{sink}}} e^{-i\mathbf{p}'\cdot\mathbf{x}_s} G_d(0, x_s) \gamma_5 G_u(x_s, x). \quad (3.9)$$

The three-point functions can now be written as

$$C_{\pi^+,u}^{\Gamma,\text{conn.}}(T_{\text{sink}}, t; \mathbf{p}', \mathbf{q}) = -\text{tr} \left[\sum_{\mathbf{x}; x_0=t} e^{-i\mathbf{q}\cdot\mathbf{x}} \Gamma G_u(x, 0) \gamma_5 S(0, x; \mathbf{p}') \right], \quad (3.10)$$

Fixing the sink has the benefit that we can vary the insertion structure, insertion time, and momentum transfer without the need for additional inversions of the Dirac matrix but with the drawback that for each new source-sink separation or new momentum of the sink, we need another inversion to calculate the sequential propagator. The insertion position can also be fixed which has the opposite benefits and drawbacks but, in this study, we only use the fixed sink method.

The nucleon two and three-point functions are calculated with projector Γ which acts on the spin indices and is needed to disentangle the form factors. An additional sequential inversion is needed for each projector. The nucleon correlators are

calculated as

$$C_N(t; \mathbf{p}) = \sum_{\mathbf{x}; x_0=t} e^{-i\mathbf{p}\cdot\mathbf{x}} \text{tr} [\Gamma_0 \langle O_N(t, \mathbf{x}) \bar{O}_N(0, \mathbf{0}) \rangle] , \quad (3.11)$$

$$C_N^{J, \text{conn.}}(\Gamma; T_{\text{sink}}, t; \mathbf{q}, \mathbf{p}') = \sum_{\mathbf{x}; x_0=t} \sum_{\mathbf{x}_s; x_{0s}=T_{\text{sink}}} \text{tr} [\Gamma \langle O_N(T_{\text{sink}}, \mathbf{x}_s) J(t, \mathbf{x}) \bar{O}_N(0, \mathbf{0}) \rangle] \quad (3.12)$$

where the full contraction has not been written out in full so that there is no confusion between the projectors and the gamma structure of the insertion current J . The four projectors used in our calculations are the unpolarized projector $\Gamma_0 = \frac{1}{2}(1 + \gamma_0)$ and the three polarized projectors $\Gamma_k = i\gamma_5\gamma_k\Gamma_0$ for $k = 1, 2, 3$.

3.2 Renormalization

The discretization of space-time is used as a regulator of UV divergences, so a multiplicative renormalization factor needs to be used to relate the bare matrix elements calculated on the lattice to their physical values. The renormalization functions needed for this study are the ones corresponding to $\mathcal{O}^{\mu\mu}$, $\mathcal{O}^{\mu\nu\rho}$ ($\mu \neq \nu \neq \rho \neq \mu$), and $\mathcal{O}^{\mu\nu\rho\sigma}$ ($\mu \neq \nu \neq \rho \neq \sigma \neq \mu$) which we refer to as Z_{vD} , Z_{vDD} , and Z_{vDDD} , respectively. The renormalization function Z_{v} for \mathcal{O}^{μ} is equal to one because we use the conserved vector currents. We calculate the renormalization functions non-perturbatively using the Rome-Southampton method (RI' scheme) [64] then, using continuum perturbation theory, convert to the pion mass independent $\overline{\text{MS}}$ scheme and evolve at a renormalization scale of $\bar{\mu} = 2$ GeV.

The renormalization functions are determined in the RI' scheme by the conditions

$$Z_q = \frac{1}{12} \text{Tr} [(S^L(p))^{-1} S^{\text{Born}}(p)] \Big|_{p^2=\mu_0^2} , \quad (3.13)$$

$$Z_q^{-1} Z_{\mathcal{O}} \frac{1}{12} \text{Tr} [\Gamma_{\mathcal{O}}^L(p) (\Gamma_{\mathcal{O}}^{\text{Born}}(p))^{-1}] \Big|_{p^2=\mu_0^2} = 1 , \quad (3.14)$$

where p is the momentum of the vertex function, set to the RI' renormalization scale, μ_0 , S^{Born} ($\Gamma_{\mathcal{O}}^{\text{Born}}$) is the tree-level value of the fermion propagator (operator), and the trace is taken over spin and color. We use the momentum source method [65] which has been successfully used for twisted mass fermions by other studies [66, 67, 68]. Momentum sources give a high signal with just a few gauge configurations at the

cost of requiring multiple momentum sources to be calculated as opposed to a single point source. In these results, we use 10 configurations which already results in per mil accuracy. Based on empirical arguments [69], we use momenta that have the same spatial components which suppresses non-Lorentz invariant contributions ($\sum_i p_i^4 / (\sum_i p_i^2)^2 < 0.3$) and reduces discretization effects. The chosen momenta are:

$$(ap) \equiv 2\pi \left(\frac{n_t}{L_t} + \frac{1}{2L_t}, \frac{n_x}{L_s}, \frac{n_x}{L_s}, \frac{n_x}{L_s} \right), \quad n_t \in [2, 9], \quad n_x \in [2, 5], \quad (ap)^2 \in [0.9 - 6.7], \quad (3.15)$$

where L_t (L_s) is the temporal (spatial) extent of the lattice.

$\beta = 1.726, \quad a = 0.093 \text{ fm}$		
$a\mu$	aM_{PS}	lattice size
0.0060	0.1680	$24^3 \times 48$
0.0080	0.1916	$24^3 \times 48$
0.0100	0.2129	$24^3 \times 48$
0.0115	0.2293	$24^3 \times 48$
0.0130	0.2432	$24^3 \times 48$

Table 3.3: Parameters for the $N_f = 4$ ensembles used for the renormalization functions.

We perform a chiral extrapolation so as to calculate the renormalization functions at the chiral limit. To achieve this, the renormalization functions are calculated on five ensembles, whose parameters are given in Table 3.3, with varied pion masses. The quark masses of each ensemble are all degenerate ($N_f = 4$) and all ensembles have the same β value as the ensemble used to calculate the matrix elements. We then fit the renormalization functions quadratically with respect to the pion mass and extrapolate to the chiral limit. We use 23 momentum sources to calculate values of the initial RI' scale μ_0 ranging from ($(a\mu)^2 \in [0.9 - 6.7]$) on each $N_f = 4$ ensemble. Tables 3.4 and 3.5 show Z_{vD} and Z_{vDD} , respectively, for each ensemble and two representative renormalization scales ($(a\mu_0)^2 = 2, 4$) and we find the pion mass dependence of both renormalization factors to be negligible. Using continuum perturbation theory, we define an intermediate Renormalization Group Invariant scheme to convert and evolved to $\overline{\text{MS}}(2\text{GeV})$.

Z_{vD} , Z_{vDD} , and Z_{vDDD} are plotted in Fig. 3.2 in the RI' and $\overline{\text{MS}}$ schemes as a function of the initial RI' renormalization scale, μ_0 . $Z_{\mathcal{O}}^{\overline{\text{MS}}}$ are plotted at $\mu = 2$

aM_{PS}	$Z_{vD}^{RI'}$	
	$(a\mu_0)^2=2$	$(a\mu_0)^2=4$
0.1680	1.1762(2)	1.1043(1)
0.1916	1.1770(3)	1.1045(2)
0.2129	1.1773(2)	1.1046(1)
0.2293	1.1782(2)	1.1048(1)
0.2432	1.1779(2)	1.1047(1)

Table 3.4: Pion mass dependence of the renormalization function Z_{vD} . The first column is the pion mass (in lattice units) for the ensemble, the second (third) is the renormalization function at scale $(a\mu_0)^2=2$ ($(a\mu_0)^2=4$). The number in the parenthesis is the statistical error.

aM_{PS}	$Z_{vDD}^{RI'}$	
	$(a\mu_0)^2=2$	$(a\mu_0)^2=4$
0.1680	1.4870(4)	1.3722(2)
0.1916	1.4890(5)	1.3733(2)
0.2129	1.4888(5)	1.3732(3)
0.2293	1.4922(4)	1.3751(2)
0.2432	1.4914(6)	1.3748(3)

Table 3.5: Table similar to Table (3.4) but for Z_{vDD} .

GeV and the purely non-perturbative data has a small residual dependence on the initial scale μ_0 they were evolved from. In order to remove this dependence, we calculate finite lattice effects to one loop in perturbation theory and to all orders in the lattice spacing, $\mathcal{O}(g^2 a^\infty)$, and subtract them from the non-perturbative estimates of Z_{vD} as explained in Refs. [70, 68]. Though this improvement is not yet available for two- and three-derivative operators, we subtract the $\mathcal{O}(g^2 a^\infty)$ artifacts from Z_q in the renormalization conditions in Eqs. (3.13) and (3.14) for Z_{vDD} and Z_{vDDD} as a partial improvement. After removing these artifacts, the slopes are reduced for both renormalization functions, which shows artifact-subtraction procedure worked effectively.

We fit the \overline{MS} values with respect to $(a\mu_0)^2$ to a line so that we can remove any remaining dependence on the initial scale μ_0 which can be caused by finite- a effects and/or truncation of the conversion factor (to three loops in perturbation theory).

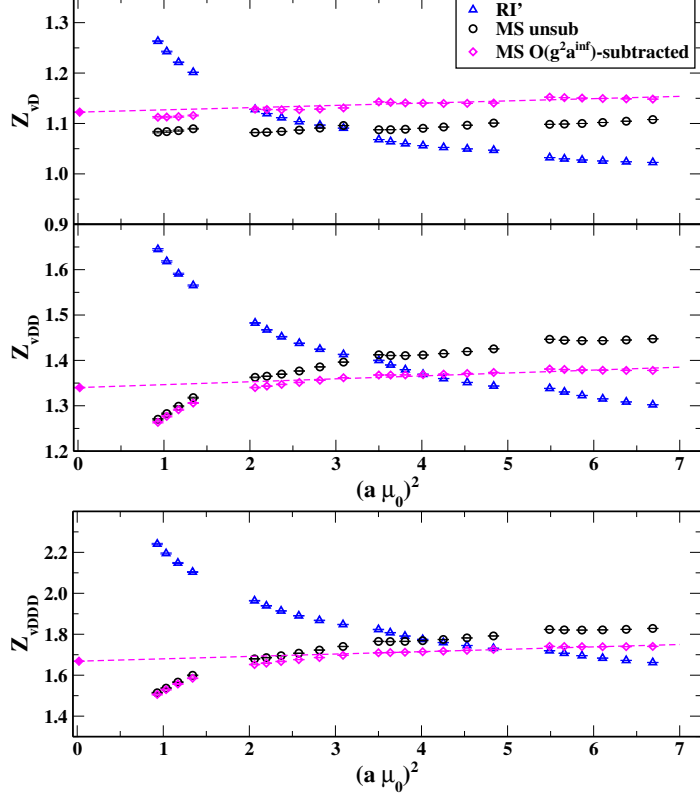


Figure 3.2: Chirally extrapolated results for Z_{vD} (top), Z_{vDD} (middle), and Z_{vDDD} (bottom). Blue triangles correspond to RI' scheme, black circles to $\overline{\text{MS}}$ scheme and magenta diamonds to the subtracted results in the $\overline{\text{MS}}$ scheme. The data are plotted as a function of the initial RI' scale $(a\mu_0)^2$. The dashed line corresponds to the fit of Eq. (3.16), and a filled magenta diamond represent our final value for \mathcal{Z}_{vD} , \mathcal{Z}_{vDD} , and \mathcal{Z}_{vDDD} .

The Ansatz we use is

$$Z_{\mathcal{O}}(ap) = \mathcal{Z}_{\mathcal{O}} + Z_{\mathcal{O}}^{(1)} \cdot (a\mu_0)^2, \quad (3.16)$$

where $\mathcal{Z}_{\mathcal{O}}$ is the final value of the renormalization function for operator \mathcal{O} . We obtain $\mathcal{Z}_{vD} = 1.123(1)(5)$, $\mathcal{Z}_{vDD} = 1.340(1)(15)$, and $\mathcal{Z}_{vDDD} = 1.668(1)(26)$, where the numbers in the first and second parentheses are the statistical and systematic errors, respectively. The systematic errors are caused by the $(a\mu_0)^2 \rightarrow 0$ extrapolation and are estimated by varying the $(a\mu_0)^2$ fit range. The final values use the fit interval $(a\mu_0)^2 \in [2 - 7]$ and the reported uncertainty is the difference between the final value and the value obtained at $(a\mu_0)^2 \in [4 - 7]$.

3.3 Controlling Excited State Systematic Uncertainties

As discussed above in Sec. 2.5, excited states of the hadron under study can contaminate the ground state results. To ensure suppression of excited states, the source-sink time separation is varied and we extract the matrix elements at sufficiently large sink times ($T_{\text{sink}}/a \gg 1$) where the ground state of the hadron gives the dominant contribution to the correlators. For the three-point functions, the insertion times also need to be sufficiently far from the source and sink ($t/a \gg 1$ and $(T_{\text{sink}} - t)/a \gg 1$).

In addition, multiple fitting routines can be employed to study the convergence to the ground state and the significance of excited-states contributions. We also used this study in the rest frame, as a guidance for the T_{sink} values to be employed in the boosted frame. We use two fits in this study, the plateau fit and the two-state fit, which are described below.

Plateau method: The first method relies on a single-state fit applied to an appropriate ratio of three-point and two-point functions. The ratio is fitted to a constant with respect to t for each T_{sink} separately and in a region where the ratio becomes time independent (plateau region). We then look for the extracted plateau values to converge as T_{sink} increases. The process is slightly different when calculating the effective mass, which only depends on two-point functions and so has no insertion current. The effective mass is instead fit to a constant with respect to T_{sink} .

Two-state fit: The second method is a two-state fit which takes into account the contribution from the first-excited state. The two-state fit Ansatz is given by

$$C_N(t) = c_0 e^{-E_0 t} + c_1 e^{-E_1 t}, \quad (3.17)$$

for the nucleon two-point functions,

$$C_M(t; Q^2) = c_0(Q^2) \left(e^{-E_0(Q^2)t} + e^{-E_0(Q^2)(T-t)} \right) + c_1(Q^2) \left(e^{-E_1(Q^2)t} + e^{-E_1(Q^2)(T-t)} \right), \quad (3.18)$$

for the meson two-point functions, and

$$C_H^\Gamma(t, T_{\text{sink}}) = A_{00}e^{-E_0 T_{\text{sink}}} + A_{01}e^{-E_0(T_{\text{sink}}-t)-E_1 t} + A_{10}e^{-E_1(T_{\text{sink}}-t)-E_0 t} + A_{11}e^{-E_1 T_{\text{sink}}}, \quad (3.19)$$

for the zero momentum transfer three-point functions. For the finite momentum transfer three-point functions in the rest frame, the two-state fit Ansatz is

$$C_H^\Gamma(t, T_{\text{sink}}; Q^2) = A_{00}e^{-m_0(T_{\text{sink}}-t)-E_0(Q^2)t} + A_{01}e^{-m_0(T_{\text{sink}}-t)-E_1(Q^2)t} + A_{10}e^{-m_1(T_{\text{sink}}-t)-E_0(Q^2)t} + A_{11}e^{-m_1(T_{\text{sink}}-t)-E_1(Q^2)t}. \quad (3.20)$$

We do not fit finite momentum transfer three-point functions in the boosted frames because the errors are too high to obtain a reliable fit. We first fit the zero momentum transfer two-point functions using the amplitudes c_0 and c_1 and the ground and first excited state energies E_0 , E_1 as fitting parameters. We then fit the non-zero momentum transfer two-point functions using the dispersion relation $E_0 = \sqrt{m_{\text{eff}}^2 + Q^2}$ and m_{eff} calculated from the plateau fit to increase the reliability of the fit. Lastly, we fit the zero momentum transfer three-point functions. Given the large number of parameters in Eqs. (3.19) and (3.20), we use E_0 and E_1 extracted from the two-state fit of the two-point functions. Therefore, the actual free parameters are the amplitudes A_{00} , A_{01} , and A_{11} ($A_{01} = A_{10}$ for zero momentum transfer). The desired matrix element of the ground state is extracted via

$$\langle H(0) | \mathcal{O}^\Gamma | H(0) \rangle = \frac{A_{00}}{c_0}. \quad (3.21)$$

3.4 Decomposition

Once we have calculated and fit the two- and three-point functions, we calculate the ratio

$$R_H(J^\Gamma; T_{\text{sink}}, t; \mathbf{p}_f, \mathbf{q} = \mathbf{p}_f - \mathbf{p}_i) = \frac{C_H^\Gamma(T_{\text{sink}}, t; \mathbf{p}_f, \mathbf{q})}{C_H(t; \mathbf{p}_f^2)} \times \sqrt{\frac{C_H(T_{\text{sink}} - t; \mathbf{p}_i^2) C_H(t; \mathbf{p}_f^2) C_H(; T_{\text{sink}} \mathbf{p}_f^2)}{C_H(T_{\text{sink}} - t; \mathbf{p}_f^2) C_H(t; \mathbf{p}_i^2) C_H(T_{\text{sink}}; \mathbf{p}_i^2)}}, \quad (3.22)$$

For the mesons, instead of using the two-point functions directly, we use the modified two-point functions

$$C_M(t; \mathbf{p}^2) = c_0 e^{-E_0(\mathbf{p}^2)t} \quad (3.23)$$

which removes the T_{sink} dependence from the plateau value of the ratio caused by the backward propagating meson. This ratio is constructed to remove unknown factors common to the two- and three-point functions while optimizing for t values with lower errors. Note that in the case of zero momentum transfer $\mathbf{q} = 0$, the factor under the square root is equal to 1.

In the limit of large time separations the ratio becomes time-independent, that is

$$R_H(J^\Gamma; T_{\text{sink}}, t; \mathbf{p}_f, \mathbf{q}) \xrightarrow[t/a \gg 1]{(T_{\text{sink}} - t)/a \gg 1} \Pi_H(J^\Gamma; \mathbf{p}_f, \mathbf{q}). \quad (3.24)$$

The time independent ratios are related to the form factors by different kinematic factors depending on how they decompose. For the four meson form factors under study, the decompositions are

$$\Pi_M(\mathcal{O}_V^\mu; \mathbf{p}_f, \mathbf{q}) = C (p_f + p_i)_\mu A_{10}(Q^2), \quad (3.25)$$

$$\Pi_M(\mathcal{O}_V^{\mu\nu}; \mathbf{p}_f, \mathbf{q}) = C [2P^{\{\mu} P^{\nu\}} A_{20}(Q^2) + 2\Delta^{\{\mu} \Delta^{\nu\}} B_{20}(Q^2)] , \quad (3.26)$$

$$\Pi_M(\mathcal{O}_V^{\mu\nu\rho}; \mathbf{p}_f, \mathbf{q}) = C [2i P^{\{\mu} P^\nu P^{\rho\}} A_{30}(Q^2) + 2i \Delta^{\{\mu} \Delta^\nu P^{\rho\}} B_{30}(Q^2)] , \quad (3.27)$$

$$\begin{aligned} \Pi_M(\mathcal{O}_V^{\mu\nu\rho\sigma}; \mathbf{p}_f, \mathbf{q}) = C [& -2 P^{\{\mu} P^\nu P^\rho P^{\sigma\}} A_{40}(Q^2) - 2 \Delta^{\{\mu} \Delta^\nu P^\rho P^{\sigma\}} B_{40}(Q^2) \\ & - 2 \Delta^{\{\mu} \Delta^\nu \Delta^\rho \Delta^{\sigma\}} C_{40}(Q^2)] , \end{aligned} \quad (3.28)$$

where P is the average of the initial and final momenta of the hadron, $P = (p + p')/2$, Δ their difference, $\Delta = p' - p$, and Q^2 the momentum transferred squared. K is a kinematic factors related to the two-point function normalization and the overlap coefficient between the ground state and the interpolating field and, therefore, depends on the momentum transferred as well as the momentum frame. The convention we use for the mesons is

$$\langle 0 | J_\pi | \pi \rangle = \frac{1}{\sqrt{2E}} Z_\pi , \quad (3.29)$$

so each two-point function in Eq. (3.22) comes with a factor of $1/(2E)$ and each three-point function comes with a factor of $1/(4E_i E_f)$. The result is that, for a general frame,

$$C = \frac{1}{\sqrt{4E(p_i)E(p_f)}}, \quad (3.30)$$

where $E(p)=\sqrt{m_M^2+\mathbf{p}^2}$ is the energy at momentum \mathbf{p} and m_M is the mass of meson M . Therefore, C only depends on \mathbf{p}^2 and not on the spatial directions of the momentum. For the nucleon moments under study, the decompositions at zero momentum transfer and zero momentum boost are

$$\Pi_N(\mathcal{O}_V^{\mu\nu};\Gamma_j)=\frac{1}{2}\langle x\rangle_q\gamma^{\{\mu}p^{\nu\}} \quad (3.31)$$

$$\Pi_N(\mathcal{O}_A^{\mu\nu};\Gamma_j)=\frac{i}{2}\langle x\rangle_{\Delta q}\gamma^{\{\mu}p^{\nu\}}\gamma_5 \quad (3.32)$$

$$\Pi_N(\mathcal{O}_T^{\mu\nu\rho};\Gamma_j)=\frac{i}{2}\langle x\rangle_{\delta q}\sigma^{[\mu\nu]}p^\rho, \quad (3.33)$$

where Γ_j is the projector acting on the spin indices.

For Eqs. (3.25-3.28), we only use cases where all of the indices are different. This is to avoid unknown mixing between operators which complicates the analysis. For $\langle x^2\rangle$ and $\langle x^3\rangle$, the boosted frames are needed because the factor which multiplies it is zero for the case when $Q^2=0$ and all indices are different. We do not need the boosted frames for the electric form factors or $\langle x\rangle$ but we use them to test the signal in the boosted frames and to help select which source-sink separations to use. We find the proper combinations of projector and current indices for each moment by using

$$\Pi_N(\mathcal{O},\Gamma_j)\equiv C\text{tr}[\Gamma_j\Lambda(p')\mathcal{F}^{[\mathcal{O}]}\Lambda(p)], \quad (3.34)$$

where $\Lambda(p)$ is the renormalization

$$\Lambda(p)=\frac{-i\not{p}+m}{2m}, \quad (3.35)$$

and, for the nucleon and general momentum,

$$C=\frac{2m^2}{\sqrt{E'E(E'+m)(E+m)}}. \quad (3.36)$$

$\mathcal{F}^{[\mathcal{O}]}$ is defined via

$$F^{[\mathcal{O}]}=\bar{u}(p')[\mathcal{F}^{[\mathcal{O}]}]u(p), \quad (3.37)$$

where $F^{[\mathcal{O}]}$ is the insertion current.

In Appendix A, we give the decomposition equations for each contributing insertion current and for the four combinations of general/rest frames and general/zero momentum transfers for the mesons. For the nucleon, only the decomposition in the

rest frame and at zero momentum transfer is given, since that is all that is needed for the moments under study.

