ABSTRACT

Title of Dissertation: TOWARDS PRECISION-CONTROLLED PARTONIC STRUCTURES FROM FIRST PRINCIPLES

Jinchen He
Doctor of Philosophy, 2026

Dissertation Directed by: Professor Xiangdong Ji
Department of Physics

The internal structure of hadrons is governed by nonperturbative Quantum Chromodynamics (QCD) and is being probed with increasing precision by modern high-energy experiments. This dissertation presents first-principles calculations of partonic observables using lattice QCD together with effective field theory methods, with controlled systematic uncertainties, advancing from collinear structure to transverse-momentum-dependent distributions (TMDs) that encode the three-dimensional partonic structure of hadrons.

Within the large momentum effective theory (LaMET) framework, this work presents state-of-the-art calculations of pion distribution amplitudes and systematic studies of nucleon parton distributions, with control of renormalization, excited-state contamination, Fourier-transform systematics, and power corrections. A Coulomb-gauge formulation of quasi-distributions is implemented to simplify ultraviolet structure by avoiding Wilson-line related linear divergences. Systematic effects associated with Gribov copies are quantified and found to be negligible at current statistical precision, supporting

the effectiveness of this approach.

Building on these developments, this dissertation reports first-principles lattice determinations of key TMD quantities, including nucleon TMD parton distributions, the Collins-Soper kernel, the intrinsic soft function, and pion TMD observables. These results provide nonperturbative inputs directly relevant to global QCD analyses and to the precision hadron-structure program at current and future facilities, including the Electron-Ion Collider.

In parallel, this dissertation explores machine-learning acceleration of lattice gauge simulations through neural field transformations embedded in Hybrid Monte Carlo. In two-dimensional U(1) tests, the method significantly reduces autocorrelation and improves performance toward finer lattice spacing. Such gains support precision-controlled partonic calculations by enabling finer lattices and larger hadron boosts, thereby reducing LaMET power corrections and improving control of discretization effects. The dissertation also outlines emerging directions using AI agents and foundation models to advance lattice data-analysis workflows under physics-based constraints.

TOWARDS PRECISION-CONTROLLED PARTONIC STRUCTURES FROM FIRST PRINCIPLES

by

Jinchen He

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2026

Advisory Committee:
Professor Xiangdong Ji, Chair/Advisor
Professor Thomas Cohen
Professor Anson Hook
Professor Zackaria Chacko
Professor Harry Dankowicz

Dedication

For an intelligible universe.

Acknowledgments

Over the past four years of my Ph.D. journey, I have been fortunate to meet many friends, classmates, mentors, and elders whose support, encouragement, and companionship have enabled me to complete this long and demanding process. What appears on these pages is not the result of individual effort alone. It is the outcome of sustained guidance, collaboration, and trust from a community that has shaped both my academic growth and my personal development.

I would first like to express my deepest gratitude to my advisor, Professor Xiangdong Ji. I have always admired your profound knowledge and deep physical insight. Your guidance has been invaluable to my training as a theoretical physicist. Many of your insightful ideas will continue to influence and guide my future research. The two years I spent at the University of Maryland laid the foundation for my research training. Jinghong Yang, as both a classmate and roommate, was the person with whom I exchanged ideas most frequently. Your sharp thinking and careful consideration often left me humbled, and I am grateful to have met such a thoughtful friend. I hope we will have many more opportunities to collaborate in the future. I also thank Jack Holligan, Chen Yang, and Haocheng Zhang; discussions with you have always been inspiring.

I am sincerely grateful to Dr. Yong Zhao for your meticulous and patient academic guidance and encouragement. Your precision and rigor in physics set an example that I continue to learn from. The two years at Argonne National Laboratory were years of rapid growth for me. I am thankful for the research environment and for the Argonne theory group: Ian, Ale, Anna, Harry, and Bob. I also thank Patrick, Chiara, Bryce, Noah, and Justin. In particular, I would like to thank Joey, Daniel, Josh, Jane, Arthur, Eric, and Bill. I am very glad to have met all of you; you made my time at Argonne both productive and memorable.

I would like to give special thanks to several senior colleagues who provided generous academic support. Yushan Su offered me substantial help and encouragement; your dedication to physics has always impressed me. I thank Rui Zhang, Yuxun Guo, and Xiang Gao for your guidance and support through collaboration and discussion. I also appreciate the pleasant collaborations with Fei Yao and Wei-Yang Liu. I am particularly grateful to Xiangyu Jiang for extensive assistance in coding and software-related matters.

I also acknowledge my collaborators at Brookhaven National Laboratory, Dr. Swagato Mukherjee, Dr. Peter Petreczky, and Dr. Dennis Bollweg, as well as James Osborn and Xiao-Yong Jin at Argonne. I would especially like to thank Andre and Jason at Berkeley Lab. You were my introduction to lattice QCD, and what I learned at CalLat continues to benefit my research to this day.

I am deeply grateful to Professor Wei Wang. It was under your guidance that I first became engaged with partonic physics and LaMET. I greatly respect your pragmatic and rigorous approach to research and teaching. I also thank the teachers and colleagues of the Lattice Parton Collaboration. In particular, I appreciate the guidance of Yi-Bo Yang, Jianhui Zhang, and Xiaonu Xiong. I am especially thankful to Qi-An Zhang, Jun Hua, Jun Zeng, Minhuan Chu, Zhifu Deng, Chang Yang, Fei Huang, and Hang Liu. I will always remember the time we spent working and discussing together, and I sincerely wish you all continued success.

I am indebted to my parents and family in China. Your unwavering support has been a necessary condition for me to complete this Ph.D. program. Studying abroad for many years and being far from home has often left me feeling that I have not accompanied you enough. I wish you health and longevity.

I thank my friends from afar: Boyue Li, Bocheng Zhang, Shenghang Xu, Le Zhang, and Weiqi Shi. Your conversations and companionship have been an essential part of my Ph.D. life.

I am especially grateful to my girlfriend, Yuwei Li. Thank you for accompanying me through the entirety of this Ph.D. program. A relationship sustained across long distances is rare and precious, and I value it deeply.

There are many more people who have contributed to my journey than can be fully listed here. Among the forces that have brought me to this point, some are visible, like the support explicitly acknowledged above. Others are less visible: the expectations of those who know me, the trajectory I have followed within academia, and the curiosity and desire for understanding that drew me to theoretical physics in the first place. I have always regarded the opportunity to pursue theoretical physics as a luxury. I am grateful for the chance to find, within an otherwise ordinary and at times monotonous life, something that reaches beyond it. I am grateful for the opportunity to seek, within a finite lifetime, laws that aspire to permanence. Above all, I am grateful for the opportunity to believe in an intelligible universe.

Table of Contents

Dedication ii

Acknowledgements iii

Table of Contents vi

List of Tables viii

List of Figures ix

List of Abbreviations xi

Chapter 1: Introduction 1
- 1.1 List of Research Contributions 3
- 1.2 Outline 5

Chapter 2: Partonic Physics in Hadrons 6
- 2.1 From Scaling to QCD: The Parton Framework 6
- 2.2 Light-front Quantization 8
- 2.3 Collinear Distributions 12
- 2.4 Momentum Tomography: TMDs 19
- 2.5 Spatial Tomography: GPDs 25
- 2.6 Conclusion 26

Chapter 3: First-Principles Calculation of Strong Interaction: Lattice QCD 28
- 3.1 Lattice QCD as a First-Principles Framework 30
- 3.2 Fermion Formulations on the Lattice 32
- 3.3 Monte Carlo Algorithms for Lattice Gauge Ensembles 35
- 3.4 Benchmark Calculations: Axial Charge 41
  - 3.4.1 Nucleon Axial Charge 41

Chapter 4: Systematics Controlled Calculations of Partonic Physics from Lattice QCD 48
- 4.1 Large Momentum Effective Theory 48
  - 4.1.1 Recent Developments in LaMET 49
  - 4.1.2 Quasi Distributions in Coulomb Gauge 50
  - 4.1.3 General Analysis Procedure 54
- 4.2 Calculation of Pion DAs 60
- 4.3 Calculation of Nucleon PDFs 67
  - 4.3.1 Theoretical Framework 70
  - 4.3.2 Numerical Results 75
  - 4.3.3 Conclusion 84
- 4.4 Calculation of Pion TMDs 86

4.4.1 Theoretical Framework . . . 88
4.4.2 Numerical Results . . . 94
4.4.3 Conclusion . . . 107
4.5 Calculation of Nucleon TMDs . . . 108
4.5.1 Theoretical Framework . . . 110
4.5.2 Numerical Results . . . 112
4.5.3 Conclusion . . . 125