3.5 Jackknife Resampling

Because the gauge configurations are generated using a Markov chain, there are auto-correlations between configurations. To reduce the auto-correlation, multiple trajectories are calculated before writing a configuration so that the configurations which are used for simulations are farther apart. The configurations might still be correlated to each other, however, so we use a resampling technique called jackknife. Auto-correlations can make the variance between configuration artificially low so jackknifing gives a better estimation of statistical errors by calculating the variance between results which are averaged over a subset of configurations.

First, we group the N_C correlators together into N_B different sets, called bins. The number of configurations in each bin is called the bin size S_B and is related to the number of configurations and bins by

$$S_B = \frac{N_C}{N_B}. \quad (3.38)$$

The next step is to average the correlators over configurations, omitting each bin in turn so that we now have N_B resampled correlators G_b . This can be stated mathematically as

$$G_b = \frac{1}{N_C - S_B} \sum_{i=0, i \notin B_b}^{N_C-1} G_i, \quad (3.39)$$

$$B_b : [S_B b, S_B b + S_B), \quad b = 0, 1, 2, \dots, N_B - 1. \quad (3.40)$$

Then, we calculate our quantity of interest Q_b (e.g., effective mass, quark momentum fraction, etc.) for each bin from the resampled correlators. Our result is then the

mean over bins with the associated error $\bar{Q} \pm dQ$,

$$\bar{Q} = \frac{1}{N_B} \sum_{b=0}^{N_B-1} Q_b, \quad (3.41)$$

$$dQ = \sqrt{\frac{N_B - 1}{N_B} \sum_{b=0}^{N_B-1} (\bar{Q} - Q_b)^2}. \quad (3.42)$$

Note that all steps for calculating Q_b should be done for each bin before calculating \bar{Q} so that there is no need to propagate the uncertainty.

CHAPTER 4

NUCLEON STRUCTURE

Understanding the nucleons and their constituent particles is one of the major goals of hadron physics, since they are the most common hadrons in matter. Calculating nucleon structure quantities from first principles is highly desirable, which is why nucleons have been the focus of most LQCD hadron structure studies. As explained in Section 1.2, we calculate the Mellin moments of the PDFs on the lattice. For the nucleon, we calculate the quark momentum fraction $\langle x \rangle_{u-d}$, the helicity $\langle x \rangle_{\Delta u-\Delta d}$, and the transversity $\langle x \rangle_{\delta u-\delta d}$. This work has been published in Physical Review D [2], along with results on the generalized form factor of the vector and axial one derivative operators.

4.1 Calculation Details

The nucleon quark momentum fraction is calculated using three ensembles whose parameters are given in Table 4.1. Two of the ensembles [71], cA2.09.48 and cA2.09.64, have the same parameters except volume size, allowing us to test finite volume effects on $\langle x \rangle$ for the unpolarized, helicity, and transversity cases. Those two are generated with two mass degenerate quarks ($N_f = 2$). The other ensemble [72] also includes the charm and strange quarks in the sea ($N_f = 2 + 1 + 1$) so that we can study quenching effects and has a slightly smaller lattice spacing. All three ensembles use quarks around the physical point and were generated by the Extended Twisted Mass Collaboration (ETMC).

The correlators are calculated at different statistics, as is listed in Table 4.2. On each configurations, we calculate N_{srcs} random source positions. The correlators are

Ensemble	β	a [fm]	volume $L^3 \times T$	N_f	m_π [MeV]
cA2.09.48	2.1	0.938(3)(1)	$48^3 \times 96$	2	130.3(4)(2)
cA2.09.64	2.1	0.938(3)(1)	$64^3 \times 128$	2	130.3(4)(2)
cB211.072.64	1.778	0.801(4)	$64^3 \times 128$	2 + 1 + 1	139.3(7)

Table 4.1: Parameters of the ensembles used for the nucleon $\langle x \rangle$ calculation. The process of determining the lattice spacing from the nucleon mass is outlined in [19]. The systematic errors in the lattice spacing of cA2.09.48 and cA2.09.64 comes from an interpolation to the physical pion mass using one-loop chiral perturbation theory because the pion mass was underestimated [19]. The errors in m_π are due to the values being calculated from the lattice spacing.

calculated using N_{confs} configurations so the total statistics are $N_{\text{total}} = N_{\text{confs}} \times N_{\text{srCs}}$. The two-point functions are calculated at higher statistics for ensembles cB211.072.64 and cA2.09.48 because they were used in a separate study of the disconnected contribution, which is beyond the scope of this thesis. All of the nucleon three-point functions are calculated using the unpolarized projector Γ_0 and all of the three-point functions at source-sink separations not labeled with an asterisk in Table 4.2 are calculated with the three polarized projectors Γ_k . As discussed in Section 3.4, different projectors are needed to disentangle the form factors. We focus on the isovector contribution $u - d$ because the contribution from the disconnected diagrams, which is costly to calculate, are the same for both flavors and, up to cutoff effects, cancel one another in the subtraction.

4.2 Systematic Uncertainties

We plot the three- to two-point function ratios for the three moments in Figures 4.1, 4.2, and 4.3 for the cB211.072.64, cA2.09.64, and cA2.09.48 ensembles, respectively. That is, we plot the ratio R of Eq. 3.22 multiplied by all kinematic factors and the renormalization function according to the decomposition in Eqs. (A.1-A.4). The ratios leading to $\langle x \rangle_{u-d}$, $\langle x \rangle_{\Delta u - \Delta d}$, and $\langle x \rangle_{\delta u - \delta d}$ are plotted in the top, middle and bottom rows, respectively.

We plot the ratios for each value of T_{sink} and as a function of t so that we can study the excited-state effects. In the center panels, we compare the plateau values with the two-state fit. The gray curve is not a fit of the plateau data, but rather calculated

T_{sink}/a	N_{confs}	N_{srCs}	N_{total}
cA211.072.64			
Three-point correlators			
8	750	1	750
10	750	2	1,500
12	750	4	3,000
14	750	6	4,500
16	750	16	12,000
18	750	48	36,000
20	750	64	48,000
Two-point correlators			
N/A	750	264	198,000
cA2.09.64			
Three-point correlators			
12	333	16	5,328
14	515	16	8,240
16	515	32	16,480
Two-point correlators			
N/A	512	32	16,480
cA2.09.48			
Three-point correlators			
10,12,14	578	16	9,248
16*	530	88	46,640
18*	725	88	63,800
Two-point correlators			
N/A	2153	100	215,300

Table 4.2: Statistics used for each nucleon correlator. For cA2.09.48, an asterisk is used to label which source-sink separations were calculated using only the Γ_0 projector.

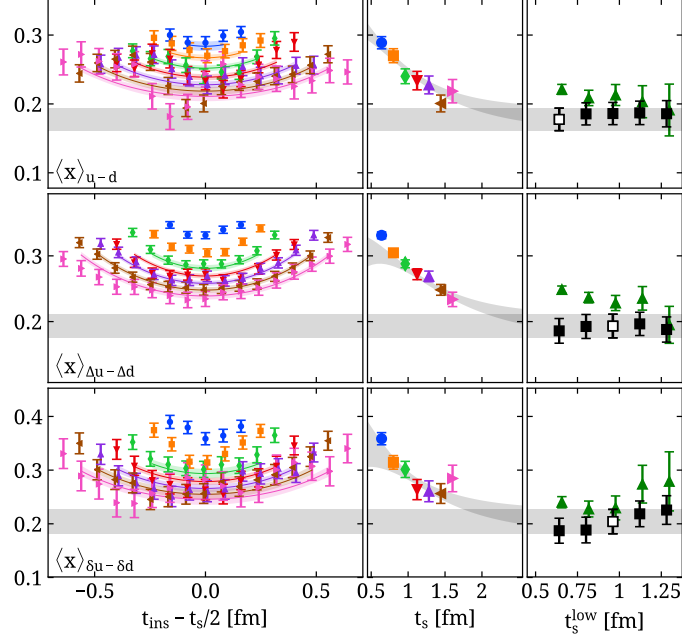


Figure 4.1: Ensemble cB211.072.64 results for $\langle x \rangle_{u-d}$ (top row), $\langle x \rangle_{\Delta u - \Delta d}$ (middle row), $\langle x \rangle_{\delta u - \delta d}$ (bottom row). In the left column, we plot the ratios in Eq. (3.22) multiplied by the renormalization and kinematic factors according to Eqs. (A.1-A.4) for $T_{\text{sink}}/a = 8, 10, 12, 14, 16, 18, 20$ as blue circles, orange squares, green diamonds, red downward-pointing triangles, purple upward-pointing triangles, brown left-pointing triangles, and magenta right-pointing triangles, respectively. The blue, orange, and green curves are functions calculated from the two-state fit parameters for each value of T_{sink} . In the center panel, the plateau fits for each T_{sink} are plotted and the gray curve is the function obtained from the two-state fit parameters using $t_i = T_{\text{sink}}/2$ band. In the right panel, the black squares are the two-state fit values using the ratio from Eq. (3.21) plotted as function of the lowest T_{sink} used in the fit. The open square is the value selected as the result and the gray bands show its error range across all three panels. The green triangles are calculated using the summation method, which is explained in [2]

from the two-state fit parameters as a function of T_{sink} and at constant $t = T_{\text{sink}}/2$. From these curves, it can be seen that higher source-sink separations are needed to obtain convergence between the plateau fits and the two-state fit. Since increased T_{sink} also means increased noise in the ratios, the statistics of the largest separation of 1.6 fm has $64\times$ more statistics than the smallest separation of 0.64 fm while the errors increased by $\sim 2.5\times$. Therefore, $\mathcal{O}(100)$ more statistics would be needed to calculate separations in the region where the curve agrees with the two-state fit. This

amount of statistics would be prohibitively expensive so we choose the two-state fit as our final values.

In order to test the dependence of the two state fit on the range of T_{sink} included in the fit, we fit the data using different ranges for T_{sink} . The results of this test are shown in the right panels, where we plot the results calculated using the ratio in Eq. (3.21) as a function of the lowest value of T_{sink} included in the fit. We find very little dependence on the fit range. For cB211.072.64, we use as our final result the fit range which agrees with the summation method, which is obtain by the process described in [2], and for the other ensembles, since there are fewer separations, the fit range including all separations is chosen. These selections are shown as open squares in the figures.

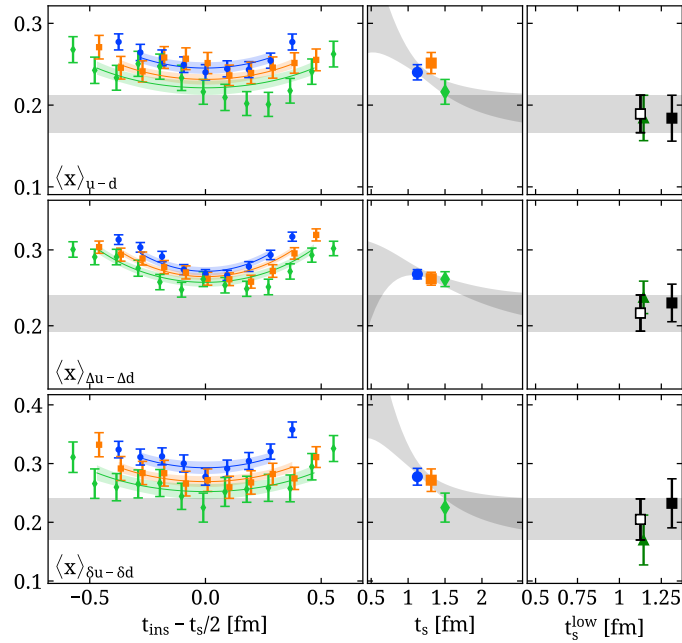


Figure 4.2: Ensemble cA2.09.64 results with notation similar to Figure 4.1. The different source-separations, $T_{\text{sink}}/a = 14, 16, 18$, are shown as blue squares for, orange squares, and green diamonds, respectively.

The final results for all three moments and three ensembles are collected in Table 4.3. We find agreement between the three ensembles within the reported errors except in the case of $\langle x \rangle_{\Delta u - \Delta d}$, for which the results of the $N_f = 2 + 1 + 1$ ensemble is within 1.5σ of the results of the $N_f = 2$ ensembles. The only difference between the two $N_f = 2$ ensembles are their volumes with $m_\pi L = 2.98$ for the smaller and

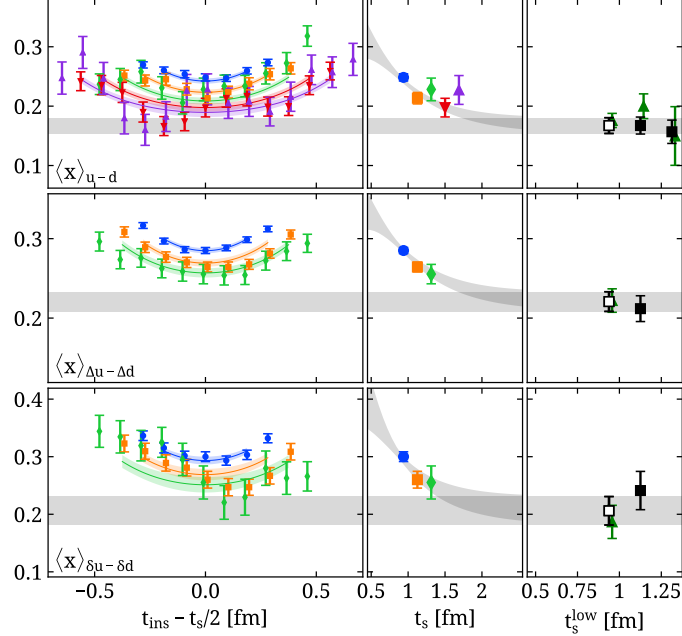


Figure 4.3: Ensemble cA2.09.48 results with notation similar to 4.1. The different source-separations, $T_{\text{sink}}/a = 10, 12, 14, 16, 18$, are shown as blue squares for, orange squares, green diamonds, red downward-pointing triangles, and purple upward-pointing triangles respectively.

$N_f = 3.97$ for the larger. Since the moments obtained from the two ensembles agree, we conclude that any finite volume systematic effects are less than the statistical errors. Ensemble cB211.072.64 has the value $m_\pi L = 3.62$ which is between the other two, so we also do not expect any finite volume systematics to affect those results. The lattice spacing of cB211.072.64 is smaller than the other two and also include the strange and charm quarks in the sea. Since the final results differ at most by 1.5σ , we find that the quenching and cutoff effects, at least between the range of the two lattice spacings, are also lower than the statistical uncertainties.

Ensemble	$\langle x \rangle_{u-d}$	$\langle x \rangle_{\Delta u - \Delta d}$	$\langle x \rangle_{\delta u - \delta d}$
cB211.072.64	0.178(16)	0.193(18)	0.204(23)
cA2.09.48	0.167(13)	0.221(12)	0.206(25)
cA2.09.64	0.189(23)	0.217(24)	0.205(35)

Table 4.3: Final results for the three nucleon moments and three ensembles. In all cases, the results of the two-state fit are chosen as the final results.

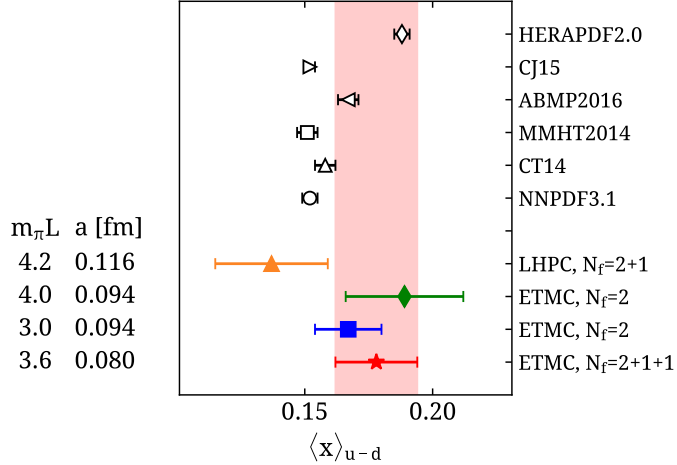


Figure 4.4: A comparison of $\langle x \rangle_{u-d}^N$ calculated by different groups and methods. Our results are shown as a red star (cB211.072.64), a green diamond (cA2.09.64), and blue square (cA2.09.48). The results from the lattice calculation of the the LHPC group [3] is plotted as an orange triangle. The open and uncolored shapes are phenomenological calculations of global fits [4, 5, 6, 7, 8, 9].

4.3 Comparison With Other Calculations

In this section, we compare our results with other studies, including phenomenological fits and other lattice calculations at or near the physical point, starting with $\langle x \rangle_{u-d}$. Many studies have calculated the isovector momentum fraction at higher than physical pion masses and a review of these are found in [23]. Figure 4.4 compares our results with the results from other studies. The red star, green diamond, and blue square are our final results for the ensembles cB211.072.64, cA2.09.64, and cA2.09.48, respectively. The orange triangle is another lattice calculation by the LHPC group [3], who use an ensemble with $N_f = 2 + 1$ HEX-smearred clover improved fermions, pion mass $m_\pi = 149$ MeV, lattice volume $48^3 \times 48$, and lattice spacing $a = 0.116$ fm. The open, uncolored shapes are phenomenological results [4, 5, 6, 7, 8, 9]. We choose the results of cB211.072.64 as our result to compare to other studies, since it has more statistics and a wider range of source-sink separations. Two of the phenomenological fit results are in agreement with ours within the reported uncertainty while the other results are all within 1.5σ .

In Figure 4.5, we compare our results for $\langle x \rangle_{\Delta u - \Delta d}$ to the phenomenological results of [73, 27, 74, 75]. We again select our results from cB211.072.64 as our final result

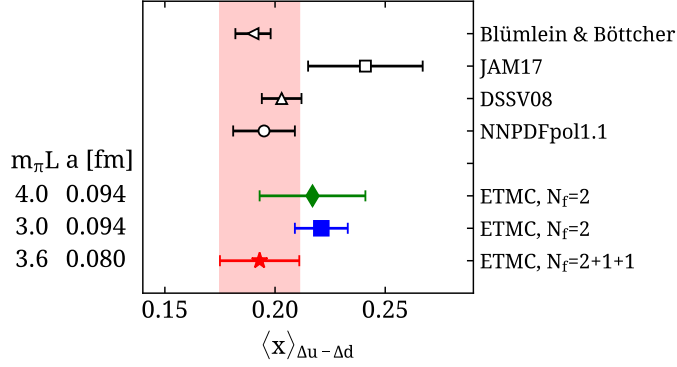


Figure 4.5: A comparison of $\langle x \rangle_{u-d}^N$ calculated by different groups and methods. Our results are shown by a green triangle and its errors by a green band. The other lattice calculations are plotted as an orange triangle for the LHPC group and a blue square and red star for two other ETMC ensembles. The open and uncolored shapes are phenomenological calculations of global fits.

due to the other ensembles only having three source-sink separations. We find good agreement between our results and those from phenomenology. We compare our $\langle x \rangle_{\delta u - \delta d}$ results for the three ensembles in Figure 4.6. There have been no other calculations of the tensor moment from either lattice or phenomenology so this study offers a valuable prediction. As we did with the previous two moments, we select the cB211.072.64 calculation as our final results and we collect our final nucleon results for each moment here:

$$\langle x \rangle_{u-d} = 0.178(16) \quad (4.1)$$

$$\langle x \rangle_{\Delta u - \Delta d} = 0.193(18) \quad (4.2)$$

$$\langle x \rangle_{\delta u - \delta d} = 0.204(23). \quad (4.3)$$

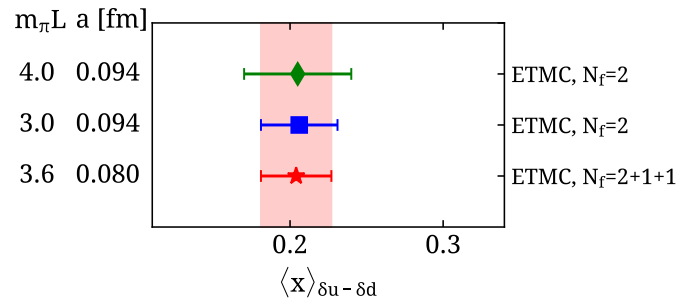


Figure 4.6: A comparison of $\langle x \rangle_{u-d}^N$ calculated by different groups and methods. Our results are shown by a green triangle and its errors by a green band. The other lattice calculations are plotted as an orange triangle for the LHPC group and a blue square and red star for two other ETMC ensembles. The open and uncolored shapes are phenomenological calculations of global fits.

CHAPTER 5

MESON STRUCTURE

Studying the pion and kaon is useful for understanding QCD dynamics at hadronic scales which answer important questions in hadronic physics. Studying the differences between nucleons on the one hand and pions and kaons on the other can give insights into how hadron masses are generated. For example, the nucleons mass is on the order of 1 GeV in the chiral limit whereas the pion and kaon are massless. Pions can explain why the proton sea has more \bar{d} than \bar{u} anti-quarks [76, 77, 78, 79]. Phenomenological studies suggest that the strange contribution to the kaon form factors decreases faster than the up quark contribution with increasing momentum transfer. This interesting phenomenon has been interpreted as a consequence of confinement [80].

Additionally, the octet of Nambu–Goldstone bosons are made up of the pion and kaon, along with the eta meson which is not studied in this thesis. The Nambu–Goldstone bosons are unique among hadrons because their masses vanish in the chiral limit. The valence quarks in these mesons are a quark and anti-quark pair with combinations of flavors u , d , and s , and the Nambu–Goldstone bosons are mass degenerate (up to electroweak effects) if the these quarks have equal masses. The strange quark is significantly heavier than the light u and d quarks— $2m_s/(m_u + m_d) = 27.46 \pm 0.15 \pm 0.41$ in the $\overline{\text{MS}}$ scheme at a renormalization scale of 2 GeV [81]—so the QCD and quark mass effects can be better understood by comparing pion and kaon observables [80].

Significant SU(3) flavor breaking effects for the pion and kaon have been observed in other studies. For example, the pion and kaon charge radii was measured by experiment to be $r_{\pi^+} = 0.672 \pm 0.008$ and $r_{K^+} = 0.560 \pm 0.031$ [81], which shows flavor breaking effects of around 10%. The neutral pion and kaon have even more striking

flavor symmetry breaking effects, where r_{π^0} vanishes and $r_{K^0} = -0.277 \pm 0.018$ [81]. The pion and kaon PDFs are expected to have similar effects but data is limited, especially for the kaon. There are some pion data from pion induced Drell-Yan [12] but the only existing information on the kaon are some early data on $\bar{u}_{K^-}(x)/\bar{u}_{\pi^-}(x)$ [82]. Naively, quark mass effects lead to the expectation that $s_K(x)$ has more support at large x (harder) while $u_K(x)$ is larger at smaller x (softer). Similarly, $s_K(x)$ is expected, based on flavor breaking, to be harder and $u_K(x)$ to be softer than the u/d quark distributions in the pion. Existing data on the pion and kaon do not contradict these naive expectations, however, the data are not precise enough to be able to make any definite conclusions.

Many LQCD studies focus on the nucleons so studies on the of pions and especially kaons are limited [82, 83, 84, 12, 85]. In addition, other LQCD studies of the pion mostly focus on the electromagnetic form factor [86, 87, 88, 89, 90] and the pion average momentum fraction [91, 92, 93, 94, 88, 95]. Alternative methods of hadron structure, mainly models, have been used to study pions and kaons such as the Dyson-Schwinger equations [96, 97, 98, 17] and the Nambu-Jona-Lasinio model [80, 99]. Since there is a relatively small amount of experimental data to date, first principle calculations of the Mellin moments are important results to obtain.

Here, we present our calculations of the non-trivial moments of the pion and kaon quark PDFs up to $\langle x^3 \rangle$. We study the size of quark mass effects by contrasting the pion and kaon moments, as well as the light and strange kaon moments. These moment calculations give important information about the large- x behavior of the pion and kaon PDFs, and can be used to check the systematic errors associated with recent methods employed by other lattice studies to determine the full x -dependence of the PDFs using, for example, quasi-PDFs [100, 101], pseudo-PDFs [102, 103], and current-current correlation methods [104, 105].

5.1 Calculation Details

We calculate the meson quantities on a single ensemble [72], whose parameters are given in Table 5.1. Such a calculation is computationally very expensive due to the increased gauge noise in operators with covariant derivatives, and the need of a kinematic setup in which the meson carries momentum (boosted frame). Furthermore, lattice calculations of two- and three-derivative operators are almost non-existing,

making this work an exploratory study with unquenched ensembles. Therefore, we chose an ensemble which allows us to do the calculation with a reasonable computational cost. In particular, we choose an ensemble iwth pion mass 200 MeV and kaon mass 530 MeV.

For our pion results, we focus on the valence quark contribution u^+ , where the notation $q^+ \equiv q + \bar{q}$ is used. We only need to calculate the up part of the three-point functions for this study, due to the pseudoscalar structure of the pion and the γ_5 -hermitian relation between the mass degenerate twisted mass quark propagators. From Eq. (3.6) and by calculating the corresponding equation for the down part of the pion, it can be shown that

$$C_{\pi,d}^\Gamma = \pm (C_{\pi,u}^\Gamma)^* . \quad (5.1)$$

The plus/minus sign is due to the γ -structure being either γ_5 -hermitian or anti- γ_5 -hermitian, that is, $\gamma_5 \Gamma^\dagger \gamma_5 = \pm \Gamma$. All of the γ -strucures of the operators used in this study are γ_5 -hermitian, so

$$C_{\pi,u+d}^\Gamma = 2C_{\pi,u}^\Gamma . \quad (5.2)$$

This relation is relevant when we compare our results with phenomenological results.

Parameters							
Ensemble	β	a [fm]	volume $L^3 \times T$	N_f	m_π [MeV]	Lm_π	L [fm]
cA211.32	1.726	0.093	$32^3 \times 64$	u, d, s, c	260	4	3.0

Table 5.1: Parameters of the ensemble used to calculate meson quantities under study.

We calculate our results for 9 momentum frames, the rest frame in which $\mathbf{p}^2 = 0$ and the eight combinations for the boosted frame in which $\mathbf{p}^2 = \frac{12\pi^2}{L^2}$, that is, frames in which the meson's final momenta are the eight permutations of $\mathbf{p} = \frac{2\pi}{L}(\pm 1, \pm 1, \pm 1)$. As discussed in Sec. 3.4, the two and three derivative form factors and their moments can only be calculated in a boosted frame. While the two derivative form factors can be calculated in momentum frames with $\mathbf{p}^2 = \frac{8\pi^2}{L^2}$ which would have a better signal, we choose to calculate $\langle x^2 \rangle$ using the $\mathbf{p}^2 = \frac{12\pi^2}{L^2}$ momentum frames for two reasons. First, there are a smaller number of permutations for the chosen frames, reducing the computational cost for the same number of configurations by 33%. The second reason is that we already need the $\mathbf{p}^2 = \frac{12\pi^2}{L^2}$ frame for the three derivative operators.

Quantity	\mathbf{p}^2	\mathbf{p} combos.	T_{sink}/a	N_{confs}	N_{srCs}	N_{total}
$\langle x \rangle$	0	1	12, 14, 16, 18, 20, 24	122	16	1,952
$\langle x \rangle$	$\frac{12\pi^2}{L}$	8	14, 16, 18	122	16	15,616
$\langle x^2 \rangle, \langle x^3 \rangle$	$\frac{12\pi^2}{L}$	8	12	122	16	15,616
$\langle x^2 \rangle, \langle x^3 \rangle$	$\frac{12\pi^2}{L}$	8	14, 16, 18	122	72	70,272

Table 5.2: Statistics used for each meson correlator.

Therefore, we also get data on the two-derivative operator at no additional cost. For the vector and one derivative form factors and their moments, we take their values calculated in the rest frame as final estimates, since they have a higher-quality signal. We also calculate them in the boosted frame which provides a valuable comparison for the statistical noise and excited-state contamination.

In our rest frame analysis, we include 122 configurations, separated by 20 trajectories to reduce auto-correlation effects. We use 16 randomly chosen source positions on each configuration, resulting in 1952 total statistics for each value of T_{sink} . We calculate six source-sink time separations, $T_{\text{sink}}/a = 12, 14, 16, 18, 20, 24$ which corresponds to $T_{\text{sink}} = 1.12 - 2.23$ fm so that we can thoroughly study and remove any possible excited-state contamination. Mesons structure quantities can be calculated at higher source-sink separations with a reasonable computational cost because the statistical error remains the same with increasing T_{sink} for mesons in the rest frame .

We concluded, based on the analysis of the results in the rest frame, that the source-sink separations $T_{\text{sink}}/a = 12, 14, 16, 18$ are sufficient for removing excited-state effects and extracting the ground state matrix elements. We therefore calculate the matrix elements using only these four time separations in the boosted frames. We use the same 122 configurations for the boosted frames as we used for the rest frame, but include 72 source positions for $T_{\text{sink}}/a = 14, 16, 18$. Since the statistical uncertainties are smaller for $T_{\text{sink}}/a = 12$, we need only to calculate 16 source positions per configuration. The statistics used for the meson quantities are summarized in Table 5.2.

5.2 Effective Mass

As can be seen in the decomposition expressions in Appenix A, the mass of the hadron is needed to calculate the various structure quantities in this study. Because of the boundary conditions, the two-point functions are symmetric around $T/2$. This fact is useful because it allows us to average the correlators corresponding to t and $T - t$, for $t \in (0, T/2)$, a process we refer to as symmetrization. The effective mass is calculated from the symmetrized two-point function according to

$$m_{\text{eff}}^M(t) = \frac{1}{2} \ln \left[\frac{C_M^{2\text{pt}}(t-1) + \sqrt{(C_M^{2\text{pt}}(t-1))^2 - (C_M^{2\text{pt}}(\frac{T}{2}))^2}}{C_M^{2\text{pt}}(t+1) + \sqrt{(C_M^{2\text{pt}}(t+1))^2 - (C_M^{2\text{pt}}(\frac{T}{2}))^2}} \right] \quad (5.3)$$

before it is fit to a plateau for a chosen range. The effective mass can also be calculated using the two state fit, since, in the rest frame, E_0 in Eq. (3.18) is just the mass. Note that the two state fit is also needed to determine the amplitude c_0 that is needed in order to calculate the ratios between the three-point and two-point functions according to Eq. (3.23).

We use both the plateau and two-state fits to confirm that the results are consistent and to select the fit ranges. For both fits, we take the fit over the range $t \in [t_{\text{low}} - 31a]$ and test the dependence of m_{eff} on the range by varying t_{low} . We show this test in Figure 5.1 by plotting m_{eff} as a function of t_{low}/a for both mesons. We consider m_{eff} resulting from the fits to be good if the condition

$$\frac{1}{2} \delta m_{\text{plat}} \geq |m_{\text{plat}} - m_{2\text{st.}}| \quad (5.4)$$

is met, where m_{plat} ($m_{2\text{st.}}$) is the mass resulting from the plateau (two-state) fit and δm_{plat} is the statistical error associated with the value extracted from the plateau method. For $t_{\text{low}}/a \geq 11$ in the plateau fit, there is a very good agreement between the effective masses calculated by the two fitting methods.

The calculation of the Mellin moments in the boosted frame requires the amplitude c_0 calculated from the two-point functions with the same momentum so we repeat a similar process for the two-point functions with momenta $\mathbf{p}_i^2 = \frac{12\pi^2}{L^2}$. We perform the fits on the averaged two-point functions over the eight values of the momentum boost leading to the same \mathbf{p}_i^2 ($\mathbf{p}_i = \frac{2\pi}{L}(\pm 1, \pm 1, \pm 1)$) which increases the accuracy of

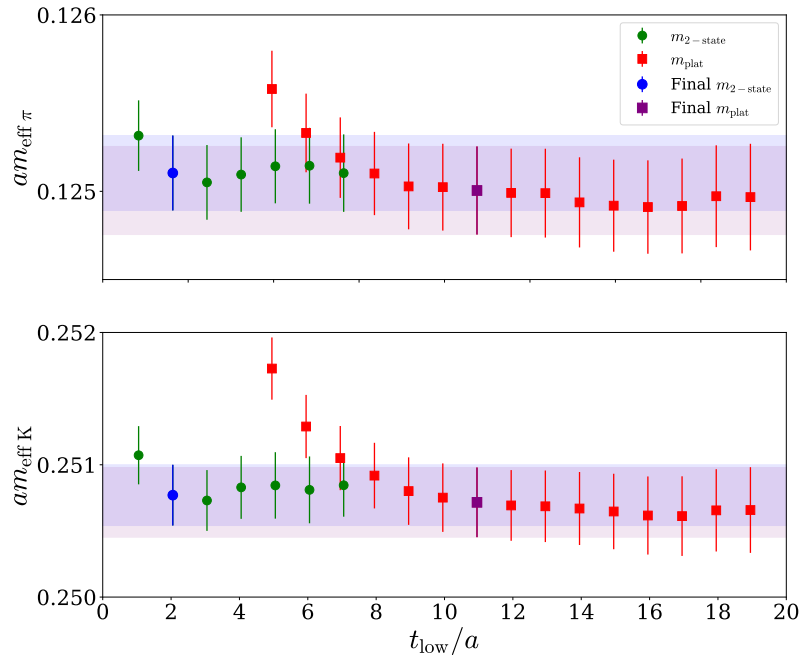


Figure 5.1: Pion (top) and kaon (bottom) mass in the rest frame as a function of the the lowest value of t/a , t_{low}/a entering the fit. The red squares are the results calculated the plateau method and the green circles are those calculated from the two-state fit. We also show the values we choose as our final results of the plateau and two-state fits as a purple square and blue circle, respectively.

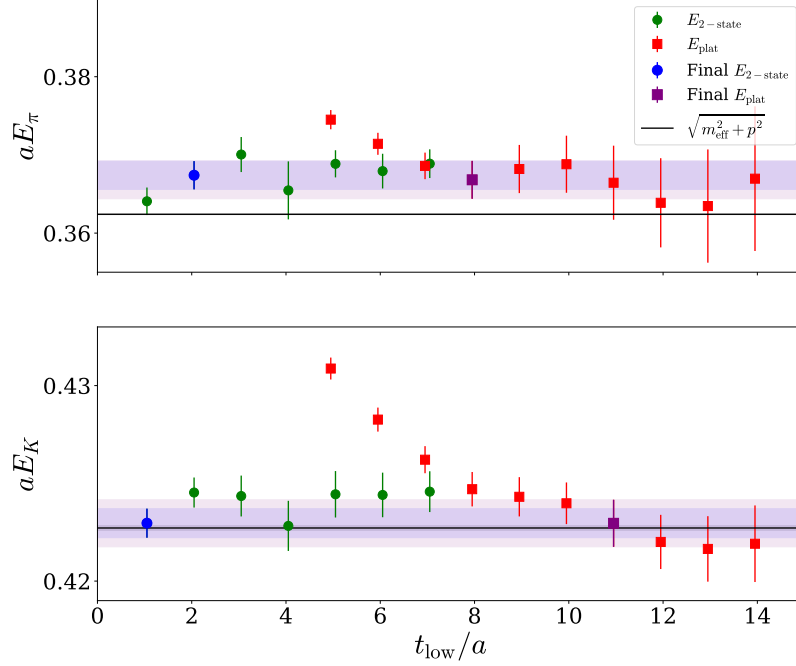


Figure 5.2: Pion (top) and kaon (bottom) mass in the boosted frame as a function of the the lowest value of t/a entering the fit. The notation is the same as in Fig. 5.1.

the results and improves the stability of the fit. Fig. 5.2 shows the results of a similar test of t_{low} done in the boosted frame.

In Table (5.3), we present our results for the ground state energies for both frames, fits, and mesons, along with the corresponding value of t_{low} used in the fit. In Fig. 5.3 we show the effective mass in the rest frame calculated from Eq. (5.3) along with the final fits plotted as a function of t/a . The two fits agree fully for both the pion and the kaon.

5.3 Mellin Moments of PDFs

As discussed previously, we calculate the Mellin moments up to $\langle x^3 \rangle$ for the pion and kaon. We present our results for all three moments in this section. A main aspect of our calculation, never done before, is the study of the excited-state contamination. Lastly, we compare the moments among the different orders as well as among the pion and kaon. The former is a measure of how quickly the PDFs loss support at large x and the latter a measure of SU(3) flavor breaking.

Pion			
\mathbf{p}^2	Fit	aE_0	t_{low}/a
0	Plateau	0.1250(2)	11
	2-state	0.1251(2)	2
$\frac{12\pi^2}{L^2}$	Plateau	0.3668(24)	8
	2-state	0.3674(18)	2
Kaon			
\mathbf{p}^2	Fit	aE_0	t_{low}/a
0	Plateau	0.2507(3)	11
	2-state	0.2508(2)	2
$\frac{12\pi^2}{L^2}$	Plateau	0.4230(12)	11
	2-state	0.4230(7)	1

Table 5.3: Final results of the ground state energies and the corresponding value of t_{low} used in the fit. Note that for $\mathbf{p}^2 = 0$, $E_0 = m_{\text{eff}}$.

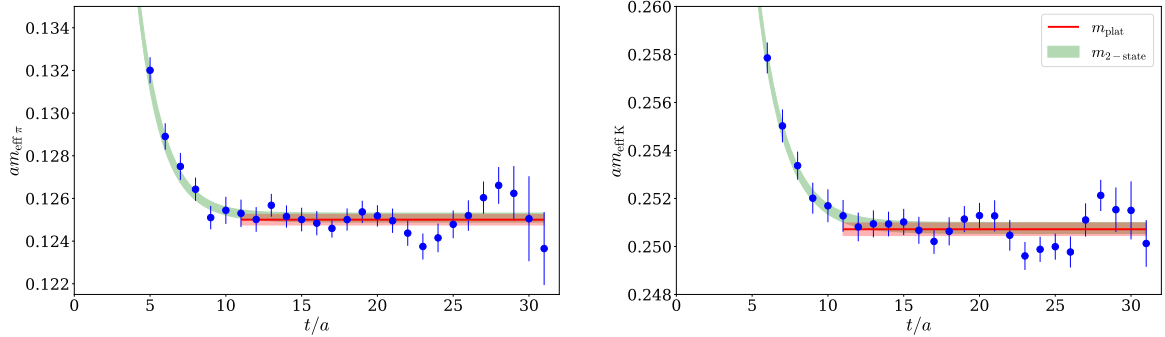


Figure 5.3: Pion (left) and kaon (right) m_{eff} in the rest frame. The fitted value from the plateau m_{plat} is shown with a red band, and from the two-state fit applied on m_{eff} with a green band.

T_{sink}/a	$\langle x \rangle_u^\pi$	$\langle x \rangle_u^K$	$\langle x \rangle_s^K$
12	0.309(3)	0.278(2)	0.339(2)
14	0.287(3)	0.264(2)	0.330(2)
16	0.275(3)	0.257(2)	0.325(2)
18	0.267(3)	0.252(2)	0.322(2)
20	0.261(4)	0.248(2)	0.319(2)
24	0.255(4)	0.244(3)	0.316(2)
2-state (a)	0.261(3)	0.246(2)	0.317(2)
2-state (b)	0.262(4)	0.246(2)	0.317(2)

Table 5.4: Renormalized data for $\langle x \rangle$ for various T_{sink} values and the 2-state fit ((a) $T_{\text{sink}} \in [12 - 24]$, (b) $T_{\text{sink}} \in [14 - 18]$). The numbers shown in the parenthesis are the statistical errors.

5.3.1 First Mellin Moment $\langle \mathbf{x} \rangle$

Fig. 5.4 shows three plots of $\langle x \rangle$ for the mesons as we have previously shown for the nucleon. The three plots from top to bottom are for the up contribution to the pion and the up and strange contributions of the kaon, respectively. The statistical uncertainties do not change with increased source-sink separation, as we expect for mesons at rest. We find that the statistical errors associated with the plateau values are 2% or less.

The excited-states contamination is similar for both the pion and kaon and we find convergence for $T_{\text{sink}} \gtrsim 18a$. We collect the plateau fits of $\langle x \rangle$ for each source-sink separation in Table 5.4 as well as the two-state fit using two ranges, $T_{\text{sink}}/a \in [12 - 24]$ and $T_{\text{sink}}/a \in [14 - 18]$. The two-state fit result plotted in Fig. 5.4 uses the fit range $T_{\text{sink}}/a \in [12 - 24]$. Since we find the excited states to be suppressed for $T_{\text{sink}} \gtrsim 18a$, we test the dependence of $\langle x \rangle$ on the two-state fit range to guide our decision of what T_{sink} values to calculate in the boosted frame. Since the two-state fit calculated from both fit ranges are compatible, and in fact identical for the kaon, we chose the values $T_{\text{sink}} = 14, 16, 18$ in the boosted frame. This is necessary because, in the boosted frame, the statistical errors increase with T_{sink} . Therefore, more statistics are needed and going beyond three to four values of T_{sink} becomes very computationally expensive.

To investigate the signal in the $\mathbf{p}^2 = \frac{12\pi^2}{L^2}$ boosted frame, we perform a similar calculation of $\langle x \rangle$ for the pion and kaon with boosted momentum, shown in Fig. 5.5.

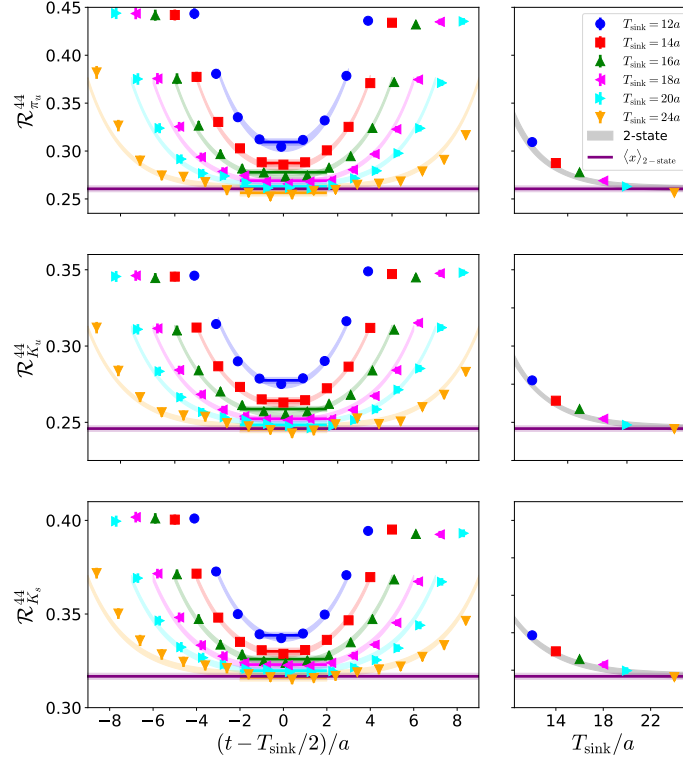


Figure 5.4: Plots of the meson $\langle x \rangle$ results in the rest frame. The three rows are for the up part of the pion (top) and the up (center) and strange (bottom) parts of the kaon. In the left column, the ratios corresponding to $\langle x \rangle$ calculated according to Eq. A.19 are plotted for $T_{\text{sink}}/a = 12, 14, 16, 18, 20, 24$ as blue circles, red squares, green downward-pointing triangles, magenta left-pointing triangles, cyan right-pointing triangles, and orange downward-pointing triangles, respectively. In the right column, the plateau fits are plotted as a function of T_{sink} . The gray bands in the right column are calculated from the two-state parameters as a function of T_{sink} and at a constant $t = T_{\text{sink}}/2$. The purple bands going across all plots are the results from the two-state fit.

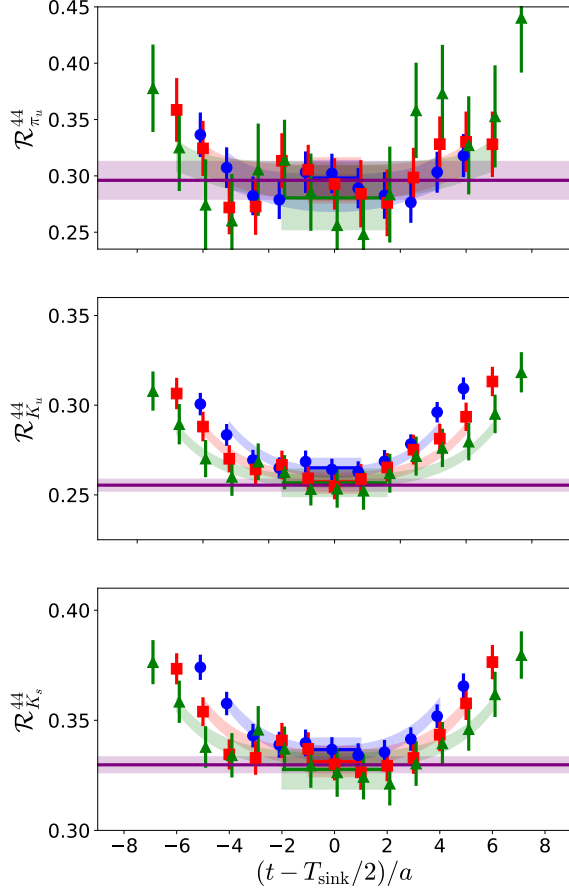


Figure 5.5: $\langle x \rangle$ results in the $\mathbf{p}^2 = \frac{12\pi^2}{L^2}$ boosted frame plotted with notation similar to the left column of Figure 5.4. The different source-separations, $T_{\text{sink}}/a = 14, 16, 18$, are shown as blue circles, red squares, and green triangles, respectively.

In the boosted frame, the decomposition is given by Eq. A.12. We use the same number of source positions as for the rest frame but the statistics are 8 times higher since there are eight momentum permutations which have $\mathbf{p}^2 = \frac{12\pi^2}{L^2}$. In the boosted frame, all of the plateau values are consistent with the two-state fit results. This indicates that the systematic uncertainties caused by excited-states are within the statistical uncertainties, which are larger than those in the rest frame. We collect the boosted frame results in Table 5.5.