Chapter 5: Applications of Machine Learning in Lattice Gauge Theory 127
5.1 Applications in Gauge Configuration Generation . . . 129
5.1.1 Overview . . . 130
5.1.2 Neural Field Transformations for Hybrid Monte Carlo . . . 131
5.2 Applications in Data Analysis . . . 138
5.2.1 Operator Optimization . . . 138
5.2.2 Parameter Inference from Correlators and Ensembles . . . 139
5.3 AI Agent and Foundation Model . . . 140

Chapter 6: Summary and Outlook 143

Bibliography 147

List of Tables

4.1 Lattice setup for the lattice calculation of pion DAs 61
4.2 Collection of ground state fit settings for nucleon PDFs 76

List of Figures

1.1 Reductionism . . . 1

2.1 Deep-inelastic scattering . . . 6
2.2 The CT18 NNLO PDFs . . . 16
2.3 The NNPDF4.0 NNLO PDFs . . . 16
2.4 The Drell–Yan process . . . 20
2.5 Graphs of the Wilson line structure . . . 22

3.1 Lattice results and FLAG averages for the isovector axial charge g_A^{u-d} . . . 42
3.2 Decomposition of excited state contributions to the axial charge extraction . . . 45
3.3 Psychedelic Moose and FH ratio plot for the axial charge extraction . . . 47

4.1 Pion valence quark PDF . . . 49
4.2 Measurement distortion from Gribov copies . . . 53
4.3 Self-renormalization of quasi-DA . . . 64
4.4 DAs of pion . . . 65
4.5 Renormalized quasi-PDF matrix elements . . . 78
4.6 Valence and full contribution of unpolarized nucleon PDF . . . 80
4.7 Unpolarized nucleon PDF . . . 81
4.8 Valence and full contribution of helicity nucleon PDF . . . 82
4.9 Helicity nucleon PDF . . . 82
4.10 Valence and full contribution of transversity nucleon PDF . . . 83
4.11 Transversity nucleon PDF . . . 84
4.12 Renormalized quasi-TMD matrix elements . . . 98
4.13 CS kernel . . . 99
4.14 Extrapolations of reduced quasi-TMDWF . . . 100
4.15 Intrinsic soft function . . . 101
4.16 Pion TMDWF . . . 103
4.17 Pion TMDPDF in x space . . . 104
4.18 Pion TMDPDF in $b_\perp$ space . . . 104
4.19 Extrapolated pion TMDPDF in $b_\perp$ space . . . 106
4.20 Pion TMDPDF in $k_\perp$ space . . . 106
4.21 Quasi TMDPDFs in GI formalism . . . 110
4.22 Wilson loop and renormalized zero-momentum matrix element . . . 114
4.23 L-dependence of subtracted quasi TMDPDF matrix elements . . . 117
4.24 Renormalized quasi TMDPDFs of nucleon . . . 118
4.25 Quasi TMDPDF v.s. TMDPDF of nucleon . . . 119
4.26 Physical limit of nucleon TMDPDF . . . 120
4.27 Nucleon TMDPDF with different physical limit . . . 121
4.28 Difference of nucleon TMDPDF from different Lorentz structures . . . 122
4.29 Systematic uncertainties of nucleon TMDPDF . . . 123
4.30 Nucleon TMDPDF in x space . . . 123

4.31 Nucleon TMDPDF in $b_\perp$ space 124

5.1 Cropped section of a lattice configuration illustrating the field transformation 133
5.2 Comparison of regular HMC and NTHMC with the `Base` model 136
5.3 Comparison of NTHMC with various network architectures 137

List of Abbreviations

AI	Artificial Intelligence
AION-1	Astronomical Omni-modal Network
AMA	All-Mode Averaging
CG	Coulomb Gauge
CNN	Convolutional Neural Network
CSS	Collins–Soper–Sterman
DA	Distribution Amplitude
DCSB	Dynamical Chiral Symmetry Breaking
DGLAP	Dokshitzer–Gribov–Lipatov–Altarelli–Parisi
DIS	Deep-Inelastic Scattering
DSE	Dyson–Schwinger Equations
DVCS	Deeply Virtual Compton Scattering
DVMP	Deeply Virtual Meson Production
DY	Drell–Yan
EIC	Electron–Ion Collider
ERBL	Efremov–Radyushkin–Brodsky–Lepage
FH	Feynman–Hellmann
GEVP	Generalized Eigenvalue Problem
GNoME	Graph Networks for Materials Exploration
GPD	Generalized Parton Distributions
GI	Gauge-Invariant
HERA	Hadron-Elektron-Ringanlage
HISQ	Highly Improved Staggered Quark
HMC	Hybrid Monte Carlo
HYP	Hypercubic
LaMET	Large Momentum Effective Theory
LHC	Large Hadron Collider
LGT	Lattice Gauge Theory
LL	Leading Logarithmic
LO	Leading Order
LPC	Lattice Parton Collaboration
LRR	Leading Renormalon Resummation
MCMC	Markov Chain Monte Carlo
MD	Molecular Dynamics
ML	Machine Learning
NLL	Next-to-Leading Logarithmic
NLO	Next-to-Leading Order
NNLO	Next-to-Next-to-Leading Order
NTHMC	Neural Network Transformed Hybrid Monte Carlo

OPE	Operator Product Expansion
PDF	Parton Distribution Function
QCD	Quantum Chromodynamics
RG	Renormalization Group
RGR	Renormalization-Group Resummation
RI/MOM	Regularization-Independent Momentum Subtraction
SCET	Soft-Collinear Effective Theory
SIDIS	Semi-Inclusive Deep-Inelastic Scattering
SLAC	Stanford Linear Accelerator Center
SNR	Signal-to-Noise Ratio
TMD	Transverse-Momentum-Dependent
TMDs	Transverse-Momentum-Dependent Distributions
TMDPDF	Transverse-Momentum-Dependent Parton Distribution Function
TMDWF	Transverse-Momentum-Dependent Wave Function
TR	Threshold Resummation
UV	Ultraviolet

Chapter 1: Introduction

From the earliest notions of atoms to the modern picture of quarks and gluons, the history of physics has been guided by reductionism: the idea that complex phenomena can be understood by identifying more elementary constituents and the interactions among them. The path from atomic structure to nuclei, from nuclei to nucleons, and from nucleons to quarks has repeatedly reshaped our view of matter, each successive layer unveiling a new effective scale where smaller degrees of freedom organize the behavior of the whole. In the visible universe, the smallest known structural description of ordinary matter is at the parton level, where quarks and gluons form the internal degrees of freedom of hadrons.

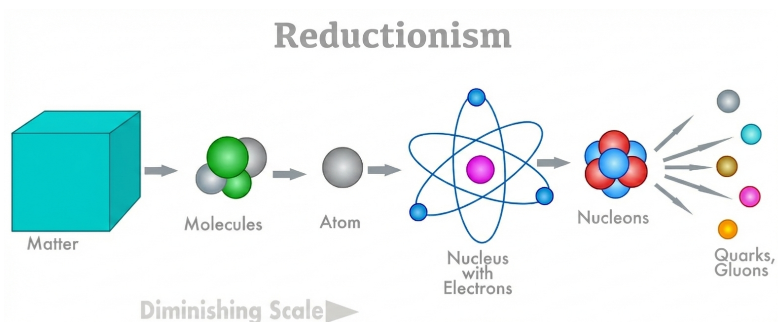

Figure 1.1: The concept of reductionism in physics, showing how complex systems can be understood through their simpler constituents.

At the same time, reductionism has never been the sole organizing principle. The celebrated essay "More is Different" [1] argued that new organizing principles can emerge at higher levels of complexity, challenging the idea that a full microscopic description automatically yields macroscopic understanding. This critique remains influential, especially in strongly coupled systems where emergent behavior is not easily reduced to a few simple constituents.

In the present era, however, the balance between these perspectives is being reshaped by the rapid growth of computing power and algorithms. High-performance computing and simulation now make it possible to solve problems that were previously out of reach, enabling direct numerical access to nonperturbative regimes of quantum field theory. This is particularly true in quantum chromodynamics (QCD), where analytic control is limited and where the strong coupling at low energies renders perturbation theory ineffective. In this sense, computing is not merely a tool but an enabling paradigm: it reopens the possibility that "more" can be approached directly, with complex many-body behavior computed from first principles.

These developments sit within the broader framework of the Standard Model, which organizes known fundamental particles and interactions into a unified field-theoretic description. The Standard Model is a renormalizable quantum field theory based on the gauge group $SU(3)_c \times SU(2)_L \times U(1)_Y$, describing the strong and electroweak interactions, while gravity is not incorporated within this framework. In a broader physics context one often speaks of four fundamental interactions: electromagnetic, weak, strong, and gravitational. The non-Abelian gauge structure of QCD explains asymptotic freedom, which underlies the partonic picture and the emergence of scaling violations in high-energy scattering [2, 3]. The same framework provides the theoretical foundation for factorization in hard processes and for the operator-based definitions of parton distributions [4].