We also compare the ratios leading to $\langle x \rangle$ in the rest frame to those in the boosted frame in Fig. 5.6. The two frames agree within statistical uncertainty for all particles and flavors. The statistical uncertainties of the plateau fits in the boosted frame grows larger with increased T_{sink} , as is expected for mesons with nonzero momentum. The

T_{sink}/a	$\langle x \rangle_u^\pi$	$\langle x \rangle_u^K$	$\langle x \rangle_s^K$
14	0.298(16)	0.265(5)	0.337(5)
16	0.296(20)	0.257(6)	0.331(7)
18	0.280(28)	0.257(8)	0.328(9)
2-state	0.296(17)	0.255(3)	0.330(4)

Table 5.5: Plateau fit results for $\langle x \rangle$ in the $\mathbf{p}^2 = \frac{12\pi^2}{L^2}$ boosted frame for the three T_{sink} values and the 2-state fit using $T_{\text{sink}} \in [14 - 18]$. The number shown in the parenthesis are the statistical errors.

errors increase from 5% to 10% for $\langle x \rangle_u^\pi$ and from 2% to 3% for $\langle x \rangle_u^K$ and $\langle x \rangle_s^K$. Compare this to the error in the rest frame which is constant for all source-sink separations and is less than or equal to 2%.

We take as our final results for $\langle x \rangle$ the values calculated by the two-state fits in the rest frame. They are

$$\langle x \rangle_{u^+}^\pi = 0.261(3)(6), \quad (5.5)$$

$$\langle x \rangle_{u^+}^K = 0.246(2)(2), \quad (5.6)$$

$$\langle x \rangle_{s^+}^K = 0.317(2)(1). \quad (5.7)$$

The quantities in the first parentheses are the statistical uncertainties and in the second are the systematics uncertainties due to excited-state effects. We use for this uncertainty the difference between $\langle x \rangle$ calculated using the two-state fit and the plateau fit at the highest value of T_{sink} , which, here, is $T_{\text{sink}}/a = 24$. These results were published in Physical Review D [106],

5.3.2 The Second Mellin Moment $\langle \mathbf{x}^2 \rangle$

There are several factors which make calculating the second nontrivial Mellin moment, $\langle x^2 \rangle$ a greater challenge than for $\langle x \rangle$. Firstly, as we discussed previously, $\langle x^2 \rangle$ needs to be calculated in a boosted momentum frame, since the kinematic factor in Eq. (A.13) vanishes in the rest frame. The signal of the meson state gets worse when we add momentum and, since we have decided to use the same momentum frame as we use for $\langle x^3 \rangle$ ($\mathbf{p}^2 = \frac{12\pi^2}{L^2}$), the noise is even higher than if we had used the minimum momentum necessary to calculate $\langle x^2 \rangle$. Secondly, the two covariant derivatives in the insertion operator require information about more gauge link around the insertion

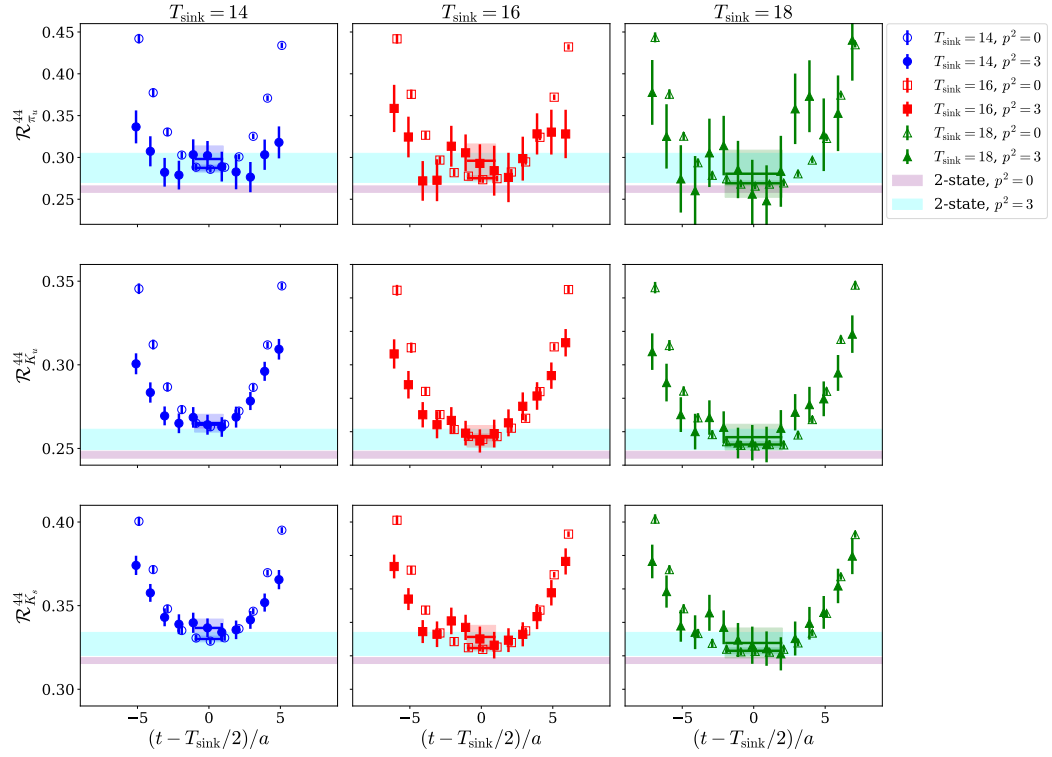


Figure 5.6: Comparison of $\langle x \rangle$ in the boosted (filled symbols) and rest frame (open symbols). From top to bottom we show $\langle x \rangle$ for the pion and kaon up and strange contribution. Results at $T_{\text{sink}}/a = 14, 16, 18$ are shown in the left, center and right panels, respectively.

point which also increases noise. Thirdly, there are now three Dirac indices in the insertion operator which makes the renormalization pattern more complicated. The indices also have to be different so that there is no mixing with other operators. We choose one of the indices to always be temporal which simplifies the kinematic factor so the operator we use is $\mathcal{O}^{\{ij4\}}$ with $i \neq j$. As a result, we increased the statistics included in the $\langle x^2 \rangle$ analysis so that there are three and a half times more than we used for $\langle x \rangle$.

In Fig. 5.7, we show the ratio leading to $\langle x^2 \rangle$ for the pion and the kaon. We plot the data for the three values of T_{sink} considered, that is $T_{\text{sink}}/a = 12, 14, 16, 18$, and compare with the two-state results. For the pion, we find that all plateau values agree within uncertainties with the two-state results as was the case with $\langle x \rangle$ in the boosted frame and which again indicates that excited-states contamination are small compared to the errors on this quantity. For the kaon, we observe larger excited-state effects at $T_{\text{sink}} = 12a$ but find convergence at higher separations. We collect the plateau and two-state fit values in Table 5.6.

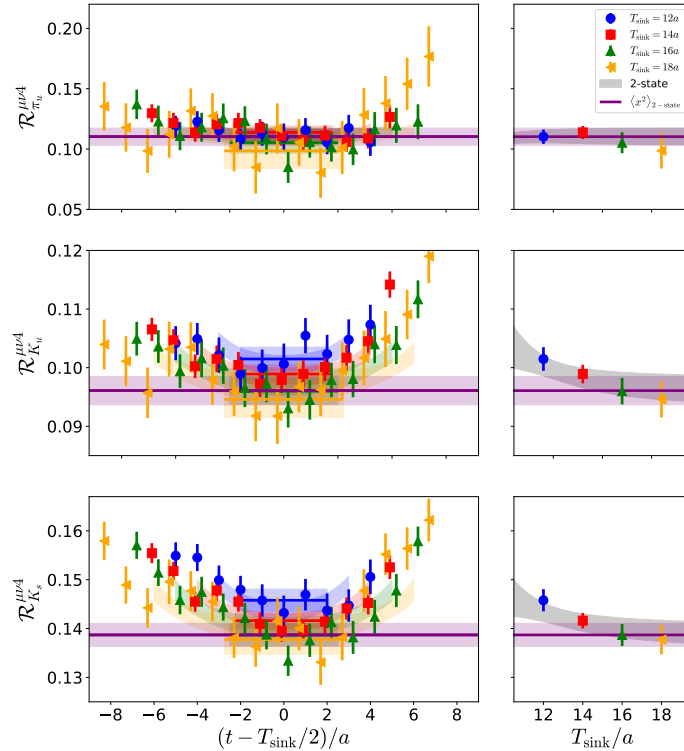


Figure 5.7: $\langle x^2 \rangle$ plotted with notation similar to Figure 5.4. The different source-separations, $T_{\text{sink}}/a = 12, 14, 16, 18$, are shown as blue circles, red squares, green upward-pointing triangles, and magenta left-pointing triangles, respectively.

T_{sink}/a	$\langle x^2 \rangle_u^\pi$	$\langle x^2 \rangle_u^K$	$\langle x^2 \rangle_s^K$
12	0.110(6)	0.101(2)	0.146(2)
14	0.114(5)	0.099(2)	0.142(2)
16	0.105(9)	0.096(2)	0.139(2)
18	0.099(15)	0.095(3)	0.138(3)
2-state	0.110(7)	0.096(2)	0.139(2)

Table 5.6: Renormalized data for $\langle x^2 \rangle$ at each value of T_{sink} and the two-state fit ($T_{\text{sink}} \in [12 - 18]$). The number shown in the parenthesis is statistical error.

We again use the two-state fit for our final results of $\langle x^2 \rangle$. Our final results are then

$$\langle x^2 \rangle_{u^+}^\pi = 0.110(7)(12), \quad (5.8)$$

$$\langle x^2 \rangle_{u^+}^K = 0.096(2)(2), \quad (5.9)$$

$$\langle x^2 \rangle_{s^+}^K = 0.139(2)(1). \quad (5.10)$$

The results were published with fewer statistics along with our $\langle x \rangle$ results in Physical Review D [106].

5.3.3 The Third Mellin Moment $\langle \mathbf{x}^3 \rangle$

The challenges in the $\langle x^2 \rangle$ calculation are not only present in the calculation of $\langle x^3 \rangle$ but they are more pronounced. The boosted momentum frame we calculate $\langle x^3 \rangle$ in is the same as we used for $\langle x^2 \rangle$ and is the lowest momentum without a vanishing kinematic factor in Eq. A.14. The insertion operator needed for $\langle x^3 \rangle$ has three covariant derivatives, which consequently needs more gauge links. Additionally, there are now four Dirac indices, leading to a renormalization which is even more complicated and limiting the number of combinations with unique indices. We use the only combinations available, which are the permutations of \mathcal{O}^{1234} . The number of statistics used in the analysis of $\langle x^3 \rangle$ are the same as those used for $\langle x^2 \rangle$.

We plot the ratios leading to $\langle x^3 \rangle$ in Fig. 5.8 for each value of $T_{\text{sink}} = 12, 14, 16, 18$ and compare the two-state fit results. The fit results for all particles and flavors are consistent with one another, though the low signal can be seen by the fact that the lower range of the $\langle x^3 \rangle_u^\pi$ errors are just barely above zero for most fits and the plateau fit at $T_{\text{sink}}/a = 18$ is zero within statistical errors. For the kaon, however, there is

already a clear signal at these statistics. The plateau fits of $\langle x^3 \rangle$, along with the two-state fit, are shown in Table 5.7.

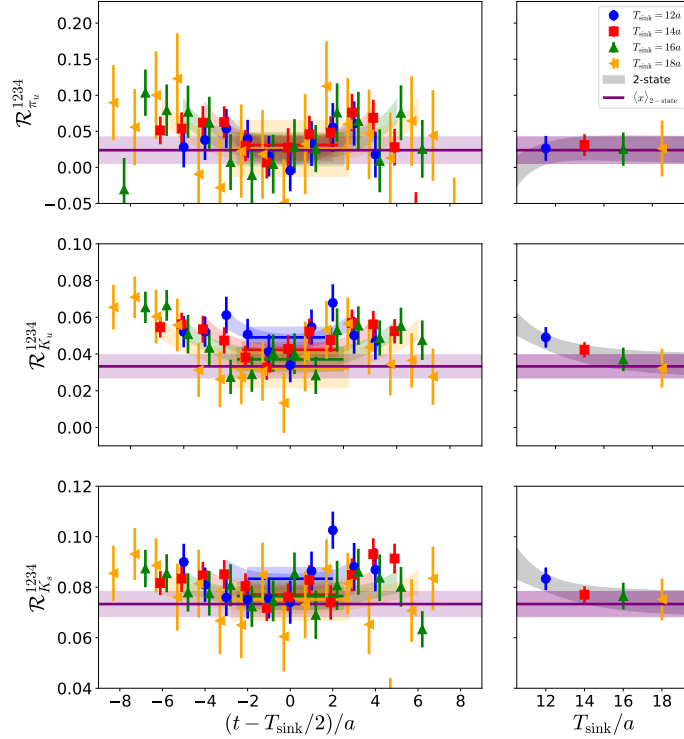


Figure 5.8: $\langle x^3 \rangle$ plotted with notation similar to Figure 5.4. The different source-separations, $T_{\text{sink}}/a = 12, 14, 16, 18$, are shown as blue circles, red squares, green upward-pointing triangles, and magenta left-pointing triangles, respectively.

T_{sink}/a	$\langle x^3 \rangle_u^\pi$	$\langle x^3 \rangle_u^K$	$\langle x^3 \rangle_s^K$
12	0.026(17)	0.043(7)	0.079(6)
14	0.031(15)	0.042(4)	0.077(3)
16	0.025(23)	0.037(6)	0.077(5)
18	0.026(39)	0.032(11)	0.075(8)
2-state	0.024(18)	0.033(6)	0.073(5)

Table 5.7: Renormalized data for $\langle x^3 \rangle$ at each value of T_{sink} and the two-state fit ($T_{\text{sink}} \in [12 - 18]$). The number shown in the parenthesis is statistical error.

Like it was for $\langle x^2 \rangle$, the statistical errors on the $\langle x^3 \rangle$ two-state fit are lower than for the $T_{\text{sink}}/a = 18$ plateau, though we take the two-state fit to be our final results since the pion plateau fit at $tsink/a = 18$ does not have signal. Including the excited-state

systematic uncertainties, these results are

$$\langle x^3 \rangle_{u^+}^{\pi} = 0.024(18)(2), \quad (5.11)$$

$$\langle x^3 \rangle_{u^+}^K = 0.033(6)(1), \quad (5.12)$$

$$\langle x^3 \rangle_{s^+}^K = 0.073(5)(2). \quad (5.13)$$

5.3.4 Comparison Between Moments

With each higher moment, the value we find becomes smaller than the last, that is, $\langle x \rangle > \langle x^2 \rangle > \langle x^3 \rangle$. This is caused by the PDFs decreasing at high- x and is expected based on model calculations. There are several model results which disagree on how quickly the PDFs decrease at high- x . A way to quantify this using only the moments is to calculate the ratios of $\langle x^2 \rangle$ and $\langle x^3 \rangle$ to $\langle x \rangle$. These ratios are a measure of the support a PDF has at high- x , since the moments will be higher or lower depending on if the PDF skews more toward high or low x with the higher moments being affected more strongly. We take as our results the ratio between the two-state fits. The $\langle x^2 \rangle$ to $\langle x \rangle$ ratios are

$$\frac{\langle x^2 \rangle_{u^+}^{\pi}}{\langle x \rangle_{u^+}^{\pi}} = 0.423(28)(57), \quad (5.14)$$

$$\frac{\langle x^2 \rangle_{u^+}^K}{\langle x \rangle_{u^+}^K} = 0.391(10)(16), \quad (5.15)$$

$$\frac{\langle x^2 \rangle_{s^+}^K}{\langle x \rangle_{s^+}^K} = 0.438(8)(11), \quad (5.16)$$

the $\langle x^3 \rangle$ to $\langle x \rangle$ ratios are

$$\frac{\langle x^3 \rangle_{\pi}^{u^+}}{\langle x \rangle_{\pi}^{u^+}} = 0.092(71)(6), \quad (5.17)$$

$$\frac{\langle x^3 \rangle_K^{u^+}}{\langle x \rangle_K^{u^+}} = 0.135(26)(8), \quad (5.18)$$

$$\frac{\langle x^3 \rangle_K^{s^+}}{\langle x \rangle_K^{s^+}} = 0.232(16)(1), \quad (5.19)$$

and the $\langle x^3 \rangle$ to $\langle x^2 \rangle$ ratios are

$$\frac{\langle x^3 \rangle_{\pi}^{u^+}}{\langle x^2 \rangle_{\pi}^{u^+}} = 0.216(167)(50), \quad (5.20)$$

$$\frac{\langle x^3 \rangle_K^{u^+}}{\langle x^2 \rangle_K^{u^+}} = 0.346(66)(5), \quad (5.21)$$

$$\frac{\langle x^3 \rangle_K^{s^+}}{\langle x^2 \rangle_K^{s^+}} = 0.529(36)(18). \quad (5.22)$$

The ratio in Eqs. (5.17), (5.18), and (5.19) are lower than those in Eqs. (5.14), (5.15), and (5.16) which supports the fact that the pion and kaon PDF's diminish at high- x .

We also compare $\langle x^n \rangle_{\pi}$ and $\langle x^n \rangle_K$ which is interesting because it is related to the SU(3) flavor symmetry breaking. The flavor symmetry breaking is caused by the heavier strange quark mass and is the reason that kaons are heavier compared to pions. Studying the SU(3) flavor symmetry breaking is important because some quantities, such as the pion and kaon radii [81], may be sensitive to these effects. We find the ratios of the up quark contributions to be

$$\frac{\langle x \rangle_{\pi}^{u^+}}{\langle x \rangle_K^{u^+}} = 1.060(9)(7), \quad (5.23)$$

$$\frac{\langle x^2 \rangle_{\pi}^{u^+}}{\langle x^2 \rangle_K^{u^+}} = 1.148(57)(106), \quad (5.24)$$

$$\frac{\langle x^3 \rangle_{\pi}^{u^+}}{\langle x^3 \rangle_K^{u^+}} = 0.717(488)(94), \quad (5.25)$$

while for the ratio between the strange quark in the kaon and the up quark in the pion, we find

$$\frac{\langle x \rangle_{\pi}^{u^+}}{\langle x \rangle_K^{s^+}} = 0.823(8)(10), \quad (5.26)$$

$$\frac{\langle x^2 \rangle_{\pi}^{u^+}}{\langle x^2 \rangle_K^{s^+}} = 0.795(45)(80), \quad (5.27)$$

$$\frac{\langle x^3 \rangle_{\pi}^{u^+}}{\langle x^3 \rangle_K^{s^+}} = 0.325(244)(23). \quad (5.28)$$

The SU(3) symmetry breaking in our calculates is $\sim 5 - 10\%$ for $\langle x \rangle$ and $\sim 10 - 20\%$ for $\langle x^2 \rangle$. The $\langle x^3 \rangle$ results have larger uncertainties but indicate a $\sim 30 - 50\%$ symmetry breaking. The increased flavor breaking with increasing moment supports the intuitive argument based on the quark masses of the different flavors which expects that the strange quark PDF has its support at higher x than the up quark PDF. We will also discuss SU(3) flavor symmetry breaking effects when reconstructing the PDFs in Section 5.4.

5.3.5 Comparison with Other Studies

Of the moments we investigate, the pion $\langle x \rangle$ has been studied the most in LQCD [91, 107, 108, 93, 95], including results which are calculated directly at the physical point [94]. The kaon $\langle x \rangle$, on the other hand, has not been studied as much. There is a method which has been used for both mesons to calculate PDF moments using non-local operators [109, 15, 110, 14, 111, 112, 101], however, the systematic uncertainties which affect such calculation are very different from ours and so we do not attempt to compare with them. We choose to compare $\langle x \rangle_{u^+}^{\pi}$ with lattice results calculated on ensembles which have the same or similar pion masses, that is 240 - 270 MeV.

In Ref. [95], several $N_f = 2 + 1 + 1$ ensembles which have similar lattice spacings ($a = 0.09$ fm) and lattice size to the one used in this work were used to calculate the pion moments. That study used only twisted mass fermions without the clover improvement. For one esemble (A30.32), they find $\langle x \rangle_{u^+}^{\pi} = 0.2586(41)(28)$. They also use another ensemble (B25.32) with a smaller lattice spacing $a = 0.082$ fm, $m_{\pi} = 260$ MeV and $m_{\pi}L = 3.5$, for which they found $\langle x \rangle_{u^+}^{\pi} = 0.2523(51)(71)$. Both values agree with our $N_f = 2 + 1 + 1$ clover-improved results $\langle x \rangle_{u^+}^{\pi} = 0.261(3)(6)$.

It is also interesting to compare our results with phenomenology calculations, which can be found in Refs. [10] and [113]. The analysis of Ref. [113] is older and gives a value of $\langle x \rangle_u^{\pi} = 0.217(10)$, in the $\overline{\text{MS}}$ at a scale at $(5.2 \text{ GeV})^2$, while ours is $(2 \text{ GeV})^2$. Converting their value to 2 GeV results in $\langle x \rangle_u^{\pi} = 0.361(17)$. The JAM Collaboration [10] performed a more recent analysis on a large set of experimental data including Drell-Yan data and leading neutron electroproduction from HERA. Our result, $2\langle x \rangle_{u^+}^{\pi} = 0.522(13)$, where the error in the parenthesis is the combined statistical and systematic uncertainties added in quadrature, agrees reasonably well with $\langle x \rangle_{valence}^{\pi} = 0.480(10)$, the results found by the JAM collaboration. Our value being higher, by $\sim 4\%$, might be because our calculation does not account for errors

caused by non-physical quark masses or a finite volume. This value will decrease when taking both the chiral extrapolation and the continuum limit $a \rightarrow 0$ as demonstrated in Ref. [95]. We also note our result, as well as all lattice calculations to date, do not calculate the disconnected contributions which must be included in a final comparison with phenomenology. We summarize the comparisons of $\langle x \rangle_u^\pi$ results in Table 5.8.

There are not a lot of experimental data available which can be used in a global fit of the kaon PDF so, instead, we compared our moments to what is expected from model calculations. We find the relation between our lattice results $\langle x \rangle_{u^+}^K < \langle x \rangle_{u^+}^\pi < \langle x \rangle_{s^+}^K$ which is in agreement with many phenomenological calculations, including the DSE results of Ref. [17]. This relative relation between the kaon momentum fractions and the pion can be explained by the fact that the s quark is heavier than the u quark so $s_K(x)$, therefore, skews to larger x which is compensated by a shift in $u_K(x)$ to smaller x . If the quark masses were degenerate, then the moments should also be the same, and so we find that flavor breaking affects these moments by up to 20%.

Even for the pion, there are not as many LQCD calculations of $\langle x^2 \rangle$ as there are of $\langle x \rangle$. It has been calculated, though, using various lattice formulations in Refs. [91, 107, 95]. We can compare our results directly with Ref. [95], who find $\langle x^2 \rangle_{u^+}^\pi = 0.131(18)(24)$ and $\langle x^2 \rangle_{u^+}^\pi = 0.132(40)(53)$ for ensembles A30.32 and B25.32, respectively and our value is in agreement with these estimates, within their large uncertainties. Other studies, in Refs. [91, 107] calculate $\langle x^2 \rangle^\pi$ to be $\langle x^2 \rangle_{u^+}^\pi = 0.128(9)(4)$. While this is consistent with the value we find, the other study uses a different operator with two of the same Dirac indices. We expect this choice of operator to lead to mixing and, therefore, a more complicated renormalization pattern, which Refs. [91, 107] do not address.

Again, there are a limited number of studies done for the kaon moments, so we compare our values to what we find for the pion, like we did for $\langle x \rangle$. Due to the larger uncertainties, we can only definitively say that $\langle x^2 \rangle_{u^+}^K, \langle x^2 \rangle_{u^+}^\pi < \langle x^2 \rangle_{s^+}^K$ and cannot determine the order of $\langle x^2 \rangle_{u^+}^K$ and $\langle x^2 \rangle_{u^+}^\pi$. The pion and kaon $\langle x^2 \rangle$ moments are expected to have the same ordering as $\langle x \rangle$ which we find to be the case for the kaon strange contribution. It would be very interesting if, unlike the quark momentum fraction, $\langle x^2 \rangle_{u^+}^K$ were higher than $\langle x^2 \rangle_{u^+}^\pi$ but more precise results are needed to make that determination.

The phenomenological study in Ref. [10] estimates a value of $\langle x^2 \rangle_{valence}^\pi = 0.210(5)$, which is compatible within uncertainties with our value of $2\langle x^2 \rangle_{u^+}^\pi = 0.214(35)$. It

should be noted, however, that the sea quarks, which are automatically included in the LQCD calculation, are not included in the phenomenological value. We also compare to the results from Ref. [113], for completeness, who find $\langle x^2 \rangle_u^\pi = 0.087(5)$ at scale of $(5.2 \text{ GeV})^2$. We convert this result to 2 GeV which results in $\langle x^2 \rangle_u^\pi = 0.169(10)$ and is compatible with our value.

Reference	$\langle x \rangle^\pi$	$\langle x^2 \rangle^\pi$
This work (lattice)	0.522(13)	0.220(28)
Ref. [95] (lattice)	0.517(99)	0.262(60)
Ref. [95] (lattice)	0.505 (174)	0.264(133)
Ref. [113] (global fits)	0.361(17)	0.169(10)
Ref. [10] (global fits)	0.480(10)	0.210(5)

Table 5.8: Comparison of lattice results and phenomenological data for $\langle x \rangle^\pi$ and $\langle x^2 \rangle^\pi$.

We calculate $\langle x^3 \rangle$ directly from the three-derivative local vector operator but there are only a few other studies which have calculated it using the same method. The first calculation was the pioneering work of QCDSF-UKQCD in 1997 [114], which was later extended in 2007 [91, 107], and reanalyzed in Ref. [115]. All of these study the pion and so ours are the first reported results for the kaon which use the three-derivative local operator. There are a number of alternative methods to extract the x -dependence of PDFs using LQCD, such as the quasi-PDFs [100, 116], pseudo-Ioffe-time-distributions (ITD) [117], and current-current correlators [118, 119, 104]. The lowest moments of the pion, and less for the kaon, have been obtained from such methods [14, 112, 101], by either integrating on the x -dependent PDF, or by the so-called OPE without an OPE method [120]. Since there are so few calculations of $\langle x^3 \rangle$ using the three-derivative local operators, we compare our results with all of these method. The comparison is qualitative for now, however, as the calculations use different methodologies and have their own systematic uncertainties which have not all been quantified.

The results of Ref. [114] are calculated using the quenched approximation for Wilson fermions and ensembles with pion masses of 712, 1013, and 1208 MeV. They use an operator with two different pairs of indices ($\mathcal{O}_{v4} = \mathcal{O}^{ii44}$), so that only one spatial direction is needed for momentum boost. This operator, however, mixes with lower dimensional operators. This mixing is difficult to eliminate so it is ignored in the

calculation and results from perturbation theory are used to renormalize $\langle x^3 \rangle_\pi$. The results for $\langle x^3 \rangle_\pi^u$ are reported in the $\overline{\text{MS}}$ scheme at a scale of 2.4 GeV as 0.0619(45), 0.0580(65), and 0.054(18) for the ensembles with pion mass 1208, 1013, and 712 MeV, respectively. We evolve their estimate for the 712 MeV ensemble to 2 GeV which is listed in Table 5.9. The results of Ref. [114] have been reanalyzed using different methods for the the chiral extrapolation on the three ensembles by Ref. [115]. For two of the more recent extrapolation methods they report $\langle x^3 \rangle_\pi^u = 0.043(15)(3)$, and 0.05(2) at a scale of 2.4 GeV. We evolve the first value to 2 GeV and include it in Table 5.9.

QCDSF-UKQCD have since performed an improved calculation of $\langle x^3 \rangle$ [91]. The calculation is improved compared to the previous study by using ensembles with unquenched ($N_f = 2$) clover fermions and pion masses between 450 - 1180 MeV. The operator \mathcal{O}_{v4} is used which still has mixing. The mixing is again ignored but the renormalization is done non-perturbatively. Their result is a chiral extrapolation to the physical pion mass which gets its accuracy from the accuracy of the $m_\pi \geq 800$ MeV ensembles. The chirally extrapolated result is $\langle x^3 \rangle_\pi = 0.074(9)(4)$ in the $\overline{\text{MS}}$ at a scale of 2 GeV. Another updated calculation which again uses $N_f = 2$ clover fermions was performed at four β values, lattice spacings ranging between 0.068–0.115 fm, and a wide range of pion masses between 440–1173 MeV. The results are reported for the valence contributions in the RGI scheme in the thesis of Ref. [107] as $\langle x^3 \rangle_\pi^u = 0.158(49)$, 0.117(53), 0.166(28), 0.152(44), for the four β values. We choose to compare to their result at $\beta = 5.40$ which they obtain from an extrapolation containing $a = 0.068 - 0.077$ fm and $m_\pi = 444 - 1037$ MeV. We use the conversion factors provided in Ref. [107] to evolve this result to $\overline{\text{MS}}$ at 2 GeV and include it in Table 5.9.

Much more recently, there have been exploratory studies using LQCD to calculate the x -dependent pion and kaon PDFs [109, 15, 110, 14, 111, 112, 101]. Some of these studies calculate the first moments indirectly from the PDFs. These new methods do not use local operators like we do in our study or in the previously discussed studies. Instead, they calculate matrix elements of non-local operators by separating the quark fields by a finite spatial distance connecting them by a straight Wilson line. Therefore, there is none of the mixing among operators which we avoid by carefully choosing our operator indices. These methods have other systematic uncertainties, however, which ours do not so only a qualitative comparison can be made.

In Ref. [14], the pseudo-ITD approach is used to calculate the x -dependence of PDFs from two ensembles of $N_f = 2 + 1$ clover fermions, a 415 MeV pion mass, and different volumes. The lowest moments are extracted via two methods: OPE without an OPE which results in $\langle x^3 \rangle_\pi^u = 0.057(4)$ and an integration of the PDFs that gives $\langle x^3 \rangle_\pi^u = 0.046(19)$. Both values are given in the $\overline{\text{MS}}$ at 2 GeV, and we compare the latter to our estimate in Table 5.9. The quasi-PDFs method is used in Ref. [112] with a mixed-action setup (clover on HISQ) on three ensembles with pion mass 310, 217, 319 MeV and two lattice spacings ($a = 0.06, 0.12$ fm). They use non-perturbative renormalization and chirally extrapolate to the physics point. The result, $\langle x^3 \rangle_\pi^u = 0.057(10)$, is reported at a scale of 5.2 GeV. They also perform a similar analysis for the kaon with the results $\langle x^3 \rangle_K^u = 0.042(6)$ and $\langle x^3 \rangle_K^u = 0.070(6)$. We evolve these results to 2 GeV and compare them to our results for the pion and kaon in Table 5.9 and Table 5.10, respectively. Finally, the work of Ref. [101] calculate the pion PDF using both the quasi-PDFs and pseudo-ITD approaches and also uses a mixed-action setup of clover valence fermions on $N_f = 2 + 1$ HISQ configurations. They use two ensembles at $m_\pi = 300$ MeV and lattice spacing 0.04 and 0.06 fm. They fit the pion PDF with a 2- and 4-parameter ansatz to extract the third moment. Their reported results for the valence case are $\langle x^3 \rangle_\pi^{val} = 0.0652(49)(36)$ and $0.0647(47)(38)$ for the 2-parameter and 4-parameter fits in the $\overline{\text{MS}}$ scheme at 3.2 GeV. Again, we evolve their $\langle x^3 \rangle_\pi^u$ results to 2 GeV and include it in Table 5.9.

Reference	Method	mixing	m_π (MeV)	$\langle x^3 \rangle_\pi^u$ (2GeV)
This work	local operator	not present	260	0.024(18)
Ref. [114]	local operator	present	chiral extrap.	0.051(21)
Ref. [115]	local operator	present	chiral extrap.	0.046(16)
Ref. [107]	local operator	present	chiral extrap.	0.074(10)
Ref. [14]	pseudo-ITD	N/A	415	0.046(19)
Ref. [112]	quasi-PDF	N/A	chiral extrap.	0.073(13)
Ref. [101]	pseudo-ITD	N/A	300	0.075(61)

Table 5.9: Comparison of lattice results for $\langle x^3 \rangle_\pi^u$ in the $\overline{\text{MS}}$ scheme at 2 GeV. The evolution from the reported scale (“initial” scale) to 2 GeV is applied to NNLO. Statistical and systematic uncertainties have been added in quadrature where applicable.

We compare the range of pion $\langle x^3 \rangle$ values obtained through the different methods discussed above in Table 5.9. We find agreement within uncertainties between our

results and the other local operator calculations Refs. [114, 115, 91, 107]. The higher value of Ref. [114] compared to the other studies can be explained by their use of a perturbative renormalization prescription. We also find that our results are compatible with the indirect methods which integrate or fit the pion PDF to obtain the Mellin moments, except for the results in Ref. [112], which calculates a higher value than the local operator calculations. Most of the uncertainties in Table 5.9, however, are large so the comparison is inconclusive. Within the ranges of values, our result is at the lower end which is due to the excited states being suppressed in our calculation. For example, the separations used in in other studies are around our value of $T_{\text{sink}}/a = 14$, where are calculation is higher (0.031(15)).

In Table 5.10, we compare our results for the kaon $\langle x^3 \rangle$ with the results of Ref. [112]. The two calculations are compatible with one another within errors. We evolved the value of Ref. [112] from 5.2 GeV to 2 GeV.

Reference	Method	m_π (MeV)	$\langle x^3 \rangle_K^u$ (2GeV)	$\langle x^3 \rangle_K^s$ (2GeV)
This work	local operator	260	0.033(6)	0.073(5)
Ref. [112]	quasi-PDF	chiral extrap.	0.042(6)	0.090(8)

Table 5.10: Comparison of lattice results for $\langle x^3 \rangle_K^u$ and $\langle x^3 \rangle_K^s$ in the $\overline{\text{MS}}$ scheme at 2 GeV. The evolution from the reported scale (“initial” scale) to 2 GeV is applied to NNLO. Statistical and systematic uncertainties have been added in quadrature where applicable.

We also compare our $\langle x^3 \rangle$ results to model calculations and global fits on experimental data, of which there are a few for the pion and kaon. As was the case for the comparisons with many of the other lattice calculations, this comparison is only qualitative many of the calculation do not have quantified uncertainties. The pion mass used for our calculation is also higher than the physical value and we have only used the connected diagrams.

Ref. [121] is one of the earliest calculations and is a next-to-leading-order analysis of several $\pi \pm N$ experimental data, including Drell-Yan and prompt photon production. They find $\langle x^3 \rangle_\pi^u = 0.058(4)$ at a scale of 2 GeV. An updated analysis was presented much later in Ref. [113]. That study analyzes the Fermilab E-615 pionic Drell-Yan data to obtain the moments of pion PDFs to next-to-leading order with a value of $\langle x^3 \rangle_\pi^u = 0.045(3)$ at 5.2 GeV. The JAM global fit analysis calculates the third non-trivial pion moment at 1.3 GeV with a result of $\langle x^3 \rangle_\pi^u = 0.074$. [10].

In Ref. [122], Dyson-Schwinger equations (DSEs) are used to calculate the valence quark PDF for the pion and obtained result is $\langle x^3 \rangle_\pi^u = 0.049$ at 2 GeV. A more recent DSE studies, Ref. [16], also studies the kaon $\langle x^3 \rangle$ and finds $\langle x^3 \rangle_\pi^u = 0.052$, $\langle x^3 \rangle_K^u = 0.048$, and $\langle x^3 \rangle_K^s = 0.092$ at 2 GeV. Another recent DSE study in Ref. [17] calculates $\langle x^3 \rangle_\pi^u = 0.109$, $\langle x^3 \rangle_K^u = 0.092$ and $\langle x^3 \rangle_K^s = 0.143$ at 0.78 GeV. The chiral constituent quark model is used in Ref. [18], which calculates $\langle x^3 \rangle_\pi^u = 0.048$, $\langle x^3 \rangle_K^u = 0.045$, and $\langle x^3 \rangle_K^s = 0.049$ at a scale of 5.2 GeV. Ref. [123] uses the Bethe-Salpeter equation and reports $\langle x^3 \rangle_\pi^u = 0.049(7)$ at 2 GeV. Finally, Ref. [11] calculates the pion and kaon PDFs with a combination of QCD evolution and light front quantization. They also calculate the the moments up to $\langle x^4 \rangle$ and find $\langle x^3 \rangle_\pi^u = 0.057(8)$, $\langle x^3 \rangle_K^u = 0.050(6)$, and $\langle x^3 \rangle_K^s = 0.066(9)$ at 2 GeV.

Table 5.11 and Table 5.12 list the results discussed above for the pion and kaon, respectively. All fo the results are evolved to 2 GeV. An extended list can be found in Ref. [11]. We find that our results are lower than most of the other calculations for the pion and up part of the kaon. Not all of the calculations include systematic uncertainties, however, so a complete comparison is not possible. Our results for the strange part of the kaon are in the middle of the range from model calculations.

Reference	$\langle x^3 \rangle_\pi^u$ (2GeV)	initial scale
This work (lattice)	0.024(18)	2 GeV
Ref. [121] (global fit)	0.058(4)	2 GeV
Ref. [122] (DSE)	0.049	2 GeV
Ref. [113] (global fit)	0.058(4)	5.2 GeV
Ref. [16] (DSE)	0.052	2 GeV
Ref. [18] (CC quark model)	0.061	5.2 GeV
Ref. [17] (DSE)	0.065	0.78 GeV
Ref. [10] (JAM global fit)	0.063	1.3 GeV
Ref. [123] (DSE, BSEs)	0.049(7)	2 GeV
Ref. [11] (BLFQ-NJL)	0.057(8)	2 GeV

Table 5.11: Comparison of $\langle x^3 \rangle_\pi$ with global fits and model calculations in the $\overline{\text{MS}}$ scheme at 2 GeV. The evolution from the reported scale (“initial” scale) to 2 GeV is applied to NNLO. Statistical and systematic uncertainties have been added in quadrature where applicable.

Reference	$\langle x^3 \rangle_K^u$ (2GeV)	$\langle x^3 \rangle_k^s$ (2GeV)	initial scale
This work (lattice)	0.033(6)	0.073(5)	2 GeV
Ref. [16] (DSE)	0.048	0.092	2 GeV
Ref. [18] (CC quark model)	0.058	0.063	5.2 GeV
Ref. [17] (DSE)	0.55	0.85	0.78 GeV
Ref. [11] (BLFQ-NJL)	0.050(7)	0.066(9)	2 GeV

Table 5.12: Comparison of $\langle x^3 \rangle_K^u$ and $\langle x^3 \rangle_K^s$ with global fits and model calculations in the $\overline{\text{MS}}$ scheme at 2 GeV. The evolution from the reported scale (“initial” scale) to 2 GeV is applied to NNLO. Statistical and systematic uncertainties have been added in quadrature where applicable.

5.4 Reconstruction of PDFs

There are many challenges to reconstructing PDFs from their Mellin moments. Besides the increasingly difficult challenges of calculating $\langle x \rangle$, $\langle x^2 \rangle$, and $\langle x^3 \rangle$, which have been explained previously in this chapter, the local derivative operators needed to calculate $\langle x^n \rangle$ for $n \geq 4$ cannot avoid mixing among operators by choosing all indices to be different, as we have done for $n \leq 3$. Because of these difficulties, it has long been argued that it is unfeasible to use Mellin moment to reconstruct PDFs [124, 125]. Computer technologies and solving algorithms have progressed, however, and we have shown in this thesis that it is possible to calculate the first three nontrivial moments with sufficient signals. In this section, we attempt to reconstruct the full x -dependence of the PDFs from only these three moments.

We use for the PDFs the standard functional form

$$q_M^f(x) = Nx^\alpha(1-x)^\beta(1+\rho\sqrt{x}+\gamma x), \quad (5.29)$$

where α , β , γ , and ρ are the parameters we will calculate and N is a normalization factor. We obtain the normalization factor from the trivial moment which, for the connected contributions, is equal to the number of valence quarks

$$\langle 1 \rangle_M = \int_0^1 q_M(x) = 1, \quad (5.30)$$

which results in

$$N = \frac{1}{B(\alpha+1, \beta+1) + \gamma B(2+\alpha, \beta+1)}, \quad (5.31)$$

where $B(x, y)$ is the Euler beta function. The parameter ρ in Eq. 5.29 is generally assumed to be negligibly small [16], so we do not include the \sqrt{x} term in our reconstruction. By integrating Eq. 5.29, we get the general formula for the moments in terms of the parameters

$$\langle x^n \rangle = \frac{\left(\prod_{i=1}^n (i + \alpha) \right) \left(n + 2 + \alpha + \beta + (i + 1 + \alpha)\gamma \right)}{\left(\prod_{i=1}^n (i + 2 + \alpha + \beta) \right) \left(2 + \alpha + \beta + (1 + \alpha)\gamma \right)}. \quad (5.32)$$

We fit our three moments $\langle x \rangle$, $\langle x^2 \rangle$, and $\langle x^3 \rangle$ to Eq. 5.32 for $n = 1, 2$, and 3 , respectively to calculate the parameters. Before fitting the moments, we evolve them to a scale of 5.2 GeV using NNLO expressions so that we can compare them to other calculations of the PDFs.

As a test, we fit our two-state fit moments to a two-parameter fit, where we only consider α and β and set $\gamma = 0$, and to a three-parameter fit, where we consider all three parameters α , β , and γ . In Table 5.13, we list the parameters calculated from each type of fit, along with their statistical uncertainties and their goodness of fit in the form of χ^2 per degree of freedom and compare the reconstructed PDF from each type of fit in Fig. 5.9. We find that including γ in the fit does not affect the shape of the PDFs much but does increase the errors in $q_K^s(x)$. Additionally, the errors associated with γ are large for each fit and so we decide to only look at the two-parameter fit from here on.

fit type	α_π^u	β_π^u	γ_π^u	$\chi^2/\text{d.o.f.}$
2-parameter	-0.04(20)	2.23(65)	0	1.48
3-parameter	-0.54(22)	2.76(64)	22.17(17.87)	2.77
fit type	α_K^u	β_K^u	γ_K^u	$\chi^2/\text{d.o.f.}$
2-parameter	-0.05(7)	2.42(24)	0	2.21
3-parameter	-0.56(72)	3.01(23)	25.11(5.23)	3.78
fit type	α_K^s	β_K^s	γ_K^s	$\chi^2/\text{d.o.f.}$
2-parameter	0.21(8)	2.13(20)	0	0.0036
3-parameter	0.18(95)	2.051(3.46)	0.347(16.10)	0.0018

Table 5.13: The values for the fit parameters, α , β and γ for q_π^u , q_K^u and q_K^s at 5.2 GeV. The error in the parenthesis is statistical.

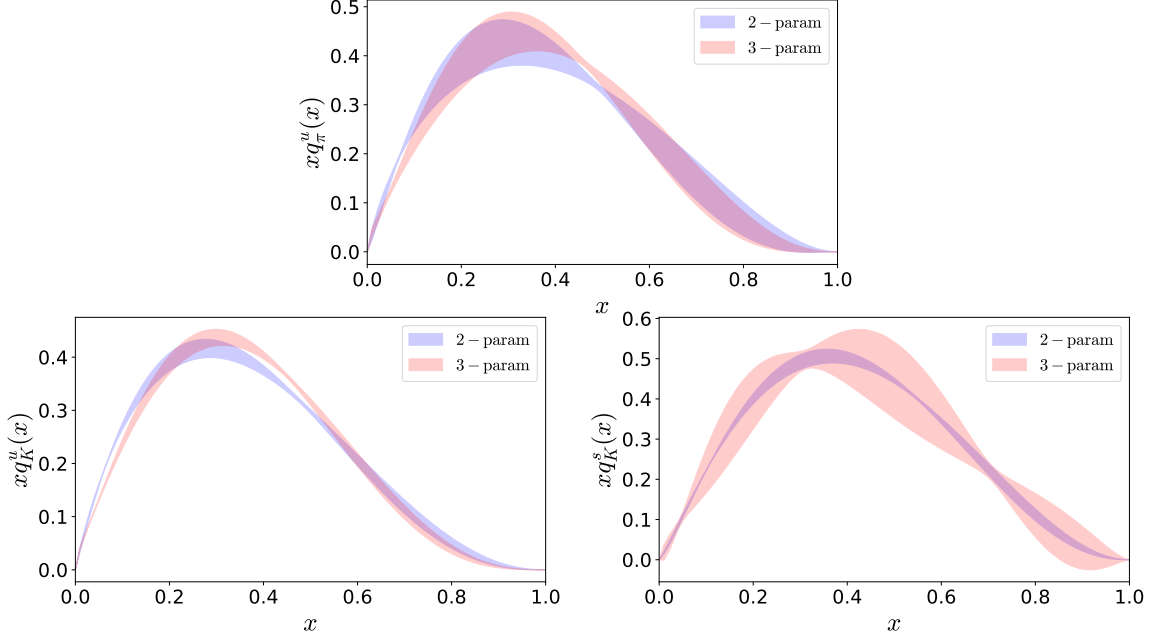


Figure 5.9: Top: The x dependence of $xq_{\pi}^u(x)$ for the 2-parameter (blue band) and 3-parameter (pink band) fits. Bottom: Same as top panel for $xq_K^u(x)$ (left) and $xq_K^s(x)$ (right). Results are given in the $\overline{\text{MS}}$ scheme at 27 GeV^2 .

We also test the excited-state effects on the reconstructed PDFs since α and β do not depend linearly on the moments and so may be nontrivial. In Figure 5.10, we compare the PDFs which are reconstructed by fitting the parameters to the moments at $T_{\text{sink}} = 14a, 16a$, and $18a$, as well as to the two-state fit moments. We find that as T_{sink} increases, the peak of the PDFs decrease and we see good convergence among the higher values of T_{sink} and the two-state fit. As a result of this test, we choose as our final result the PDFs calculated from the two-state fit data.