Experimentally, our knowledge of structure at ever smaller distance scales has been enabled by high-energy facilities and precision detectors. Deep-inelastic scattering at SLAC revealed Bjorken scaling and the parton model [5, 6]. Later, colliders such as HERA, the Tevatron, and the Large Hadron Collider (LHC) pushed these probes to higher energies, while dedicated nuclear facilities such as Jefferson Lab mapped the valence region with high precision. The planned Electron–Ion Collider (EIC) will further extend this program by enabling three-dimensional imaging of hadrons and precision studies of parton dynamics in nucleons and nuclei [7, 8].

Cosmological observations further sharpen the motivation. While dark matter and dark energy dominate the cosmic energy budget, their microscopic nature remains unknown. Within the framework of the Standard Model, the dynamics of the visible sector are predominantly governed by QCD and the electroweak theory. Within this visible sector, the partonic structure of hadrons is the most elementary known organization of matter, and a complete description of partons is essential for connecting fundamental theory with experimental measurements in nuclear and particle physics.

Despite decades of progress, the strong interaction still presents fundamental challenges [9]. Confinement, the emergence of hadrons as relativistic bound states, and the dynamical generation of mass and structure remain only partially understood. Describing a relativistic bound state from first principles is one of the deepest problems in quantum field theory. The internal partonic structure of hadrons encodes these nonperturbative dynamics, yet it is precisely this structure that is most difficult to compute analytically.

This dissertation is motivated by these intertwined themes. It aims to advance first-principles calculations of partonic structure using lattice QCD and modern effective field theory frameworks. The core objective is to connect QCD at the quark–gluon level to measurable parton observables, with controlled systematics and clear theoretical interpretation.

1.1 List of Research Contributions

The project on lattice calculations of the nucleon axial charge in Chapter 3 is based on the following research contribution:

[1] J. He, D. A. Brantley, C. C. Chang, I. Chernyshev, E. Berkowitz, D. Howarth, C. Körber, A. S. Meyer, H. Monge-Camacho and E. Rinaldi, *et al.* “Detailed analysis of excited-state systematics in a lattice QCD calculation of gA,” Phys. Rev. C **105**, no.6, 065203 (2022) [arXiv:2104.05226

[hep-lat]].

The project on systematic uncertainties from Gribov copies in lattice calculation of parton distributions in the Coulomb Gauge in Chapter 4 is based on the following research contribution:

[2] X. Gao, J. He, R. Zhang and Y. Zhao, "Systematic Uncertainties from Gribov Copies in Lattice Calculation of Parton Distributions in the Coulomb Gauge," Chin. Phys. Lett. **41**, no.12, 121201 (2024) [arXiv:2408.05910 [hep-lat]].

The projects on lattice calculations of partonic physics from lattice QCD described in Chapter 4 are based on the following research contributions:

[3] J. Hua *et al.* [Lattice Parton], "Pion and Kaon Distribution Amplitudes from Lattice QCD," Phys. Rev. Lett. **129**, no.13, 132001 (2022) [arXiv:2201.09173 [hep-lat]].

[4] J. C. He *et al.* [Lattice Parton Collaboration (LPC)], "Unpolarized transverse momentum dependent parton distributions of the nucleon from lattice QCD," Phys. Rev. D **109**, no.11, 114513 (2024) [arXiv:2211.02340 [hep-lat]].

[5] M. H. Chu *et al.* [Lattice Parton], "Transverse-momentum-dependent wave functions of the pion from lattice QCD," Phys. Rev. D **109**, no.9, L091503 (2024) [arXiv:2302.09961 [hep-lat]].

[6] M. H. Chu *et al.* [Lattice Parton (LPC)], "Lattice calculation of the intrinsic soft function and the Collins-Soper kernel," JHEP **08**, 172 (2023) [arXiv:2306.06488 [hep-lat]].

[7] D. Bollweg, X. Gao, J. He, S. Mukherjee and Y. Zhao, "Transverse-momentum-dependent pion structures from lattice QCD: Collins-Soper kernel, soft factor, TMDWF, and TMDPDF," Phys. Rev. D **112**, no.3, 034501 (2025) [arXiv:2504.04625 [hep-lat]].

[8] X. Gao, J. He, J. Lin, S. Mukherjee, P. Petreczky, R. Zhang and Y. Zhao, "Nucleon Parton Distribution Functions from Boosted Correlations in the Coulomb gauge," [arXiv:2602.11283 [hep-lat]].

The projects on applications of machine learning in lattice gauge theory described in Chapter 5 are based on the following research contributions:

[9] M. Zhu, M. Tian, X. Yang, T. Zhou, L. Yuan, P. Zhu, E. Chertkov, S. Liu, Y. Du and Z. Ji, *et al.* "Probing the Critical Point (CritPt) of AI Reasoning: a Frontier Physics Research Benchmark," [arXiv:2509.26574 [cs.AI]].

[10] J. He, X. Y. Jin, J. C. Osborn and Y. Zhao, "Neural Field Transformations for Hybrid Monte Carlo: Architectural Design and Scaling," [arXiv:2511.02018 [hep-lat]].

1.2 Outline

The remainder of this dissertation is organized as follows. Chapter 2 reviews the phenomenology of partonic physics in hadrons, with a focus on collinear distributions and transverse-momentum-dependent distributions (TMDs). Chapter 3 introduces the lattice QCD framework and presents a benchmark calculation of the nucleon axial charge. Chapter 4 describes systematic calculations of partonic observables from lattice QCD, including pion distribution amplitudes (DAs), nucleon parton distribution functions (PDFs), and TMDs for both pion and nucleon. Chapter 5 explores applications of machine learning in lattice gauge theory, including applications in configuration generation, data analysis, AI agent and foundation model. Finally, Chapter 6 summarizes the thesis and discusses future directions.

Chapter 2: Partonic Physics in Hadrons

Understanding the internal structure of hadrons in terms of quarks and gluons is a central goal of QCD. Over the past five decades a coherent theoretical framework has been developed to describe how partons behave inside hadrons and how they can be probed experimentally. This chapter provides a brief overview of partonic physics, beginning with the historical foundations of the parton model and QCD factorization. We then discuss collinear parton structures, including PDFs and DAs, followed by TMD structures and their evolution. We then briefly mention the generalized parton distributions (GPDs) and spatial tomography, emphasizing the experimental and phenomenological program that extracts these objects from data and the role of future facilities in completing a three-dimensional picture of hadrons [7, 8, 10].

2.1 From Scaling to QCD: The Parton Framework

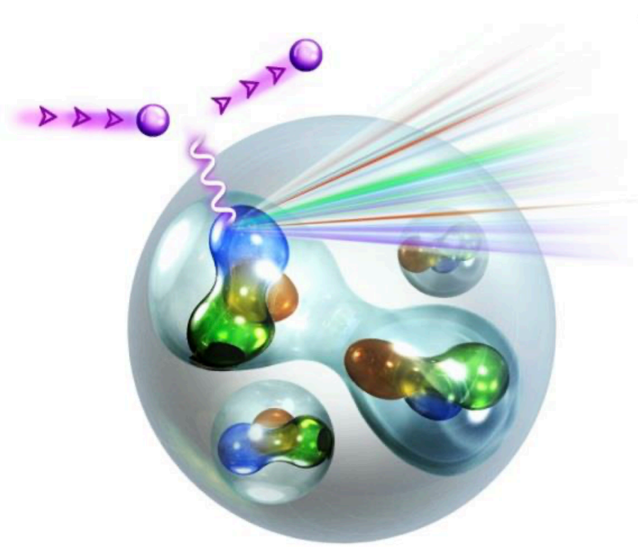

Figure 2.1: Diagramatic illustration of the deep-inelastic scattering process. Credit to Ref. [11]

In the late 1960s, the electron-nucleon deep-inelastic scattering (DIS) experiments at SLAC revealed approximate Bjorken scaling, indicating scattering from point-like constituents inside the proton, as illustrated schematically in Fig. 2.1. The parton model proposed by Feynman provided a simple interpretation: a fast-moving hadron contains quasi-free partons carrying fractions of its momentum [6]. Bjorken and Paschos showed that such a picture leads to scale-invariant structure functions and sum rules, relating the DIS structure function $F_2(x)$ to momentum fraction distributions of charged partons [5, 12].

Embedding the parton model into a quantum field theory required asymptotic freedom. The discovery of asymptotic freedom in non-Abelian gauge theories established QCD as the theory of strong interactions and explained the observed scaling violations [2, 3]. Asymptotic freedom implies that the strong coupling becomes weak at short distances, with the running coupling (at one-loop accuracy) governed by the renormalization group equation,

$$\alpha_s^{(1)}(Q^2) = \frac{1}{\beta_0 \ln(Q^2/\Lambda_{\rm QCD}^2)} \ , \tag{2.1}$$

so that radiative corrections at high momentum transfer generate logarithmic contributions of the form $\alpha_s \ln(Q^2/\mu^2)$, such terms appear as powers $(\alpha_s \ln Q^2)^n$ and must be resummed. These logarithmic contributions arise from collinear parton splitting processes such as $q \to qg$, $g \to gg$, and $g \to q\bar{q}$, which redistribute longitudinal momentum among partons and induce a scale dependence of parton distribution functions.

The QCD evolution of PDFs is described by the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) equations, derived in the 1970s, which resum these leading logarithmic contributions and account for the logarithmic scaling violations observed in deep-inelastic structure functions [13–15].

Another crucial development was Wilson's operator product expansion (OPE) and the concept of factorization [16]. The OPE provides a separation of short-distance effects from long-distance physics in the product of currents, enabling the formal proof that inclusive DIS cross-sections factorize into perturbative coefficients and universal PDFs [4]. This foundation underlies the universality of PDFs across processes such as DIS and Drell–Yan [17], up to higher-twist corrections. In exclusive processes, the analogous collinear object are the DAs, formalized independently by Brodsky–Lepage [18, 19] and by Efremov–Radyushkin [20, 21], leading to the ERBL evolution equations. Together with the DGLAP description of scaling violations in DIS, this framework laid the foundation for modern QCD phenomenology, where global analyses combine a wide range of experimental data to extract universal parton distributions through parameterized fits.

2.2 Light-front Quantization

The experimental discovery of approximate Bjorken scaling and its interpretation within the parton model naturally emphasize the Bjorken limit, defined by $Q^2 \to \infty$ with the Bjorken variable $x_B = \frac{Q^2}{2P\cdot q}$ held fixed. In this regime the virtual photon probes distance scales of order $1/Q$, and the hadronic tensor can be written as the Fourier transform of a current–current correlator [4],

$$W_{\mu\nu}(P,q) = \frac{1}{4\pi}\int d^4\xi\, e^{iq\cdot\xi}\langle P|[J_\mu(\xi), J_\nu(0)]|P\rangle \,, \tag{2.2}$$

where $J_\mu(x) = \sum_f e_f\, \bar{\psi}_f(x)\gamma_\mu\psi_f(x)$ is the electromagnetic current. The state $|P\rangle$ denotes a single-hadron momentum eigenstate with four-momentum P^μ, normalized relativistically as

$$\langle P'|P\rangle = 2P^0(2\pi)^3\delta^{(3)}(\vec{P} - \vec{P}') \,. \tag{2.3}$$

For large Q, the rapidly oscillating phase factor $e^{iq\cdot\xi}$ suppresses contributions from large space–time separations, so that the integral is dominated by short distances $|\xi| \sim 1/Q$. More specifically, the current product becomes most singular as the invariant separation approaches the light cone, $\xi^2 \to 0$, implying that the leading contributions arise from space–time intervals near light-like directions.

The dominance of light-like separations admits a natural interpretation within the parton picture. In the Bjorken limit, the hard probe couples primarily to partons carrying a finite fraction of the hadron momentum and moving approximately collinearly with the parent hadron. These partons are nearly on shell compared with the hard scale Q^2, and their propagation between the two currents is therefore effectively light-like. Consequently, the dominant contributions to the current–current correlator originate from correlations of quark and gluon fields separated along the light cone. This provides a direct bridge between the phenomenology of DIS and light-cone correlation functions of quark and gluon fields. Indeed, PDFs and DAs are most naturally defined as hadronic matrix elements of nonlocal operators separated along a light-like direction.

From a Hamiltonian perspective, these considerations motivate quantization directly on a light-like hypersurface, following Dirac's front-form dynamics [22–24]. In light-front quantization one introduces light-cone coordinates

$$x^{\pm} = \frac{1}{\sqrt{2}}(x^0 \pm x^3) \; , \qquad \vec{x}_\perp = (x^1, x^2) \; , \tag{2.4}$$

and chooses x^+ as the evolution parameter, so that fields are quantized on the light-front hypersurface $x^+ = 0$. The generator of translations in x^+,

$$P^- = \frac{1}{\sqrt{2}}(P^0 - P^3), \tag{2.5}$$

plays the role of the Hamiltonian.

Dirac classified the generators of the Poincaré algebra into *kinematical* and *dynamical* ones according to whether they preserve the quantization hypersurface. Kinematical generators leave the surface $x^+ = 0$ invariant and therefore do not involve the Hamiltonian in their realization, while dynamical generators transform this hypersurface into another one and consequently depend on the dynamics of the theory. In the front form the kinematical generators include the longitudinal momentum P^+, the transverse momenta $P^\perp$, the transverse rotations J^{12}, the longitudinal boost J^{+-}, and the transverse boosts J^{+i} $(i = 1, 2)$. The remaining generators, $P^-,\ J^{-i}$ $(i = 1, 2)$, are dynamical.

This distribution differs from the conventional equal-time quantization, where fields are quantized on the hypersurface $x^0 = 0$. In that case the Hamiltonian P^0 and the Lorentz boosts J^{0i} are dynamical, while the spatial translations P^i and spatial rotations J^{ij} are kinematical. Consequently, the front form reduces the number of dynamical generators from four to three. This enlargement of the kinematical subgroup greatly simplifies the description of relativistic bound states and boosts along the longitudinal direction.

Another distinctive feature of this framework is that the longitudinal momentum $p^+ = (p^0 + p^3)/\sqrt{2}$ is positive for physical particles, which severely restricts vacuum fluctuations. As a consequence, the light-front vacuum in many theories is considerably simpler than the equal-time vacuum [22], and particle number changing processes are encoded directly in the structure of the light-front Hamiltonian rather than in vacuum condensates.

Within light-front quantization, a hadronic state admits a Fock space expansion at fixed light-front time [23],

$$|P\rangle = \sum_n \int [\mathrm{d}x_i\, \mathrm{d}^2\vec{k}_{\perp i}]\, \psi_n(x_i, \vec{k}_{\perp i}, \lambda_i)\, |n; x_i P^+, \vec{k}_{\perp i}, \lambda_i\rangle \ , \tag{2.6}$$

where n is the number of partons in the Fock component, $\vec{k}_{\perp i}$ are the transverse momenta of the partons, and λ_i are the light-front helicities, $x_i = p_i^+/P^+$ are longitudinal momentum fractions satisfying $\sum_i x_i = 1$, and ψ_n are light-front wave functions. These wave functions provide a complete set of amplitudes characterizing hadronic structure in the light-front framework. Because longitudinal boosts act multiplicatively on all p_i^+, the momentum fractions x_i remain invariant, and the light-front wave functions furnish a frame-independent generalization of the nonrelativistic wave-function description of bound states [22]. The intuitive quantum-mechanical picture of a bound state as a superposition of multi-particle configurations thus carries over to relativistic quantum field theory, with the light-front wave functions encoding the probability amplitudes for finding partonic constituents with specified longitudinal momentum fractions and transverse momenta.

This perspective gives a transparent amplitude interpretation of both PDFs and DAs. PDFs can be expressed as diagonal matrix elements of nonlocal light-cone operators and are related to squared light-front wave functions summed over spectator configurations. At leading twist, they correspond to probability densities in longitudinal momentum fraction, modulo the necessary gauge links ensuring color gauge invariance. DAs arise in exclusive processes as projections of light-front wave functions onto states with vanishing transverse separation at leading power. In particular, the leading-twist meson DA is obtained from the valence Fock component by integrating over transverse momenta up to the factorization scale, thereby encoding the longitudinal momentum distribution of the minimal quark–antiquark configuration. The evolution equations governing PDFs and DAs can then be understood as renormalization-group equations for these light-cone operators, reflecting the scale dependence of the underlying light-front amplitudes.

Despite these appealing features, light-front quantization does not eliminate the dynamical complexity of relativistic bound states. The simplicity of the vacuum is accompanied by subtleties associated with constrained degrees of freedom and zero modes, corresponding to field configurations with vanishing longitudinal momentum $p^+ = 0$ [25]. Such modes can contribute to matrix elements and are closely related to issues of spontaneous symmetry breaking, gauge invariance, and the realization of vacuum structure. In modern language, light-front quantization may be viewed as an effective description in which small-k^+ modes are integrated out, with their effects absorbed into renormalization factors [25]. In practice, zero modes and rapidity divergences complicate both analytical treatments and numerical approaches based on truncated Fock spaces. Moreover, solving the light-front Hamiltonian eigenvalue problem remains a nonperturbative challenge equivalent in difficulty to solving QCD itself. In this sense, light-front quantization reformulates rather than resolves the fundamental problem of describing relativistic bound states.