The last aspect of our reconstructed PDFs we test is their dependence on the maximum moment included in the fit. To do this, we fit up to $\langle x^n \rangle$ for $n = 2, 3, 4$. For the fits up to $n = 2$ and 3 , we use only our lattice data but, since we do not calculate $\langle x^4 \rangle$, we fit Eq. 5.32 for $n = 4$ to the value calculated by global fits and model calculations for the pion and kaon, respectively. Particularly, we use $\langle x^4 \rangle_{\pi}^u = 0.027(2)$ from the JAM analysis [10], and $\langle x^4 \rangle_K^s = 0.029_{-0.004}^{+0.005}$, $\langle x^4 \rangle_K^u = 0.021_{-0.003}^{+0.003}$ from BLFQ-NJL [11]. We include these values only as a test and do not use as a final result any fits obtained from data combined from multiple calculation methods. The results of this test are plotted in Figure 5.11. We find that the resulting PDFs do not depend

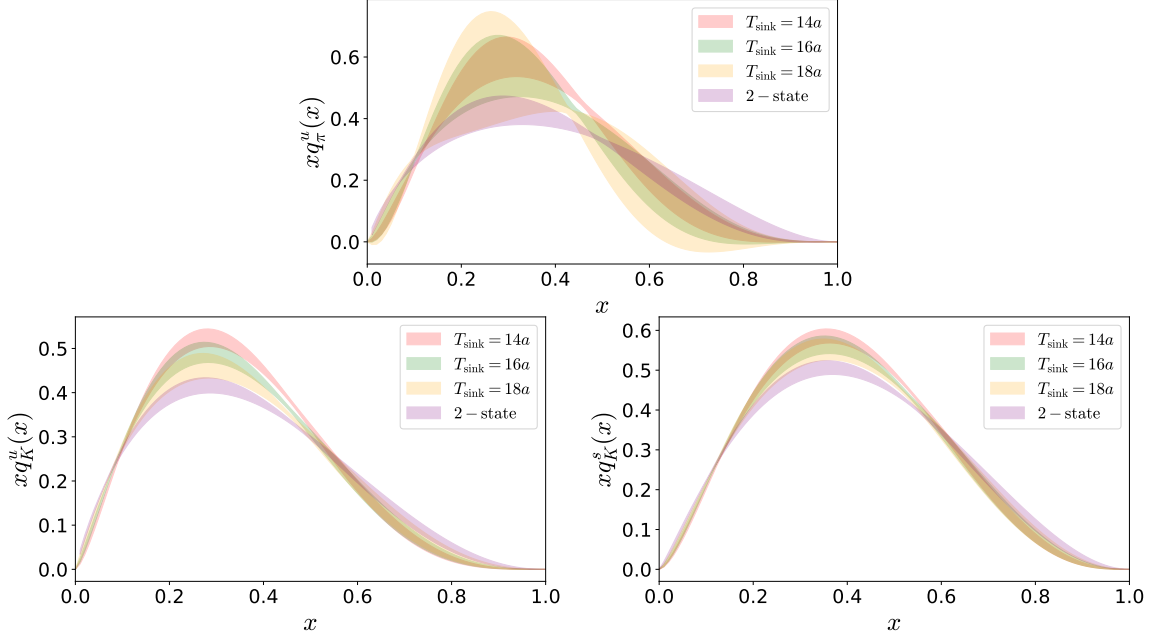


Figure 5.10: Dependence of $xq_{\pi}^u(x)$ (top), $xq_K^u(x)$ (lower left panel) and $xq_K^s(x)$ (lower right panel) on the source-sink time separation. The pink, green, yellow and purple bands correspond to $t_s/a = 14, 16, 18$ and the 2-state fit, respectively. The results are given in the $\overline{\text{MS}}$ scheme at a scale of 27 GeV^2 .

strongly on the number of moments included in the fit so we take as our final result the PDFs calculate with the full set of our lattice data.

A concern about our PDF reconstruction process is whether we are able to get the correct PDF value since we are fitting a limited number of moments. To test this, we perform the same process but use instead the pion PDF moments calculated from the JAM collaboration's global fits [10]. We include the moments from the JAM data up to $n = 3$. In Figure 5.12, we compare the PDF reconstructed from the JAM moments to the actual JAM PDF calculated from global fits. We use bootstrap resampling to calculate the errors on our reconstruction. Our comparison shows that the reconstructed PDF agrees with the actual within the uncertainty. The uncertainty on the reconstructed PDF is larger which is expected due to the information being truncated in the reconstruction. We also integrate the reconstructed PDF to predict the $n = 4$ moment. The moment calculated from the reconstructed PDF is $\langle x^4 \rangle_{\pi}^u =$

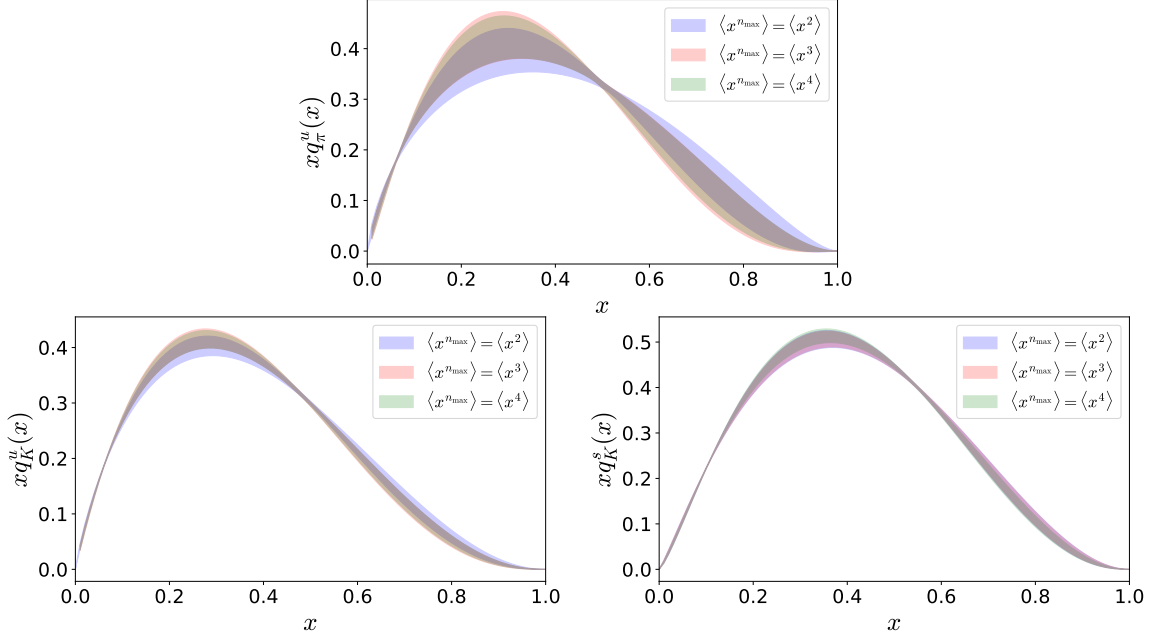


Figure 5.11: Top: The x dependence of $xq_{\pi}^u(x)$ using the 2-parameter with $\langle x^{n_{\max}} \rangle = \langle x^2 \rangle, \langle x^3 \rangle, \langle x^4 \rangle$, shown with blue, pink and green bands. For the green band we use a constraint of $\langle x^4 \rangle_{\pi}^u$ [10]. Bottom: Same as top panel for $xq_K^u(x)$ (left panel) and $xq_K^s(x)$ (right panel) using the BLFQ-NJL [11] $\langle x^4 \rangle_K$ values as constraints for the pink band.

0.026(2) which agrees excellently with the moment obtained from the actual JAM PDF, $\langle x^4 \rangle_{\pi}^u = 0.027(2)$. These results support the fact that the PDF can be correctly reconstructed from only the $n \leq 3$ moments.

To review and summarize the above discussion, our final results are the two-parameter fits on the two-state fit moments for $n = 1, 2, 3$. In Figure 5.13, we study the SU(3) flavor symmetry breaking by comparing these final results for the reconstructed PDFs for both flavors of the kaon to the reconstructed pion PDF. We find that there is agreement between $xq_{\pi}^u(x)$ and $xq_K^u(x)$ for all x regions except for a small amount of tension around $x = 0.5$. Based on our findings, then, the up quark is equally as prevalent in the pion as in the kaon. The up quark distributions are highest in the low to intermediate x regions. The strange is expected to be larger than the pion distribution based on the size of the quark masses and indeed we find that this is the case in the region between $x = 0.3$ and $x = 0.8$. We find the distribution peaks to be $xq_{\pi}^u(x = 0.30) = 0.43(5)$, $xq_{\pi}^u(x = 0.28) = 0.42(2)$, and $xq_{\pi}^u(x = 0.36) = 0.51(2)$.

We integrate over our reconstructed PDFs to calculate the moments for $n = 1-6$. We list the moments calculated from the reconstructed PDFs in Table 5.14 in the

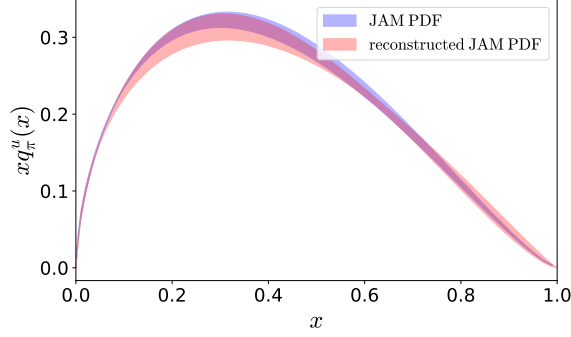


Figure 5.12: Pion PDF using the JAM data (blue band) and the reconstructed PDF using its moments with $n \leq 3$ (pink band). The reported scale is 27 GeV^2 .

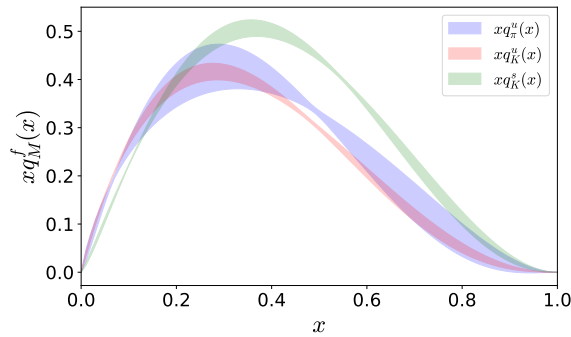


Figure 5.13: Comparison of $xq_\pi^u(x)$ (blue band), $xq_K^u(x)$ (pink band) and $xq_K^s(x)$ (green band) at 27 GeV^2 . The reconstruction uses our lattice data up to $\langle x^3 \rangle$ obtained with the 2-state fits analysis and a 2-parameter fit.

$\overline{\text{MS}}$ scheme at 27 GeV^2 . The first parentheses are the statistical uncertainties and the second are the systematic uncertainties due to excited states. Calculating the moments using this method has two major benefits over calculating them directly from local derivative operators for $n \geq 4$. The first is that this method completely avoids mixing among operators which is unavoidable for the 4- and higher-derivative operators needed for the local operator calculation. The second benefit is that the uncertainties associated with the moments are under control which would be almost impossible to achieve in a local derivative operator calculation.

There is a lot of interest in the high- x behavior of the pion and kaon PDFs because there is some disagreement between the findings of available data and model calculations. The pion PDFs are calculated from an analysis of the Drell-Yan data from the Fermilab E615 experiment [12] which uses evolution equations only to leading order in perturbation theory. The analysis suggests that the up quark distribution

q_M^f	$\langle x \rangle$	$\langle x^2 \rangle$	$\langle x^3 \rangle$
q_π^u	0.230(3)(7)	0.087(5)(8)	0.041(5)(9)
q_K^u	0.217(2)(5)	0.079(2)(1)	0.036(2)(2)
q_K^s	0.279(1)(5)	0.115(2)(6)	0.058(2)(2)
q_M^f	$\langle x^4 \rangle$	$\langle x^5 \rangle$	$\langle x^6 \rangle$
q_π^u	0.023(5)(6)	0.014(4)(5)	0.009(3)(3)
q_K^u	0.019(1)(2)	0.011(1)(2)	0.007(1)(1)
q_K^s	0.033(2)(2)	0.021(1)(2)	0.014(1)(2)

Table 5.14: Values for the first six moments for the pion and kaon. The number shown in the first parenthesis is statistical error, while the number in the second parenthesis is systematic due to excited-states contamination.

falls off like $(1-x)^1$ ($\beta = 1$) in the high- x region. However, a more recent analysis [13] of the Drell-Yan data which also includes next to leading order terms finds that the fall depends on $(1-x)^2$ ($\beta = 2$). DSE results [16] also find a value for β closer to 2. Because of this tension, a calculation of β from LQCD is desirable and some lattice studies [126, 110, 14, 15, 112] which calculate the x dependence of the PDFs have discussed the high- x dependence. As can be seen in Table 5.13 and Figure 5.9, our calculation supports the $(1-x)^2$ high- x behavior.

Finally, we compare our reconstructed PDFs to other PDF studies. This comparison is shown in Figure 5.14. In the left upper panel, we compare with lattice results which use the pseudo-ITD approach [14] and the current-current correlators (LCS) method [15]. The LCS result has borderline agreement with ours while the pseudo-ITD has a lower peak. We also include the two analyses of the Drell-Yan experimental data, E615 [12] and ASV data [13], as well as the global fits of the JAM Collaboration [10, 127]. In the upper right panel, we compare with the two DSE calculations, labeled DSE [16] DSE '18 [17], the chiral constituent quark model (χ CQ) [18], the BLFQ Collaboration results from basis light front quantization, and QCD evolution (NJL) [11]. There are more limited calculations of the kaon PDFs. In the lower panels in Fig. 5.14, we compare our kaon PDFs with the χ CQ results [18], the BLFQ-NJL data [11], and DSE '18 [17] for the up quark (left) and strange quark (right). All results agree with one another in the small- x region ($x < 0.1$) for the pion, though the slope of the χ CQ PDF at small- x is different. In the intermediate to large region ($x > 0.6$) our results are consistent with all other results except the

original E615 data. Most of the disagreement between the different methods is found in the intermediate region. Our pion results and DSE [16] are in agreement for all regions of x , but our peak is higher compared to the updated calculation, DSE'18. Our kaon PDFs agree with the other studies in the small- and large- x regions, while we again see disagreement in the intermediate- x region. There are no experimental kaon data, however, so the comparison is qualitative as all calculations do not fully quantify their systematic uncertainties.

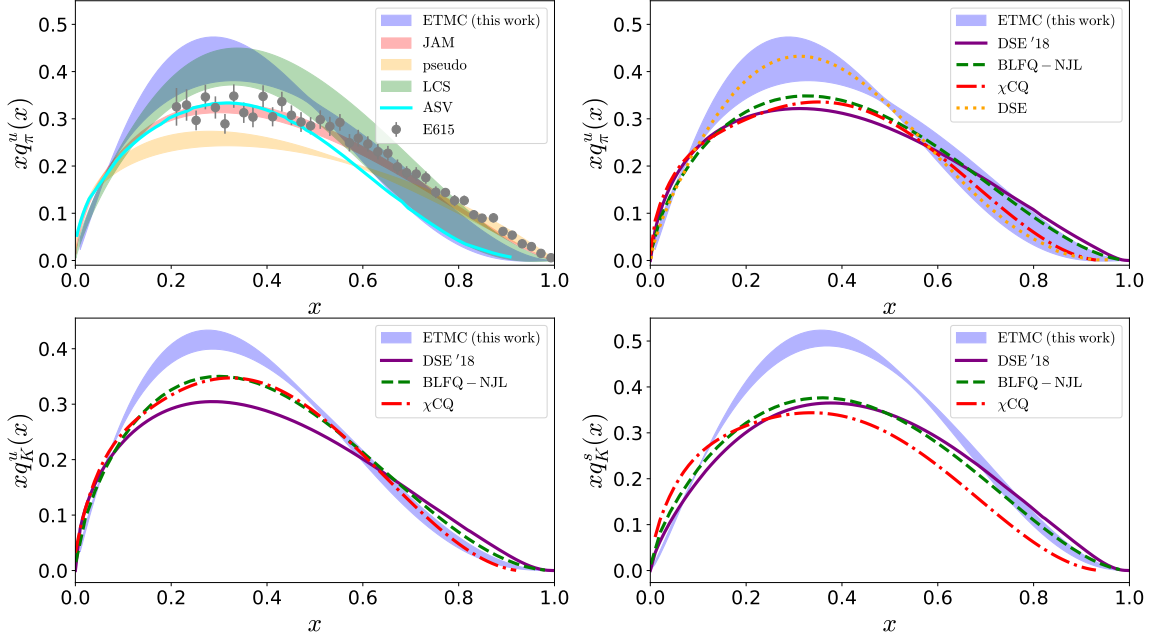


Figure 5.14: Top left panel: Comparison of $xq_{\pi}^u(x)$ with other lattice calculations, experimental data and global analysis, all in the $\overline{\text{MS}}$ at 27 GeV^2 . Our results (blue band) use data up to $\langle x^3 \rangle$ obtained with the 2-state fits analysis and a 2-parameter fit. The E615 data [12] are plotted as gray points and the rescaled ASV data [13] as a solid cyan curve. The JAM global fit is shown with a red band, and the lattice results from pseudo-ITD [14] and current current correlators (LCS) [15] are shown with orange and green band, respectively. Top right panel: Comparison of our results for $xq_{\pi}^u(x)$ with DSE [16] (dotted orange curve), the updated DSE'18 [17] (dot-dashed red curve), BLFQ-NJL [11] (dashed green curve), and χCQ [18] (solid purple curve). Bottom panel: Same as top panel for $xq_K^u(x)$ (left) and $xq_K^s(x)$ (right).

5.5 Form Factors

In this last section, we present a preliminary analysis of the scalar, vector and tensor form factors of the pion and kaon which have information about the internal structure of the particles. In fact, these are considered Mellin moments of GPDs and carry information on the momentum transfer between initial and final states. Also, the scalar form factors carry information related to the generation of the hadron mass from QCD dynamics and the Higgs boson interaction. From the vector form factors, the electromagnetic properties of the mesons, such as their radius and spatial charge distribution, can be calculated. The tensor form factors are related to physics beyond the Standard Model which might include new tensor interactions.

We calculate all three form factors in the rest frame and the vector and tensor form factors in the boosted frames. The form factors are functions of the momentum transfer square $Q^2 = (\mathbf{p}' - \mathbf{p})^2 - (E' - E)^2$. Because the boosted frames have nonzero \mathbf{p}' and E' , we can obtain a denser range of Q^2 from the boosted data. Additionally, some of the form factors in the boosted frame at large Q^2 have lower uncertainties than the form factor in the rest frame at equal or lower Q^2 due to some of the two-point functions in the ratio in Eq. 3.22 being at a lower momentum.

We begin by showing preliminary results of the scalar form factors in Figure 5.15. The form factors are plotted as a function of Q^2 for the up part of the pion (top), up part of the kaon (bottom left), and strange part of the kaon (bottom right) and for $T_{\text{sink}}/a = 12, 14, 16, 18, 20, 24$. Of all the form factors presented in this thesis, the scalar form factors have the largest statistical uncertainties, which is expected due to its coupling with the gauge field. We also find non-negligible excited-state effects and convergence around $T_{\text{sink}} \geq 18a$. We have also produced the necessary data for these form factors in the boosted frames but more work is need for their analysis.

Next, we present our preliminary results of the vector form factors F_π and F_K . F_K is calculated as the combination of its flavor parts

$$F_K(Q^2) = q_u F_K^u(Q^2) + q_s F_K^s(Q^2), \quad (5.33)$$

where of q_u and q_s are the quark charges $q_u = 2/3\epsilon$ and $q_s = -1/3\epsilon$ and ϵ is the fundamental charge. For these form factors, we have preliminary results of the rest frame as well as the boosted frames. In Figure 5.16, we plot the form factors for just the rest frame for all values of T_{sink} as well as for the two-state fit. We find small

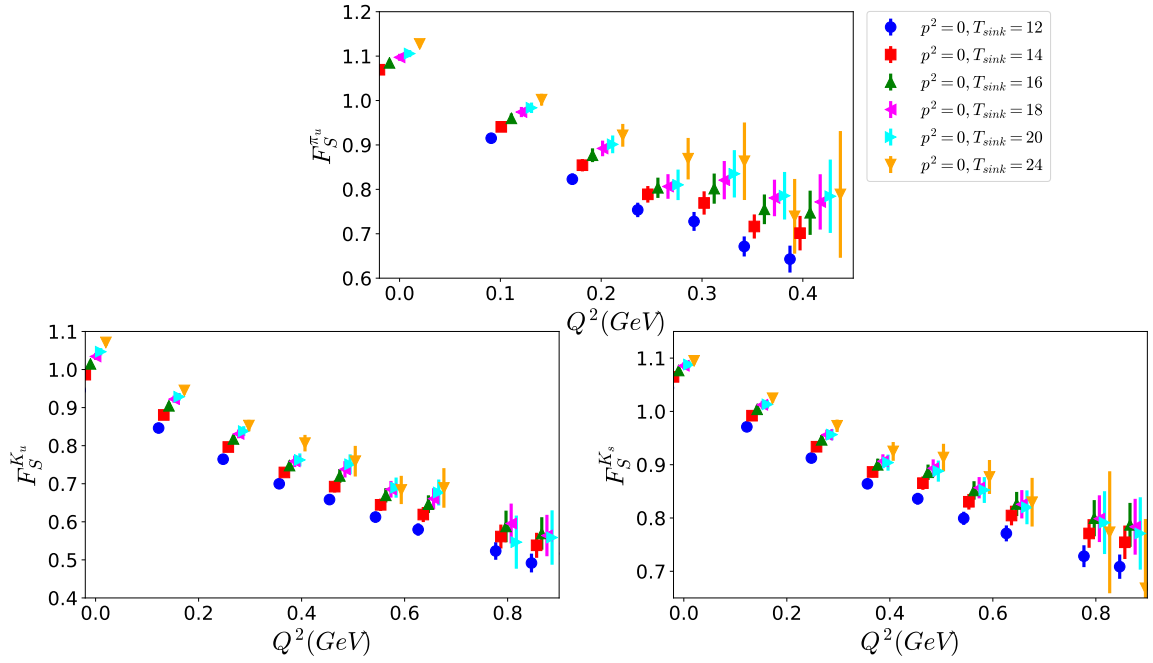


Figure 5.15: Preliminary results of the scalar form factors for the up part pion (top), up part of the kaon (bottom left), and strange part of the kaon (bottom right) plotted as a function of Q^2 . The different source-sink separations are plotted as the colored symbols as in Figure 5.4.

excited-state effects at some Q^2 and a convergence for $T_{\text{sink}} \geq 20a$ and the two-state fit. In Figure 5.17, we compare the rest frame (open symbols) to the boosted frame (closed symbols) and find good agreement between the two.