Nevertheless, the light-front framework provides a unifying language that connects the parton model intuition emerging from the Bjorken limit with the operator definitions of PDFs and DAs in QCD. It clarifies why light-cone correlations dominate high-energy processes, furnishes a boost-invariant description of hadronic structure, and supplies an amplitude-level interpretation of the universal distributions that enter factorization theorems. In the following sections, we make these connections explicit by introducing the operator definitions and evolution properties of PDFs and DAs in QCD.

2.3 Collinear Distributions

QCD factorization separates short-distance perturbative dynamics from long-distance nonperturbative structure in high-energy processes. In processes characterized by a large momentum transfer $Q \gg \Lambda_{\text{QCD}}$, the cross-section can be written schematically as a convolution of perturbatively calculable hard-scattering kernels with universal hadronic matrix elements. In the collinear framework, transverse momenta of partons are integrated out up to the factorization scale μ, and the dominant non-perturbative objects are PDFs for inclusive reactions and DAs for exclusive reactions. Their operator definitions make explicit their light-cone structure and ensure gauge invariance through appropriate Wilson lines.

For a quark of flavor q in a hadron H with momentum P^μ, the leading-twist collinear PDF is defined as the forward matrix element of a gauge-invariant nonlocal light-cone operator,

$$f_{q/H,\Gamma}(x;\mu) = \int \frac{d\lambda}{2\pi}\, e^{-i\lambda x}\, \langle PS|\bar{\psi}_q(\xi^-)\, W(\xi^-,0)\,\Gamma\,\psi_q(0)|PS\rangle\Big|_{\overline{\mathrm{MS}},\,\mu}, \tag{2.7}$$

where $|PS\rangle$ denotes a proton plane-wave state with momentum $P^\mu = (P^t, 0, 0, P^z)$ and spin S. The Wilson line $W(\xi^-, 0)$ runs along the light-like direction to ensure color gauge invariance. Light-cone coordinates are defined as $\xi^\pm = (t \pm z)/\sqrt{2}$, and the variable $\lambda = P^+\xi^-$ renders the Fourier transform dimensionless. The distribution is renormalized in the $\overline{\mathrm{MS}}$ scheme at scale μ.

When spin degrees of freedom are considered, the cross-section depends on the polarization of the parton relative to the nucleon spin, leading to different types of spin-dependent PDFs. At leading twist, three independent collinear quark PDFs appear. For unpolarized, helicity, and transversity channels, the Dirac structures are chosen as $\Gamma = \gamma^+$, $\gamma^+\gamma^5$, and $i\sigma^{+\perp}\gamma^5$, respectively. These correspond to the number density, longitudinal spin density, and transverse spin density of quarks in a longitudinally or transversely polarized nucleon. Analogous gauge-invariant definitions apply to gluon distributions, constructed from light-cone correlators of the gluon field strength tensor.

In inclusive deep-inelastic scattering (DIS), $eH \to eX$, the cross-section at large momentum transfer Q^2 admits a leading-twist factorization into perturbatively calculable hard-scattering coefficients and universal PDFs [26]:

$$\sigma_{eH}(x, Q^2) = \sum_a \int_x^1 d\xi \, f_{a/H}(\xi; \mu_F) \, \hat{\sigma}_{ea}\left(\frac{x}{\xi}, Q^2; \mu_F, \mu_R\right) , \tag{2.8}$$

where $f_{a/H}(\xi; \mu_F)$ denotes the PDF for a parton of type a carrying longitudinal momentum fraction ξ of the hadron H at factorization scale μ_F. The quantity $\hat{\sigma}_{ea}$ represents the perturbatively calculable partonic cross section for electron–parton scattering, evaluated at renormalization scale μ_R.

The factorization scale μ_F separates long-distance, nonperturbative physics encoded in the PDFs from short-distance dynamics contained in the hard-scattering kernel, while μ_R governs the running of the strong coupling in the perturbative expansion. Although these scales are, in principle, independent, a common practical choice is $\mu_F = \mu_R = \mu$. As mentioned above, due to the running of the strong coupling and the presence of collinear divergences, the PDFs acquire a scale dependence governed by the DGLAP evolution equation,

$$\frac{\partial f(x; \mu)}{\partial \ln \mu} = \int_x^1 \frac{d\nu}{|\nu|} \mathcal{P}(\nu; \mu) f(x/\nu; \mu) , \tag{2.9}$$

where $\mathcal{P}(\nu; \mu)$ is the DGLAP kernel, also known as the splitting function, that encodes the probability of a parton splitting into two partons with momentum fractions ν and $1 - \nu$. The DGLAP kernel has been calculated to 5-loop in perturbation theory [27].

A crucial consequence of QCD factorization is the universality of PDFs: the same distributions extracted from DIS appear in other hard processes involving the same hadron. For example, in Drell–Yan lepton-pair production [28], electroweak boson production [29], and inclusive jet production [30], the cross sections factorize into convolutions of the same PDFs with process-dependent hard-scattering kernels. This universality enables global analyses in which a wide range of experimental data sets are combined to determine PDFs with controlled uncertainties, which can then be used to make precision predictions for other reactions.

Over the past decade, global analyses of proton PDFs have entered a high-precision era, largely driven by the final combined HERA I+II inclusive DIS measurements and by the wealth of differential data from LHC Runs I and II. Modern global fits are now performed consistently at next-to-next-to-leading order (NNLO) in QCD for all processes that play a significant role in constraining PDFs, often supplemented by electroweak corrections and refined treatments of correlated experimental systematics. These developments have reduced PDF uncertainties in phenomenologically crucial regions of Bjorken-x, while simultaneously exposing nontrivial tensions among datasets and methodological differences between fitting groups.

Several recent global PDF determinations, including CT18 [31], MSHT20 [32], and NNPDF4.0 [33], play a central role in current LHC phenomenology. These analyses, together with related global efforts, provide the main inputs to the PDF4LHC21 combination [34], which establishes a common framework for PDF usage and uncertainty estimation in LHC studies. Although these analyses share a common experimental foundation—including HERA inclusive and heavy-flavor DIS, fixed-target Drell–Yan data, Tevatron measurements, and an extensive set of LHC observables such as inclusive jets, vector-boson production, and top-quark pair production—their results exhibit meaningful differences in specific flavor sectors and kinematic regions.

The CT18 analysis represents a major update of the CT14 [35] generation and places strong emphasis on incorporating high-precision LHC differential measurements. The resulting NNLO PDFs show improved constraints on the gluon distribution in the region $10^{-3} \lesssim x \lesssim 0.3$, driven primarily by inclusive jet and top-pair production data, and a revised strange-quark distribution influenced by

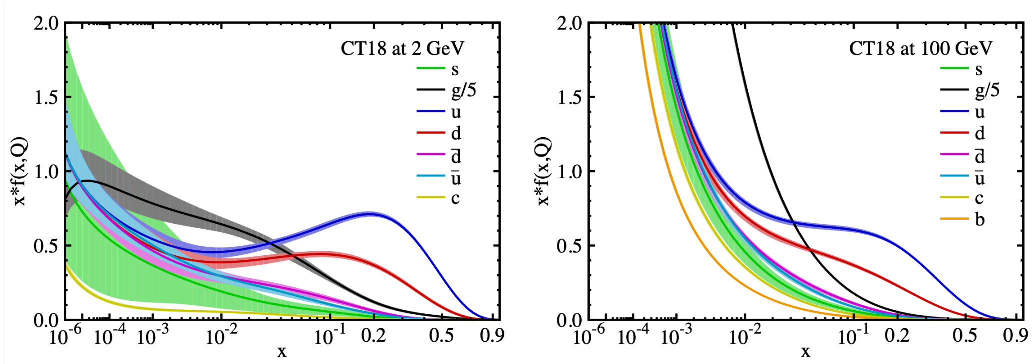

Figure 2.2: The CT18 PDFs at Q=2 GeV (left) and Q=100 GeV(right) [31].

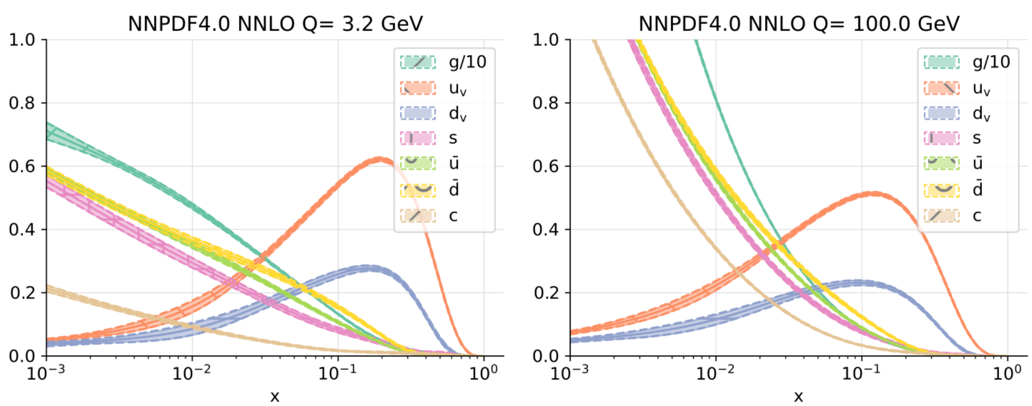

Figure 2.3: The NNPDF4.0 NNLO PDFs at Q= 3.2 GeV (left) and Q= 100 GeV (right) [33].

LHC W/Z measurements. As illustrated in Fig. 2.2, the CT18 gluon uncertainty at intermediate x is significantly reduced compared to earlier generations, while the strange-quark distribution exhibits a noticeable enhancement relative to CT14 in certain x regions. At small x, the combined HERA data continue to dominate the constraints.

The NNPDF4.0 determination [33] extends the global dataset further and reports a general reduction of PDF uncertainties relative to earlier NNPDF releases. Its phenomenological picture is largely consistent with CT18, featuring a well-constrained gluon at intermediate x and an enhanced strange sea compared to older fits, as illustrated in Fig. 2.3. Nevertheless, visible spreads among central values persist, particularly in the detailed flavor decomposition of the light sea and in the large-x gluon region.

The PDF4LHC21 combination [34], based on CT18, MSHT20, and NNPDF3.1, reflects this overall level of agreement while retaining the residual spread among global fits as a measure of systematic differences.

Taken together, these modern global analyses of unpolarized PDFs exhibit substantial consistency across most kinematic regions relevant for LHC phenomenology, while still displaying non-negligible variations in specific flavor sectors, in particular the strange distribution and the large-x gluon. These residual differences highlight the importance of complementary first-principles inputs, such as lattice QCD calculations, which offer an independent avenue to constrain and ultimately validate aspects of proton structure beyond the phenomenological global-fit framework. This complementarity is even more pronounced in the case of polarized PDFs, where the experimental constraints are considerably more limited and global fits rely on a much smaller dataset, thereby increasing the potential impact of first-principles determinations.

In comparison, DAs describe longitudinal momentum sharing among valence partons in hard exclusive processes, such as the B meson decay $B \to \pi\pi$ and $B \to l\nu\pi$ [36, 37]. For a pseudoscalar meson M with momentum p^μ, the leading-twist light-cone DA $\phi_M(x;\mu)$ is defined as

$$\int \frac{d\xi^-}{2\pi} e^{ixp^+\xi^-} \langle 0|\bar{\psi}_1(0)\not{n}\gamma_5 U(0,\xi^-)\psi_2(\xi^-)|M(P)\rangle = if_M(p\cdot n)\phi_M(x;\mu)\,, \tag{2.10}$$

where $U(0,\xi^-) = P\exp[ig_s\int_{\xi_-}^{0} ds\; n\cdot A(sn)]$ is the path-ordered gauge link defined along the minus light-cone direction n ($n^2 = 0$), and f_M the decay constant.

Similar to the factorization-scale dependence of PDFs, DAs depend on the renormalization scale

μ through the ERBL evolution equation,

$$\frac{\partial \phi(x;\mu)}{\partial \ln \mu} = \int_0^1 dy\, V(x,y;\mu)\, \phi(y;\mu)\,, \tag{2.11}$$

where $V(x,y;\mu)$ is the ERBL kernel that governs the evolution of a quark-antiquark pair carrying momentum fractions $(y, 1-y)$ into a pair with fractions $(x, 1-x)$.

As discussed above, DAs are probability amplitudes and are conceptually analogous to wave functions in quantum mechanics. Unlike PDFs, however, DAs describe the correlated longitudinal momentum distribution of a quark–antiquark pair inside a meson. Consequently, experimental observables do not probe the DA at a fixed value of x; instead, the DA enters physical amplitudes through a convolution with a perturbatively calculable hard-scattering kernel. Most phenomenological constraints on DAs are obtained from exclusive processes such as the $\pi \to \gamma\gamma^*$ transition form factor [38] and electromagnetic form factors [39]. For instance, the pion transition form factor $F_{\pi\gamma}(Q^2)$ can be written schematically as a convolution of the hard kernel $T_H(x, Q^2; \mu)$ with the pion DA $\phi_\pi(x;\mu)$. Early measurements of this observable by the CLEO collaboration [40] up to $Q^2 \sim 9$ GeV^2 were broadly consistent with perturbative QCD expectations based on the asymptotic form $\phi_\pi(x) = 6x(1-x)$.

However, the BABAR experiment [41], which extended the measurement to $Q^2 \sim 40$ GeV^2, reported a continued growth of $Q^2 F_{\pi\gamma}(Q^2)$ at large Q^2, exceeding the asymptotic prediction and suggesting a significantly broader or endpoint-enhanced pion DA. This unexpected trend triggered extensive theoretical investigations and debate regarding possible modifications of the DA shape, higher-twist effects, or limitations of collinear factorization in the accessible kinematic regime. Subsequent measurements by the Belle collaboration [42] did not confirm such a strong rise and were instead more consistent with perturbative QCD expectations, partially alleviating the so-called "BABAR puzzle" [43]. As a result, experimental data constrain only specific moments or weighted integrals of the DA rather than its full x-dependence. To bridge this gap, various nonperturbative approaches have been developed to model the shape of the DA, including QCD sum rules [44–46] and Dyson–Schwinger equation studies [47]. Nevertheless, these approaches typically rely on model assumptions or truncations, which makes first-principles calculations of x-dependent DAs particularly valuable.

2.4 Momentum Tomography: TMDs

While collinear parton distributions provide a one-dimensional description of hadron structure in terms of longitudinal momentum fractions, many high-energy processes are sensitive to the intrinsic transverse motion of partons. When the transverse momentum $q_\perp \equiv |\vec{q}_\perp| = \sqrt{q_x^2 + q_y^2}$ of an observed final state is parametrically smaller than the hard scale, $\Lambda_{\rm QCD} \lesssim q_\perp \ll Q$, a more differential description is required. In this regime, QCD factorization must retain transverse momentum degrees of freedom, leading to the framework of TMDs. These objects encode a multidimensional momentum-space structure and therefore provide access to momentum tomography beyond the collinear approximation.

The necessity of this refinement can be understood from the general logic of QCD factorization, whose validity relies on a systematic separation of scales. In inclusive observables characterized by a single hard scale Q, collinear factorization exploits the hierarchy $Q \gg \Lambda_{\rm QCD}$ and integrates over transverse momenta up to the hard scale, thereby absorbing long-distance physics into universal collinear PDFs. However, when the observable is differential in a small transverse momentum $q_\perp \ll Q$, an additional scale enters the problem. This small $q_\perp$ acts as a new infrared-sensitive scale, and the simple separation between hard and collinear modes is no longer sufficient. Large logarithms of the form

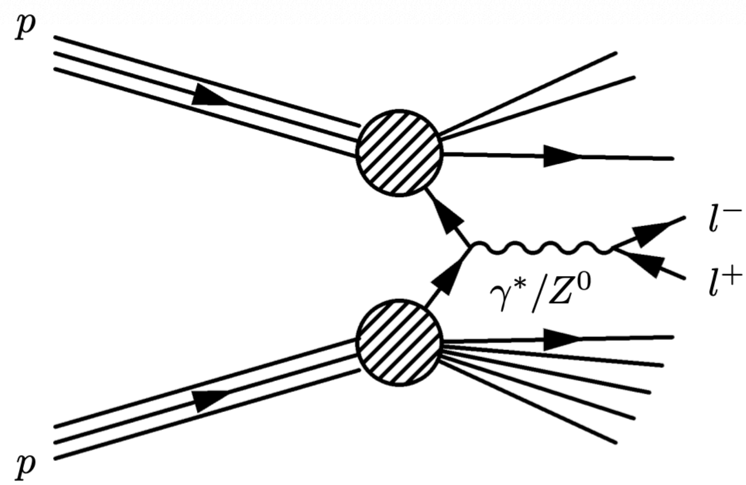

Figure 2.4: The diagrammatic illustration of the Drell–Yan process. Credit to Ref. [48].

$\ln(Q/q_\perp)$ emerge in perturbation theory, signaling the breakdown of fixed-order expansions within the collinear framework.

To restore predictive power, the factorization structure must be reorganized to consistently separate hard, collinear, and soft contributions at the level of transverse momentum. This leads to TMD factorization, in which the nonperturbative physics is encoded in TMDs, accompanied by an explicit soft factor and appropriate rapidity regularization and evolution. In this way, the appearance of a new hierarchical scale necessitates a corresponding refinement of the factorization theorem and the definition of the underlying nonperturbative functions.