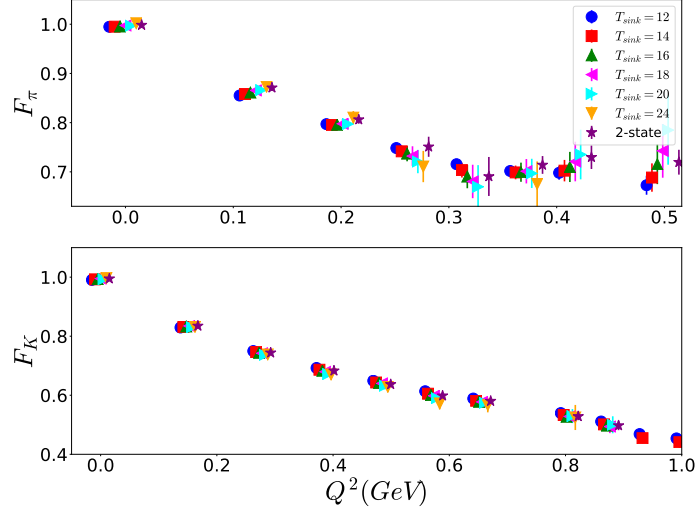


Figure 5.16: Preliminary results of the vector form factors for the pion (top) and kaon (bottom) in the rest frame. In addition to the form factors at the various values of T_{sink} , the values calculated from the two-state fit is also plotted as purple stars.

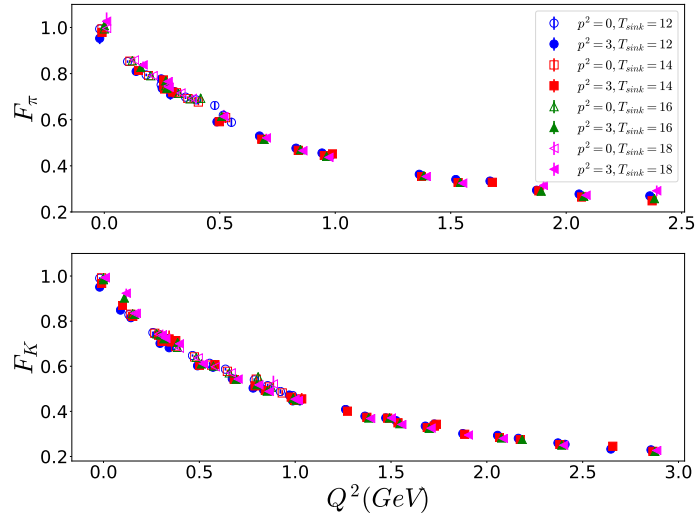


Figure 5.17: Comparison of preliminary results of the vector form factors for the pion (top) and kaon (bottom) between the rest frame (open symbols) and boosted frame (closed symbols).

In Figure 5.18, we plot the preliminary results of the tensor form factors in the rest frame, divided by the meson mass. We include source-sink separations $T_{\text{sink}}/a = 12, 14, 16, 18$ and find non-negligible excited-state effects which are suppressed around $T_{\text{sink}} = 18a$. In Figure 5.19, we compare the tensor form factors, divided by the meson mass, in the two momentum frames. We find that there is some small tension between the two frames, especially in the region of Q^2 that is considered high for the rest frame. This is around 0.6 GeV^2 for the pion and $\sim 1 \text{ GeV}^2$ for the kaon.

We plan to perform a more detailed analysis so that we can better understand the form factors and we stress that these are only preliminary results. For example, the systematic uncertainties due to excited states will be extended. In addition to the three types of form factors presented here, we have also produced the three-point functions needed to calculate the vector one-, two-, and three-derivative generalized form factors but the analysis is still under development. Investigating these quantities will be the next step in this study. In conclusion, having results in the boosted frame has been pioneering for these form factors, as it gives access to much higher values of Q^2 than the rest frame.

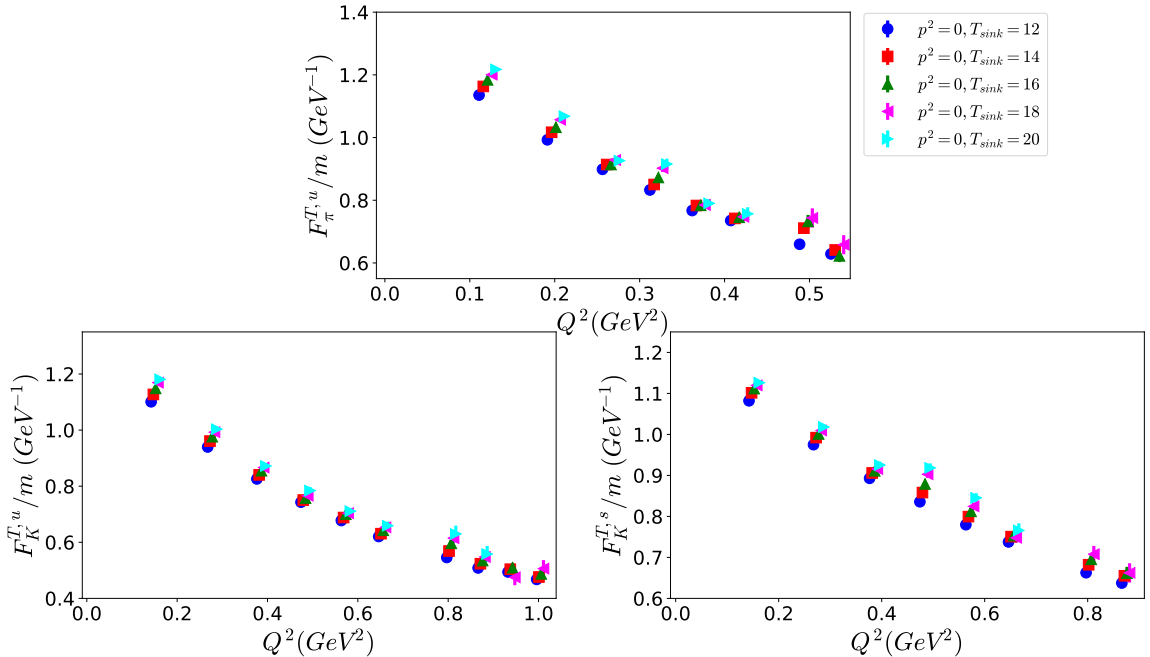


Figure 5.18: Preliminary results of the tensor form factors for the pion (top) and kaon (bottom) in the rest frame.

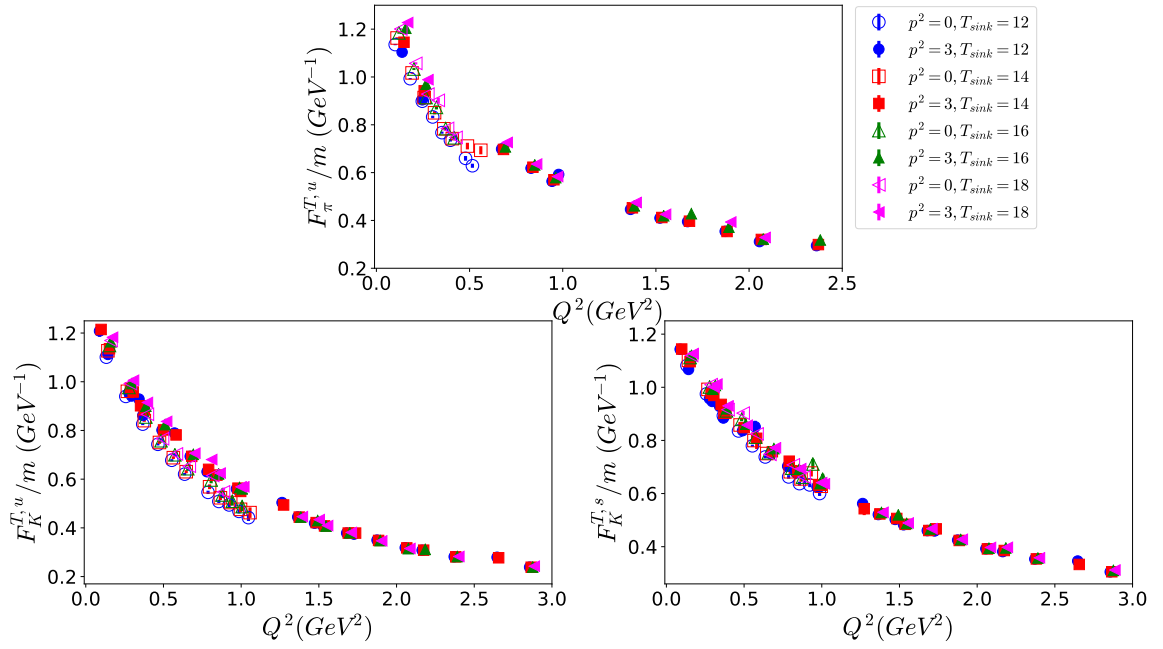


Figure 5.19: Comparison of preliminary results of the tensor form factors for the pion (top) and kaon (bottom) between the rest frame (open symbols) and boosted frame (closed symbols).

CHAPTER 6

CONCLUSION

There have been two main topics of study in this thesis. The first is the calculation of the isovector combinations of the nucleon quark momentum fraction, helicity, and transversity. Three different lattice ensembles were used for these calculations so that we can study the finite volume, discretization, strange and charm quark quenching, and excited-state effects. All three ensembles are at the physical point. Our final results for the nucleon are summarized in Eqs. (4.1–4.3). We compare our results for the quark momentum fraction and helicity, where available, to other lattice calculations at the physical point and to results from global fits and find agreement within 1.5σ of our reported errors. The results for the helicity and transversity are first predictions from LQCD at the physical point.

The second topic of study in this thesis is related to pion and kaon PDFs. We calculate the first three Mellin moments of the meson PDFs, $\langle x \rangle$, $\langle x^2 \rangle$, and $\langle x^3 \rangle$. We avoid mixing with lower dimension operators by carefully choosing the insertion operators in the matrix elements we calculate. This simplifies the renormalization process and ensures our results are free of additional systematic uncertainties which would have been caused by the mixing. Our results are summarized in Eqs. (5.5–5.7) for the first Mellin moments, Eqs. (5.8–5.10) for the second moments, and Eqs. (5.11–5.13) for the third moments. We find that our pion results are consistent with other lattice calculations at similar pion masses as well as with phenomenological results. The calculations of $\langle x^2 \rangle$ and $\langle x^3 \rangle$ are the first of their kind for the kaon. We also study how quickly the PDFs decrease at high- x by calculating the ratios $\langle x \rangle / \langle x^2 \rangle$, $\langle x \rangle / \langle x^3 \rangle$, and $\langle x^2 \rangle / \langle x^3 \rangle$, which are summarized in Eqs. (5.14–5.22). We investigate SU(3) symmetry breaking effects by calculating $\langle x^n \rangle_{u^+}^{\pi} / \langle x^n \rangle_{u^+}^K$ and $\langle x^n \rangle_{u^+}^{\pi} / \langle x^n \rangle_{s^+}^K$ for

the three moments. The results of the symmetry breaking study are summarized in Eqs.(5.23–5.28).

We also show that the full x -dependence of the PDFs can be reconstructed from our results for the first three Mellin moments. This is a very exciting development in this study because it was previously thought that moments up to $\langle x^3 \rangle$ obtained from LQCD could not be feasibly used to reconstruct PDFs. Particularly, it has been argued that the high- x behavior of the PDFs cannot be reliably understood. Our moments are precise enough, however, that our reconstruction can be determined to favor $\sim (1-x)^2$ behavior in the high- x region and we find a β parameter around 2. We are also able to study the SU(3) symmetry breaking by comparing our reconstructed PDFs for the pion with both flavors of the kaon. We find only small differences between the up quark distributions in the pion and kaon and larger symmetry breaking effects between the strange distribution in the kaon and up quark in the pion.

Looking to the future of this study, we intend to perform a complete analysis of the pion and kaon form factors. We have shown preliminary results of the scalar, vector, and tensor form factors and plan to also study the SU(3) symmetry breaking in the vector form factors. Lastly, we will calculate the vector derivative generalized form factors, the analysis for which is currently under development.

BIBLIOGRAPHY

- [1] Jacob J. Ethier and Emanuele R. Nocera. Parton distributions in nucleons and nuclei. *Annual Review of Nuclear and Particle Science*, 70(1):43–76, Oct 2020.
- [2] C. Alexandrou, S. Bacchio, M. Constantinou, P. Dimopoulos, J. Finkenrath, R. Frezzotti, K. Hadjiyiannakou, K. Jansen, B. Kostrzewa, G. Koutsou, and et al. Moments of nucleon generalized parton distributions from lattice qcd simulations at physical pion mass. *Physical Review D*, 101(3), Feb 2020.
- [3] J.R. Green, M. Engelhardt, S. Krieg, J.W. Negele, A.V. Pochinsky, and S.N. Syritsyn. Nucleon structure from lattice qcd using a nearly physical pion mass. *Physics Letters B*, 734:290–295, Jun 2014.
- [4] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Luigi Del Debbio, Stefano Forte, Patrick Groth-Merrild, Alberto Guffanti, Nathan P. Hartland, Zahari Kassabov, José I. Latorre, and et al. Parton distributions from high-precision collider data. *The European Physical Journal C*, 77(10), Oct 2017.
- [5] Sayipjamal Dulat, Tie-Jiun Hou, Jun Gao, Marco Guzzi, Joey Huston, Pavel Nadolsky, Jon Pumplin, Carl Schmidt, Daniel Stump, and C.-P. Yuan. New parton distribution functions from a global analysis of quantum chromodynamics. *Physical Review D*, 93(3), Feb 2016.
- [6] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne. Parton distributions in the lhc era: Mmht 2014 pdfs. *The European Physical Journal C*, 75(5), May 2015.
- [7] S. Alekhin, J. Blümlein, S. Moch, and R. Plačákytė. Parton distribution functions, α_s , and heavy-quark masses for lhc run ii. *Physical Review D*, 96(1), Jul 2017.

- [8] A. Accardi, L. T. Brady, W. Melnitchouk, J. F. Owens, and N. Sato. Constraints on large-xparton distributions from new weak boson production and deep-inelastic scattering data. *Physical Review D*, 93(11), Jun 2016.
- [9] H1 and ZEUS Collaborations. Combination of measurements of inclusive deep inelastic $e^\pm p$ scattering cross sections and qcd analysis of heraf data, 2015.
- [10] P.C. Barry, N. Sato, W. Melnitchouk, and Chueng-Ryong Ji. First Monte Carlo Global QCD Analysis of Pion Parton Distributions. *Phys. Rev. Lett.*, 121(15):152001, 2018.
- [11] Jiangshan Lan, Chandan Mondal, Shaoyang Jia, Xingbo Zhao, and James P. Vary. Pion and kaon parton distribution functions from basis light front quantization and QCD evolution. *Phys. Rev. D*, 101(3):034024, 2020.
- [12] J. S. Conway et al. Experimental Study of Muon Pairs Produced by 252-GeV Pions on Tungsten. *Phys. Rev.*, D39:92–122, 1989.
- [13] Matthias Aicher, Andreas Schafer, and Werner Vogelsang. Soft-gluon resummation and the valence parton distribution function of the pion. *Phys. Rev. Lett.*, 105:252003, 2010.
- [14] Bálint Joó, Joseph Karpie, Kostas Orginos, Anatoly V. Radyushkin, David G. Richards, Raza Sabbir Sufian, and Savvas Zafeiropoulos. Pion Valence Structure from Ioffe Time Pseudo-Distributions. *Phys. Rev.*, D100(11):114512, 2019.
- [15] Raza Sabbir Sufian, Joseph Karpie, Colin Egerer, Kostas Orginos, Jian-Wei Qiu, and David G. Richards. Pion Valence Quark Distribution from Matrix Element Calculated in Lattice QCD. *Phys. Rev.*, D99(7):074507, 2019.
- [16] Chen Chen, Lei Chang, Craig D. Roberts, Shaolong Wan, and Hong-Shi Zong. Valence-quark distribution functions in the kaon and pion. *Phys. Rev. D*, 93(7):074021, 2016.
- [17] Kyle D. Bednar, Ian C. Cloët, and Peter C. Tandy. Distinguishing quarks and gluons in pion and kaon parton distribution functions. *Physical Review Letters*, 124(4), Jan 2020.

- [18] Akira Watanabe, Takahiro Sawada, and Chung Wen Kao. Kaon quark distribution functions in the chiral constituent quark model. *Phys. Rev. D*, 97(7):074015, 2018.
- [19] C. Alexandrou, S. Bacchio, M. Constantinou, J. Finkenrath, K. Hadjiyianakou, K. Jansen, G. Koutsou, and A. Vaquero Aviles-Casco. Proton and neutron electromagnetic form factors from lattice qcd. *Physical Review D*, 100(1), Jul 2019.
- [20] National Academies of Sciences Engineering and Medicine. *An Assessment of U.S.-Based Electron-Ion Collider Science*. The National Academies Press, Washington, DC, 2018.
- [21] Arlene C. Aguilar et al. Pion and Kaon Structure at the Electron-Ion Collider. *Eur. Phys. J. A*, 55(10):190, 2019.
- [22] M.E. Peskin and D.V. Schroeder. *An Introduction To Quantum Field Theory*. Frontiers in Physics. Avalon Publishing, 1995.
- [23] Huey-Wen Lin, Emanuele R. Nocera, Fred Olness, Kostas Orginos, Juan Rojo, Alberto Accardi, Constantia Alexandrou, Alessandro Bacchetta, Giuseppe Bozzi, Jiunn-Wei Chen, and et al. Parton distributions and lattice qcd calculations: A community white paper. *Progress in Particle and Nuclear Physics*, 100:107–160, May 2018.
- [24] E Perez and E Rizvi. The quark and gluon structure of the proton. *Reports on Progress in Physics*, 76(4):046201, Mar 2013.
- [25] A. De Roeck and R.S. Thorne. Structure functions. *Progress in Particle and Nuclear Physics*, 66(4):727–781, Oct 2011.
- [26] Sergey Alekhin, Simone Alioli, Richard D. Ball, Valerio Bertone, Johannes Blumlein, Michiel Botje, Jon Butterworth, Francesco Cerutti, Amanda Cooper-Sarkar, Albert de Roeck, Luigi Del Debbio, Joel Feltesse, Stefano Forte, Alexander Glazov, Alberto Guffanti, Claire Gwenlan, Joey Huston, Pedro Jimenez-Delgado, Hung-Liang Lai, Jose I. Latorre, Ronan McNulty, Pavel Nadolsky, Sven-Olaf Moch, Jon Pumplin, Voica Radescu, Juan Rojo, Torbjorn Sjostrand, W. J. Stirling, Daniel Stump, Robert S. Thorne, Maria Ubiali, Alessandro

- Vicini, Graeme Watt, and C. P. Yuan. The pdf4lhc working group interim report, 2011.
- [27] Emanuele R. Nocera, Richard D. Ball, Stefano Forte, Giovanni Ridolfi, and Juan Rojo. A first unbiased global determination of polarized pdfs and their uncertainties. *Nuclear Physics B*, 887:276–308, Oct 2014.
- [28] Martha Constantinou, Aurore Courtoy, Markus A. Ebert, Michael Engelhardt, Tommaso Giani, Tim Hobbs, Tie-Jiun Hou, Aleksander Kusina, Krzysztof Kutak, Jian Liang, Huey-Wen Lin, Keh-Fei Liu, Simonetta Liuti, Cédric Mezrag, Pavel Nadolsky, Emanuele R. Nocera, Fred Olness, Jian-Wei Qiu, Marco Radici, Anatoly Radyushkin, Abha Rajan, Ted Rogers, Juan Rojo, Gerrit Schierholz, C. P. Yuan, Jian-Hui Zhang, and Rui Zhang. Parton distributions and lattice qcd calculations: toward 3d structure, 2020.
- [29] Xiangdong Ji. Qcd analysis of the mass structure of the nucleon. *Phys. Rev. Lett.*, 74:1071–1074, Feb 1995.
- [30] Xiangdong Ji. Breakup of hadron masses and the energy-momentum tensor of qcd. *Phys. Rev. D*, 52:271–281, Jul 1995.
- [31] Xiangdong Ji. Gauge-invariant decomposition of nucleon spin. *Phys. Rev. Lett.*, 78:610–613, Jan 1997.
- [32] Xiangdong Ji. Deeply virtual compton scattering. *Physical Review D*, 55(11):7114–7125, Jun 1997.
- [33] A. V. Radyushkin. Nonforward parton distributions. *Physical Review D*, 56(9):5524–5557, Nov 1997.
- [34] Markus Diehl. Introduction to gpds and tmds. *The European Physical Journal A*, 52(6), Jun 2016.
- [35] R.P. Feynman, A.R. Hibbs, and D.F. Styer. *Quantum Mechanics and Path Integrals*. Dover Books on Physics. Dover Publications, 2010.
- [36] Kenneth G. Wilson. Confinement of quarks. *Phys. Rev. D*, 10:2445–2459, Oct 1974.

- [37] Christof Gattringer and Christian B. Lang. *Quantum chromodynamics on the lattice*, volume 788. Springer, Berlin, 2010.
- [38] Jean Zinn-Justin. Quantum field theory and critical phenomena. *Int. Ser. Monogr. Phys.*, 113:1–1054, 2002.
- [39] T. Matthews and A. Salam. The Green’s functions of quantised fields. *Nuovo Cim.*, 12:563—565, 1954.
- [40] T. Matthews and A. Salam. Propagators of quantized field. *Nuovo Cim.*, 2:120—134, 1955.
- [41] R. Barret et al. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*. SIAM, Philadelphia, 1994.
- [42] P. Wesseling. *An Introduction to Multigrid Methods*. An Introduction to Multigrid Methods. R.T. Edwards, 2004.
- [43] M. Luscher and P. Weisz. On-Shell Improved Lattice Gauge Theories. *Commun. Math. Phys.*, 97:59, 1985. [Erratum: *Commun.Math.Phys.* 98, 433 (1985)].
- [44] Ph. de Forcrand, M. Garcíá Pérez, T. Hashimoto, S. Hioki, H. Matsufuru, O. Miyamura, A. Nakamura, I.-O. Stamatescu, T. Takaishi, and T. Umeda. Renormalization group flow of $su(3)$ lattice gauge theory. numerical studies in a two coupling space. *Nuclear Physics B*, 577(1-2):263–278, Jun 2000.
- [45] Y. Iwasaki. Renormalization Group Analysis of Lattice Theories and Improved Lattice Action: Two-Dimensional Nonlinear $O(N)$ Sigma Model. *Nucl. Phys. B*, 258:141–156, 1985.
- [46] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. P. Lepage, J. Shigemitsu, H. Trotter, and K. Wong. Highly improved staggered quarks on the lattice with applications to charm physics. *Physical Review D*, 75(5), Mar 2007.
- [47] Yigal Shamir. Chiral fermions from lattice boundaries. *Nuclear Physics B*, 406(1-2):90–106, Sep 1993.
- [48] Kenneth G. Wilson. Quarks and Strings on a Lattice. In *13th International School of Subnuclear Physics: New Phenomena in Subnuclear Physics*, page 99, 11 1975.