More specifically, the physical motivation for TMD factorization can be illustrated using the Drell–Yan process, $pp \to l^+l^-X$. In this reaction, we consider the light-cone components of the lepton pair momentum $q^\mu = (q^+, q^-, \vec{q}_\perp)$, where $q^+ = (q^0 + q^3)/\sqrt{2}$ and $q^- = (q^0 - q^3)/\sqrt{2}$. The invariant mass Q of the lepton pair sets the hard scale, while the measured transverse momentum $q_\perp$ of the pair probes the transverse motion of the annihilating partons. The diagrammatic illustration of this process is shown in Fig. 2.4. When $q_\perp \sim Q$, the transverse momentum arises perturbatively from hard emissions and the standard collinear factorization formula suffices. In contrast, when $q_\perp \ll Q$, the cross-section becomes sensitive to soft and collinear radiation as well as to intrinsic transverse momentum inside the incoming hadrons. In this region, the differential cross-section can be organized as [49]

$$\frac{d\sigma}{dQ^2\,dy\,d^2\vec{q}_\perp} = (W(Q,\vec{q}_\perp) + Y(Q,\vec{q}_\perp))\left[1 + \mathcal{O}\left(\frac{\Lambda^2_{\text{QCD}}}{Q^2}\right)\right] , \tag{2.12}$$

where $y \equiv \frac{1}{2}\ln\left(\frac{q^+}{q^-}\right)$ defines the rapidity of the lepton pair, the W-term resums logarithms dominant in the small-$q_\perp$ region, while the Y-term accounts for the regular fixed-order contributions that remain finite as $q_\perp \to 0$. At leading power for $q_\perp \ll Q$, the W-term factorizes as

$$W(Q,\vec{q}_\perp) = \sum_q H_{q\bar{q}}(Q;\mu)\int d^2\vec{b}_\perp e^{i\vec{b}_\perp\cdot\vec{q}_\perp} f_{q/H_1}(x_1,\vec{b}_\perp;\mu,\zeta_1) f_{\bar{q}/H_2}(x_2,\vec{b}_\perp;\mu,\zeta_2) , \tag{2.13}$$

where $H_{q\bar{q}}(Q;\mu)$ is the perturbatively calculable hard function, and $f_{q/H}(x,\vec{b}_\perp;\mu,\zeta)$ denotes the renormalized unpolarized TMD PDF in impact-parameter $b_\perp \equiv |\vec{b}_\perp|$ space, Fourier conjugate to transverse momentum. The variables $\zeta_{1,2}$ are Collins–Soper scales associated with rapidity regularization, satisfying $\zeta_1\zeta_2 = Q^4$. This structure, originally derived in the Collins–Soper–Sterman (CSS) formalism and subsequently refined in modern formulations, establishes the separation of short-distance physics at scale Q from long-distance transverse dynamics encoded in universal TMDs [49].

The necessity of working in transverse coordinate space arises from the fact that large logarithms of the form $\ln(Q^2/q_\perp^2)$ appear at fixed order when $q_\perp \ll Q$. In $b_\perp$ space these logarithms exponentiate and can be resummed systematically through renormalization group evolution. The small-$b_\perp$ region corresponds to perturbative transverse momenta $q_\perp \sim b_\perp^{-1} \gg \Lambda_{\text{QCD}}$, where TMDs can be matched

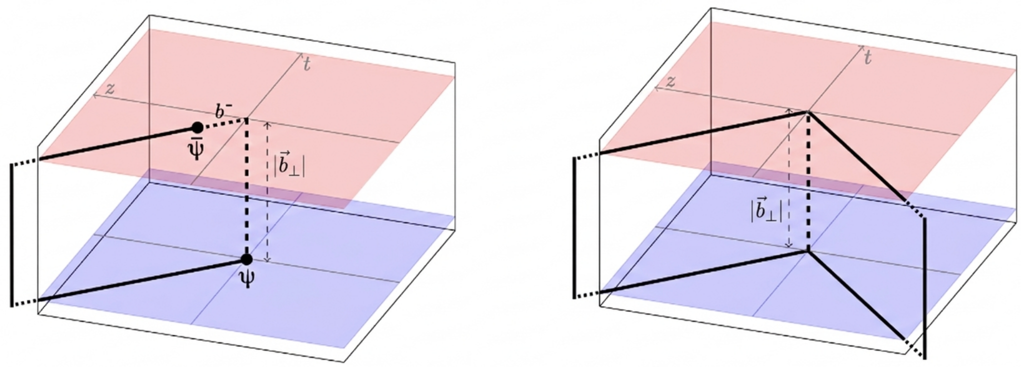

Figure 2.5: Graphs of the Wilson line structure of bare unsubtracted TMD correlator (left) and bare soft function (right). Credit to Ref. [49].

onto collinear PDFs via an OPE [50, 51],

$$f_{q/H}(x, \vec{b}_\perp; \mu, \zeta) = \sum_i \int_x^1 \frac{d\xi}{\xi} C_{q\leftarrow i}\left(\frac{x}{\xi}, \vec{b}_\perp; \mu, \zeta\right) f_{i/H}(\xi; \mu) + \mathcal{O}\big(b_\perp^2 \Lambda_{\rm QCD}^2\big) \ . \tag{2.14}$$

The coefficient functions $C_{q\leftarrow i}$ are perturbatively calculable. In this way, TMDs interpolate between perturbative transverse dynamics at small $b_\perp$ and genuinely nonperturbative physics at large $b_\perp$, providing a unified description across momentum scales.

At the operator level, TMD PDFs are defined as gauge-invariant nonlocal matrix elements of quark or gluon fields separated along both light-cone and transverse directions. For a quark of flavor q in a proton moving predominantly along the $+$ light-cone direction, the bare unsubtracted TMD correlator in coordinate space takes the schematic form

$$\begin{aligned} f_{q/H}^{0,\,\rm unsub}(x, \vec{b}_\perp, P^+, \epsilon, \tau) = \int \frac{db^-}{2\pi} e^{-ixP^+b^-} \\ \times \Big\langle PS \Big| \Big[\bar{\psi}_q^0(b^-, \vec{b}_\perp)\, \mathcal{W}_{\rm staple}^\dagger(b^-, \vec{b}_\perp) \Gamma \mathcal{W}_{\rm staple}(0) \psi_q^0(0)\Big]_\tau \Big| PS \Big\rangle \ , \end{aligned} \tag{2.15}$$

where the staple-shaped Wilson lines $\mathcal{W}_{\rm staple}$ extend to light-cone infinity and include transverse gauge links to ensure full gauge invariance. The diagrammatic illustration of the staple-shaped Wilson lines is shown in the left panel of Fig. 2.5. In contrast to collinear PDFs, TMD correlators contain rapidity divergences associated with gluons of arbitrarily large rapidity. These divergences are not regulated by dimensional regularization alone and require an additional rapidity regulator, so we have ϵ and τ as the dimensional regulator and rapidity regulator, respectively. After combining the unsubtracted bare TMD matrix element with a bare soft factor S^0, a soft subtraction factor $S^{0\text{subt}}$ that subtracts double counting of soft radiation, and performing ultraviolet renormalization in the $\overline{\text{MS}}$ scheme, one obtains the properly defined TMD PDF $f_{q/H}(x, \vec{b}_\perp; \mu, \zeta)$ [49]. The bare soft factor S^0 is defined as a vacuum matrix element of Wilson lines,

$$S^0(\vec{b}_\perp, \epsilon, \tau) = \frac{1}{N_c} \left\langle 0 \middle| \left[\mathcal{W}_n^\dagger(\vec{b}_\perp) \mathcal{W}_{\bar{n}}(\vec{b}_\perp) \mathcal{W}_{\bar{n}}^\dagger(0) \mathcal{W}_n(0) \right]_\tau \middle| 0 \right\rangle , \tag{2.16}$$

where the shape of Wilson lines is presented in the right panel of Fig. 2.5. Regarding the soft subtraction factor $S^{0\text{subt}}$, in SCET these subtractions are known as zero-bin subtractions [52] and arise from ensuring fluctuations encoded by collinear fields do not have singular overlap with those of the soft fields, thus also ensuring there is no double counting of infrared regions. More details about the subtraction factor can be found in Ref. [49].

A central structural difference relative to collinear PDFs is therefore the presence of two evolution scales. In addition to the usual renormalization scale μ, TMDs depend on a rapidity scale ζ, whose variation is governed by the Collins–Soper equation [53],

$$\frac{\partial}{\partial \ln \sqrt{\zeta}} f_{q/H}(x, \vec{b}_\perp, \mu, \zeta) = K(b_\perp, \mu)\, f_{q/H}(x, \vec{b}_\perp, \mu, \zeta) , \tag{2.17}$$

where $K(b_\perp, \mu)$ is the Collins–Soper kernel. Together with the renormalization group equation in μ, this yields a coupled evolution system that resums Sudakov logarithms and controls the scale dependence of TMD cross-sections. In contrast to DGLAP evolution, which mixes different momentum fractions x, Collins–Soper evolution is multiplicative in x and acts at fixed transverse separation $b_\perp$.