- [49] R. Frezzotti, P. A. Grassi, S. Sint, and P. Weisz. Lattice QCD with a chirally twisted mass term. *JHEP*, 08:058, 2001.
- [50] Andrea Shindler. Twisted mass lattice qcd. *Physics Reports*, 461(2-3):37–110, May 2008.
- [51] R Frezzotti and G.C Rossi. Chirally improving wilson fermions i. o(a) improvement. *Journal of High Energy Physics*, 2004(08):007–007, Aug 2004.
- [52] K. Symanzik. Continuum limit and improved action in lattice theories: (i). principles and ϕ^4 theory. *Nuclear Physics B*, 226(1):187–204, 1983.
- [53] K. Symanzik. Continuum limit and improved action in lattice theories: (ii). o(n) non-linear sigma model in perturbation theory. *Nuclear Physics B*, 226(1):205–227, 1983.
- [54] B. Sheikholeslami and R. Wohlert. Improved continuum limit lattice action for qcd with wilson fermions. *Nuclear Physics B*, 259(4):572–596, 1985.
- [55] D. Bećirević, Ph. Boucaud, V. Lubicz, G. Martinelli, F. Mescia, S. Simula, and C. Tarantino. Exploring twisted mass lattice qcd with the clover term. *Phys. Rev. D*, 74:034501, Aug 2006.
- [56] R.A. Howard. *Dynamic Probabilistic Systems, Volume I: Markov Models*. Dover Books on Mathematics. Dover Publications, 2007.
- [57] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.
- [58] Simon Duane and John B. Kogut. Hybrid stochastic differential equations applied to quantum chromodynamics. *Phys. Rev. Lett.*, 55:2774–2777, Dec 1985.
- [59] S. Duane and J. B. Kogut. The Theory of Hybrid Stochastic Algorithms. *Nucl. Phys. B*, 275:398–420, 1986.
- [60] David J. E. Callaway and Aneesur Rahman. Lattice gauge theory in the microcanonical ensemble. *Phys. Rev. D*, 28:1506–1514, Sep 1983.

- [61] J. Polonyi and H. W. Wyld. Microcanonical simulation of fermionic systems. *Phys. Rev. Lett.*, 51:2257–2260, Dec 1983.
- [62] D.H. Weingarten and D.N. Petcher. Monte carlo integration for lattice gauge theories with fermions. *Physics Letters B*, 99(4):333–338, 1981.
- [63] M. Albanese et al. Glueball Masses and String Tension in Lattice QCD. *Phys. Lett. B*, 192:163–169, 1987.
- [64] G. Martinelli, C. Pittori, Christopher T. Sachrajda, M. Testa, and A. Vladikas. A General method for nonperturbative renormalization of lattice operators. *Nucl. Phys.*, B445:81–108, 1995.
- [65] M. Göckeler, R. Horsley, H. Oelrich, H. Perlt, D. Petters, Paul E. L. Rakow, A. Schäfer, G. Schierholz, and A. Schiller. Nonperturbative renormalization of composite operators in lattice QCD. *Nucl. Phys.*, B544:699–733, 1999.
- [66] C. Alexandrou, M. Constantinou, T. Korzec, H. Panagopoulos, and F. Stylianou. Renormalization constants for 2-twist operators in twisted mass QCD. *Phys. Rev.*, D83:014503, 2011.
- [67] C. Alexandrou, M. Constantinou, T. Korzec, H. Panagopoulos, and F. Stylianou. Renormalization constants of local operators for Wilson type improved fermions. *Phys.Rev.*, D86:014505, 2012.
- [68] Constantia Alexandrou, Martha Constantinou, and Haralambos Panagopoulos. Renormalization functions for $N_f=2$ and $N_f=4$ twisted mass fermions. *Phys. Rev.*, D95(3):034505, 2017.
- [69] M. Constantinou et al. Non-perturbative renormalization of quark bilinear operators with $N_f=2$ (tmQCD) Wilson fermions and the tree-level improved gauge action. *JHEP*, 08:068, 2010.
- [70] M. Constantinou, R. Horsley, H. Panagopoulos, H. Perlt, P. E. L. Rakow, G. Schierholz, A. Schiller, and J. M. Zanotti. Renormalization of local quark-bilinear operators for $N_f=3$ flavors of stout link nonperturbative clover fermions. *Phys. Rev.*, D91(1):014502, 2015.

- [71] A. Abdel-Rehim, C. Alexandrou, F. Burger, M. Constantinou, P. Dimopoulos, R. Frezzotti, K. Hadjiyiannakou, C. Helmes, K. Jansen, C. Jost, and et al. First physics results at the physical pion mass from nf=2 wilson twisted mass fermions at maximal twist. *Physical Review D*, 95(9), May 2017.
- [72] Constantia Alexandrou et al. Simulating twisted mass fermions at physical light, strange and charm quark masses. *Phys. Rev.*, D98(5):054518, 2018.
- [73] Johannes Blümlein and Helmut Böttcher. Qcd analysis of polarized deep inelastic scattering data. *Nuclear Physics B*, 841(1-2):205–230, Dec 2010.
- [74] Daniel de Florian, Rodolfo Sassot, Marco Stratmann, and Werner Vogelsang. Extraction of spin-dependent parton densities and their uncertainties. *Physical Review D*, 80(3), Aug 2009.
- [75] J. J. Ethier, N. Sato, and W. Melnitchouk. First simultaneous extraction of spin-dependent parton distributions and fragmentation functions from a global qcd analysis. *Physical Review Letters*, 119(13), Sep 2017.
- [76] Anthony William Thomas, W. Melnitchouk, and Fernanda Monti Steffens. Dynamical symmetry breaking in the sea of the nucleon. *Phys. Rev. Lett.*, 85:2892–2894, 2000.
- [77] Jiunn-Wei Chen and Xiang-dong Ji. Large N(c) quark distributions in the delta and chiral logarithms in quark distributions of the nucleon. *Phys. Lett.*, B523:73–78, 2001.
- [78] Jiunn-Wei Chen and Xiang-dong Ji. Leading chiral contributions to the spin structure of the proton. *Phys. Rev. Lett.*, 88:052003, 2002.
- [79] Yusupujiang Salamu, Chueng-Ryong Ji, W. Melnitchouk, and P. Wang. $\bar{d} - \bar{u}$ asymmetry in the proton in chiral effective theory. *Phys. Rev. Lett.*, 114:122001, 2015.
- [80] Parada T. P. Hutauruk, Ian C. Cloet, and Anthony W. Thomas. Flavor dependence of the pion and kaon form factors and parton distribution functions. *Phys. Rev. C*, 94(3):035201, 2016.
- [81] P. A. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8):083C01, 2020.

- [82] J. Badier et al. Experimental Determination of the pi Meson Structure Functions by the Drell-Yan Mechanism. *Z. Phys.*, C18:281, 1983.
- [83] B. Betev et al. Differential Cross-section of High Mass Muon Pairs Produced by a 194-GeV/ $c\pi^-$ Beam on a Tungsten Target. *Z. Phys.*, C28:9, 1985.
- [84] S. R. Amendolia et al. A Measurement of the Space - Like Pion Electromagnetic Form-Factor. *Nucl. Phys.*, B277:168, 1986.
- [85] F. D. Aaron et al. Measurement of Leading Neutron Production in Deep-Inelastic Scattering at HERA. *Eur. Phys. J.*, C68:381–399, 2010.
- [86] A. J. Chambers et al. Electromagnetic form factors at large momenta from lattice QCD. *Phys. Rev.*, D96(11):114509, 2017.
- [87] Jonna Koponen, Francis Bursa, Christine T. H. Davies, Rachel J. Dowdall, and G. Peter Lepage. Pion electromagnetic form factor from full lattice QCD. *PoS, LATTICE2015:119*, 2015.
- [88] C. Alexandrou et al. Pion vector form factor from lattice QCD at the physical point. *Phys. Rev.*, D97(1):014508, 2018.
- [89] Gen Wang, Jian Liang, Terrence Draper, Keh-Fei Liu, and Yi-Bo Yang. Pion Form Factor with Overlap Fermion. In *36th International Symposium on Lattice Field Theory (Lattice 2018) East Lansing, MI, United States, July 22-28, 2018*, 2018.
- [90] Benjamin Owen, Waseem Kamleh, Derek Leinweber, Benjamin Menadue, and Selim Mahbub. Light Meson Form Factors at near Physical Masses. *Phys. Rev. D*, 91(7):074503, 2015.
- [91] Dirk Brommel et al. Quark distributions in the pion. *PoS, LATTICE2007:140*, 2007.
- [92] D. Brömmel et al. The Pion form-factor from lattice QCD with two dynamical flavours. *Eur. Phys. J. C*, 51:335–345, 2007.
- [93] G. Bali, S. Collins, B. Glässle, M. Göckeler, N. Javadi-Motaghi, J. Najjar, W. Söldner, and A. Sternbeck. Pion structure from lattice QCD. *PoS, LATTICE2013:447*, 2014.

- [94] A. Abdel-Rehim et al. Nucleon and pion structure with lattice QCD simulations at physical value of the pion mass. *Phys. Rev. D*, 92(11):114513, 2015. [Erratum: *Phys.Rev.D* 93, 039904 (2016)].
- [95] M. Oehm, C. Alexandrou, M. Constantinou, K. Jansen, G. Koutsou, B. Kostrzewa, F. Steffens, C. Urbach, and S. Zafeiropoulos. $\langle x \rangle$ and $\langle x^2 \rangle$ of the pion PDF from lattice QCD with $N_f = 2 + 1 + 1$ dynamical quark flavors. *Phys. Rev. D*, 99(1):014508, 2019.
- [96] Lei Chang, I. C. Clöet, J. J. Cobos-Martinez, C. D. Roberts, S. M. Schmidt, and P. C. Tandy. Imaging dynamical chiral symmetry breaking: pion wave function on the light front. *Phys. Rev. Lett.*, 110(13):132001, 2013.
- [97] L. Chang, I. C. Clöet, C. D. Roberts, S. M. Schmidt, and P. C. Tandy. Pion electromagnetic form factor at spacelike momenta. *Phys. Rev. Lett.*, 111(14):141802, 2013.
- [98] Chao Shi and Ian C. Cloët. Intrinsic transverse motion of the pion’s valence quarks. *Physical Review Letters*, 122(8), Mar 2019.
- [99] Y. Ninomiya, W. Bentz, and I. C. Clöet. Dressed Quark Mass Dependence of Pion and Kaon Form Factors. *Phys. Rev.*, C91(2):025202, 2015.
- [100] Xiangdong Ji. Parton Physics on a Euclidean Lattice. *Phys.Rev.Lett.*, 110:262002, 2013.
- [101] Xiang Gao, Luchang Jin, Christos Kallidonis, Nikhil Karthik, Swagato Mukherjee, Peter Petreczky, Charles Shugert, Sergey Syritsyn, and Yong Zhao. Valence parton distribution of the pion from lattice qcd: Approaching the continuum limit. *Physical Review D*, 102(9), Nov 2020.
- [102] A. V. Radyushkin. Quasi-parton distribution functions, momentum distributions, and pseudo-parton distribution functions. *Phys. Rev. D*, 96(3):034025, 2017.
- [103] Bálint Joó, Joseph Karpie, Kostas Orginos, Anatoly V. Radyushkin, David G. Richards, and Savvas Zafeiropoulos. Parton Distribution Functions from Ioffe Time Pseudodistributions from Lattice Calculations: Approaching the Physical Point. *Phys. Rev. Lett.*, 4 2020.

- [104] Yan-Qing Ma and Jian-Wei Qiu. Exploring Partonic Structure of Hadrons Using ab initio Lattice QCD Calculations. *Phys. Rev. Lett.*, 120(2):022003, 2018.
- [105] Raza Sabbir Sufian, Colin Egerer, Joseph Karpie, Robert G. Edwards, Bálint Joó, Yan-Qing Ma, Kostas Orginos, Jian-Wei Qiu, and David G. Richards. Pion Valence Quark Distribution from Current-Current Correlation in Lattice QCD. *Phys. Rev. D*, 102(5):054508, 2020.
- [106] Constantia Alexandrou, Simone Bacchio, Ian Cloët, Martha Constantinou, Kyriakos Hadjiyiannakou, Giannis Koutsou, and Colin Lauer. Mellin moments $\langle x \rangle$ and $\langle x^2 \rangle$ for the pion and kaon from lattice qcd. *Physical Review D*, 103(1), Jan 2021.
- [107] Dirk Brommel. *Pion Structure from the Lattice*. PhD thesis, Regensburg U., 2007.
- [108] Remi Baron, Stefano Capitani, Jaume Carbonell, Karl Jansen, Zhaofeng Liu, Olivier Pene, and Carsten Urbach. Moments of meson distribution functions with dynamical twisted mass fermions. *PoS, LATTICE2007*:153, 2007.
- [109] Joseph Karpie, Kostas Orginos, and Savvas Zafeiropoulos. Moments of Ioffe time parton distribution functions from non-local matrix elements. *JHEP*, 11:178, 2018.
- [110] Taku Izubuchi, Luchang Jin, Christos Kallidonis, Nikhil Karthik, Swagato Mukherjee, Peter Petreczky, Charles Shugert, and Sergey Syritsyn. Valence parton distribution function of pion from fine lattice. *Phys. Rev.*, D100(3):034516, 2019.
- [111] Gunnar S. Bali, Vladimir M. Braun, Simon Bürger, Meinulf Göckeler, Michael Gruber, Fabian Hutzler, Piotr Korcyl, Andreas Schäfer, André Sternbeck, and Philipp Wein. Light-cone distribution amplitudes of pseudoscalar mesons from lattice QCD. *JHEP*, 08:065, 2019.
- [112] Huey-Wen Lin, Jiunn-Wei Chen, Zhouyou Fan, Jian-Hui Zhang, and Rui Zhang. Valence-quark distribution of the kaon and pion from lattice qcd. *Physical Review D*, 103(1), Jan 2021.

- [113] K. Wijesooriya, P.E. Reimer, and R.J. Holt. The pion parton distribution function in the valence region. *Phys. Rev. C*, 72:065203, 2005.
- [114] C. Best, M. Gockeler, R. Horsley, Ernst-Michael Ilgenfritz, H. Perlt, Paul E. L. Rakow, A. Schafer, G. Schierholz, A. Schiller, and S. Schramm. Pion and rho structure functions from lattice QCD. *Phys. Rev. D*, 56:2743–2754, 1997.
- [115] William Detmold, W. Melnitchouk, and Anthony William Thomas. Parton distribution functions in the pion from lattice QCD. *Phys. Rev. D*, 68:034025, 2003.
- [116] Xiangdong Ji. Parton Physics from Large-Momentum Effective Field Theory. *Sci. China Phys. Mech. Astron.*, 57:1407–1412, 2014.
- [117] Anatoly Radyushkin. Nonperturbative Evolution of Parton Quasi-Distributions. *Phys. Lett.*, B767:314–320, 2017.
- [118] Yan-Qing Ma and Jian-Wei Qiu. Extracting Parton Distribution Functions from Lattice QCD Calculations. *Phys. Rev.*, D98(7):074021, 2018.
- [119] Yan-Qing Ma and Jian-Wei Qiu. QCD Factorization and PDFs from Lattice QCD Calculation. *Int. J. Mod. Phys. Conf. Ser.*, 37:1560041, 2015.
- [120] G. Martinelli. Hadronic weak interactions of light quarks. *Nucl. Phys. B Proc. Suppl.*, 73:58–71, 1999.
- [121] P. J. Sutton, Alan D. Martin, R. G. Roberts, and W. James Stirling. Parton distributions for the pion extracted from Drell-Yan and prompt photon experiments. *Phys. Rev. D*, 45:2349–2359, 1992.
- [122] M. B. Hecht, Craig D. Roberts, and S. M. Schmidt. Valence quark distributions in the pion. *Phys. Rev. C*, 63:025213, 2001.
- [123] Minghui Ding, Khépani Raya, Daniele Binosi, Lei Chang, Craig D Roberts, and Sebastian M Schmidt. Drawing insights from pion parton distributions. *Chin. Phys. C*, 44(3):031002, 2020.
- [124] William Detmold, W. Melnitchouk, and Anthony William Thomas. Parton distributions from lattice QCD. *Eur. Phys. J. direct*, 3(1):13, 2001.

- [125] Roy J. Holt and Craig D. Roberts. Nucleon and pion distribution functions in the valence region. *Reviews of Modern Physics*, 82(4):2991–3044, Oct 2010.
- [126] Jian-Hui Zhang, Jiunn-Wei Chen, Luchang Jin, Huey-Wen Lin, Andreas Schäfer, and Yong Zhao. First direct lattice-QCD calculation of the x -dependence of the pion parton distribution function. *Phys. Rev.*, D100(3):034505, 2019.
- [127] N. Y. Cao, P. C. Barry, N. Sato, and W. Melnitchouk. Towards the 3-dimensional parton structure of the pion: integrating transverse momentum data into global qcd analysis, 2021.

APPENDIX A

DECOMPOSITION EQUATIONS

We give the decomposition equations which relate the matrix elements to the quantities of interest. For the nucleon the decompositions are given only for the case of rest frame and zero momentum frame. For the mesons, the case of general momentum frame and general momentum transfer is given and then the combinations of also the rest frame, if applicable, and zero momentum transfer. Here, the indices 1, 2, and 3 for spatial components and the index 4 is used for the temporal component. Latin indices range from 1 to 3 and Greek indices from 1 to 4. Primed energy and momentum components are the values at the hadron sink.

A.1 Nucleon Rest Frame

$$\mathcal{Z}_V \Pi_V^{00}(\Gamma_0) = -\frac{3m_N}{4} \langle x \rangle_{u-d}, \quad (\text{A.1})$$

$$\mathcal{Z}_V \Pi_A^{kk}(\Gamma_0) = -\frac{m_N}{4} \langle x \rangle_{u-d}, \quad (\text{A.2})$$

$$\mathcal{Z}_A \Pi_V^{j0}(\Gamma_k) = -\frac{im_N}{2} \delta_{jk} \langle x \rangle_{\Delta u - \Delta d}, \quad (\text{A.3})$$

$$\mathcal{Z}_T \Pi_V^{\mu\nu\rho}(\Gamma_k) = i\epsilon_{\mu\nu\rho k} \frac{m_N}{8} (2\delta_{0\rho} - \delta_{0\mu} - \delta_{0\nu}) \langle x \rangle_{\delta u - \delta d}. \quad (\text{A.4})$$

A.2 Meson General frame

A.2.1 General momentum transfer

$$\mathcal{Z}_v \Pi^4 = C(E' + E)A_{10}, \quad (\text{A.5})$$

$$\mathcal{Z}_v \Pi^j = -iC(p'_j + p_j)A_{10}, \quad (\text{A.6})$$

$$\mathcal{Z}_{vD} \Pi^{\{44\}} = C \left(\frac{A_{20}}{4} (m^2 - 2(E' + E)^2 - p'_\rho p_\rho) + B_{20} (m^2 - 2(E' - E)^2 + p'_\rho p_\rho) \right), \quad (\text{A.7})$$

$$\mathcal{Z}_{vD} \Pi^{\{j4\}} = C \left(\frac{iA_{20}}{2} (E' + E)(p'_j + p_j) + 2iB_{20}(E' - E)(p'_j - p_j) \right), \quad (\text{A.8})$$

$$\mathcal{Z}_{vD} \Pi^{\{jk\}} = C \left(\frac{A_{20}}{2} (p'_j + p_j)(p'_k + p_k) + 2B_{20}(p'_j - p_j)(p'_k - p_k) \right), \quad (\text{A.9})$$

$$\begin{aligned} \mathcal{Z}_{vDD} \Pi^{\{ij4\}} = C & \left(\frac{A_{30}}{4} (E' + E)(p'_i + p_i)(p'_j + p_j) \right. \\ & + B_{20} \left((E' + E) \left(\frac{1}{3} p'_j p_i + \frac{1}{3} p'_i p_j \right) \right. \\ & \left. \left. + \left(\frac{E}{3} - E' \right) p'_i p'_j + \left(\frac{E'}{3} - E \right) p_i p_j \right) \right), \end{aligned} \quad (\text{A.10})$$

$$\begin{aligned} \mathcal{Z}_{vDDD} \Pi^{\{ijk4\}} = C & \left(-\frac{i}{8} A_{40} (E' + E)(p'_i + p'_i)(p'_j + p_j)(p'_k + p_k) \right. \\ & + iB_{40} \left(E' \left(\frac{1}{6} p'_k p_i p_j + \frac{1}{6} p'_j p_i p_k + \frac{1}{6} p'_i p_j p_k - \frac{1}{2} p'_i p'_j p'_k \right) \right. \\ & \left. + E \left(\frac{1}{6} p'_j p'_k p_i + \frac{1}{6} p'_i p'_k p_j + \frac{1}{6} p'_i p'_j p_k - \frac{1}{2} p_i p_j p_k \right) \right) \\ & + iC_{40} \left(\left(\frac{E}{2} - E' \right) p'_i p'_j p'_k + \left(E - \frac{E'}{2} \right) p_i p_j p_k \right. \\ & + E \left(\frac{1}{2} p'_j p'_k p_i + \frac{1}{2} p'_i p'_k p_j + \frac{1}{2} p'_i p'_j p_k \right) \\ & \left. \left. + E \left(-\frac{1}{2} p'_k p_i p_j - \frac{1}{2} p'_i p_k p_j - \frac{1}{2} p'_j p_i p_k \right) \right) \right). \end{aligned} \quad (\text{A.11})$$

A.2.2 Zero momentum transfer

$$\mathcal{Z}_{\text{vD}} \Pi^{\{44\}} = \frac{1}{2E'} \left(\frac{m^2}{2} - 2E'^2 \right) \langle x \rangle, \quad (\text{A.12})$$

$$\mathcal{Z}_{\text{vDD}} \Pi^{\{ij4\}} = -p'_i p'_j \langle x^2 \rangle, \quad (\text{A.13})$$

$$\mathcal{Z}_{\text{vDDD}} \Pi^{\{ijk4\}} = -ip'_i p'_j p'_k \langle x^3 \rangle. \quad (\text{A.14})$$

A.3 Meson Rest frame

A.3.1 General momentum transfer

$$\mathcal{Z}_{\text{v}} \Pi^0 = C(E + m)A_{10}, \quad (\text{A.15})$$

$$\mathcal{Z}_{\text{vD}} \Pi^{\{44\}} = C \left(-\frac{A_{20}}{4}(E + m)(2E + m) - A_{22}(E - m)(2E - m) \right), \quad (\text{A.16})$$

$$\mathcal{Z}_{\text{vD}} \Pi^{\{j4\}} = C \left(\frac{iA_{20}}{2}(E + m)p_j - 2iA_{22}(E - m)p_j \right), \quad (\text{A.17})$$

$$\mathcal{Z}_{\text{vD}} \Pi^{\{jk\}} = C \left(\frac{A_{20}}{2}p_j p_k - 2A_{22}p_j p_k \right). \quad (\text{A.18})$$

A.3.2 Zero momentum transfer

$$\mathcal{Z}_{\text{vD}} \Pi^{\{44\}} = -\frac{3m}{4} \langle x \rangle, \quad (\text{A.19})$$