When spin degrees of freedom are included, the TMD framework reveals a much richer structure than in the collinear case. At leading twist, the quark TMD correlator for a spin-$1/2$ nucleon can be decomposed into eight independent functions, corresponding to different correlations between parton transverse momentum and nucleon spin. In addition to the unpolarized distribution, one encounters helicity and transversity TMDs, as well as time-reversal-odd functions such as the Sivers and Boer–Mulders distributions. The presence of staple-shaped Wilson lines implies that certain TMDs are process dependent, differing by an overall sign between semi-inclusive deep-inelastic scattering and Drell–Yan. This so-called modified universality arises from the opposite directions of the gauge links in the two processes and is consistent with time-reversal invariance once Wilson lines are properly taken into account.

The phenomenological extraction of TMDs combines data from Drell–Yan, semi-inclusive deep-inelastic scattering, and back-to-back hadron production in e^+e^- annihilation. In each case, the measured small transverse momentum provides sensitivity to intrinsic transverse motion and to spin-momentum correlations. The interplay between perturbative matching at small $b_\perp$, nonperturbative modeling at large $b_\perp$, and coupled (μ, ζ) evolution equations allows for global analyses of unpolarized and polarized TMDs, analogous in spirit to global fits of collinear PDFs but with substantially richer kinematic structure.

In summary, TMD factorization extends the one-dimensional parton model into three-dimensional momentum space. By retaining transverse degrees of freedom and properly accounting for soft radiation and rapidity evolution, it provides a theoretically consistent and phenomenologically powerful framework for describing processes at small transverse momentum.

2.5 Spatial Tomography: GPDs

While collinear PDFs describe the longitudinal momentum structure of partons and TMDs extend this picture into three-dimensional momentum space, GPDs provide complementary access to the spatial structure of hadrons. In particular, GPDs correlate the longitudinal momentum of partons with the transverse spatial distribution of their color charge, thereby enabling a spatial tomography of the nucleon.

GPDs appear in hard exclusive processes in which the hadron remains intact and the full final state is reconstructed. The paradigmatic examples are deeply virtual Compton scattering (DVCS) [54, 55], and deeply virtual meson production (DVMP) [56, 57]. Unlike collinear PDFs, GPDs depend on additional kinematic variables that encode the non-forward nature of the process. In the forward limit, where the momentum transfer to the hadron vanishes, GPDs reduce to ordinary PDFs. On the other hand, integration over the momentum fraction x relates GPDs to elastic form factors. In this way, GPDs interpolate smoothly between inclusive parton distributions and exclusive hadronic observables, unifying momentum and spatial information within a single framework.

A particularly important aspect of GPDs is that their Fourier transform with respect to the transverse momentum transfer provides impact-parameter dependent distributions. In a frame where the longitudinal momentum transfer is fixed, this transformation yields parton densities in transverse coordinate space, offering a probabilistic interpretation of the transverse spatial distribution of partons at fixed longitudinal momentum fraction. This connection underlies the interpretation of GPDs as tools for spatial imaging of the nucleon.

Furthermore, GPDs contain information on the total angular momentum carried by quarks and gluons through Ji's sum rule [54], which relates the second moment of specific GPD combinations to the total partonic contribution to the nucleon spin. This establishes a direct link between exclusive scattering measurements and the decomposition of nucleon spin in QCD.

Experimentally, GPDs are constrained by measurements of DVCS and exclusive meson production at fixed-target facilities and colliders. While the kinematic coverage remains limited compared to inclusive DIS, ongoing and future programs, in particular at high-luminosity electron-ion facilities, are expected to significantly improve constraints on the multidimensional structure encoded in GPDs.

2.6 Conclusion

In this chapter, we have reviewed how the study of partonic structure has evolved into a coherent and multidimensional description of hadrons within QCD. Collinear PDFs quantify the longitudinal momentum and spin structure and underpin precision collider phenomenology; TMDs resolve intrinsic transverse motion and spin-momentum correlations, revealing nonperturbative dynamics at the few-hundred-MeV scale; and GPDs correlate momentum and transverse spatial distributions, enabling access to impact-parameter tomography and to energy-momentum tensor form factors that encode mass, angular momentum, and mechanical properties. Together, these distributions form a unified framework for hadron tomography, supported by QCD factorization, increasingly precise global analyses, and complementary first-principles input from lattice QCD. With ongoing measurements at facilities such as Jefferson Lab and the Large Hadron Collider, and future advances anticipated at the Electron-Ion Collider, partonic physics stands as a mature and rapidly progressing program aimed at a quantitative understanding of how quarks and gluons generate the observable properties of hadrons.

Chapter 3: First-Principles Calculation of Strong Interaction: Lattice QCD

Lattice QCD emerged from the need to address quark confinement and to formulate QCD beyond the domain of perturbation theory. In 1974, Kenneth G. Wilson proposed the discretization of non-Abelian gauge theories on a space-time lattice and, in his seminal paper "Confinement of Quarks" [58], demonstrated through a strong-coupling expansion that lattice Yang-Mills theory exhibits an area law behavior for large Wilson loops. This area law implies a linearly rising potential between static color sources at large separations, thereby providing a concrete realization of confinement within a gauge-invariant, nonperturbative framework. The recovery of continuum QCD from the lattice formulation was clarified later through the development of renormalization-group arguments and lattice perturbation theory. Asymptotic freedom [2, 3] implies that the continuum limit corresponds to the ultraviolet fixed point at vanishing bare coupling, and it was subsequently established that lattice gauge theory is renormalizable and belongs to the same universality class as continuum Yang–Mills theory [59–61]. Although this formulation established the conceptual foundation of lattice gauge theory, practical calculations of hadronic observables required the later development of efficient numerical algorithms and sufficient computational resources.

By 1979-1980, the development of Monte Carlo simulation techniques transformed Wilson's proposal into a practical computational framework. Michael Creutz and collaborators carried out the first large-scale numerical simulations of non-Abelian gauge theory on the lattice, including SU(2) Yang-Mills theory, and directly measured Wilson loops and the static quark-antiquark potential [62]. The observed linear rise of the potential at large separations provided numerical evidence for confinement, consistent with the area-law behavior anticipated in the strong-coupling expansion. Shortly thereafter, it was demonstrated that hadron spectroscopy could be addressed from first principles: early calculations of meson and baryon masses by Weingarten, and by Hamber and Parisi, showed that lattice QCD could reproduce key features of the hadron spectrum [63–65]. These pioneering works established lattice gauge theory as a viable ab initio approach to nonperturbative QCD.

From the 1980s onward, progress in algorithms and computing power significantly broadened the scope of lattice QCD. The formulation of efficient algorithms to include dynamical fermions—most notably the Hybrid Monte Carlo (HMC) method [66]—made it possible to incorporate the fermion determinant in a controlled manner. Over time, simulations progressed from the quenched approximation to fully dynamical calculations with 2, 2+1, and eventually 2+1+1 quark flavors. By the 2000s and 2010s, lattice ensembles at or near physical quark masses and large volumes became feasible [67]. As a result, modern lattice QCD calculation provides high-precision determinations of hadronic observables, ranging from spectroscopy and decay constants to form factors and weak matrix elements, after systematic extrapolations to the continuum limit, infinite volume, and physical quark masses.

In recent years, lattice QCD has further expanded beyond traditional spectroscopy and matrix elements to address the internal structure of hadrons in terms of quarks and gluons. The proposal of the large momentum effective theory (LaMET) [68–71] opened a new avenue for computing x-dependence of parton distribution functions directly on the lattice through quasi-distributions defined at finite hadron momenta. Building on this framework, current lattice calculations now achieve controlled theoretical uncertainties for PDFs in the moderate-x region [72]. Exploratory studies of TMDs and GPDs are also underway, marking a transition from static properties to a multidimensional, partonic description of hadrons within first-principles QCD. More detailed discussions on lattice calculations of partonic structure will be presented in Chapter 4.

3.1 Lattice QCD as a First-Principles Framework

Lattice QCD provides a nonperturbative definition of QCD by promoting the formal path integral of the continuum theory to a mathematically well-defined statistical system. After Wick rotation to Euclidean spacetime, vacuum expectation values take the form

$$\langle \mathcal{O} \rangle = \frac{1}{Z} \int \mathcal{D}U\, \mathcal{D}\bar{\psi}\, \mathcal{D}\psi\; \mathcal{O}\, e^{-S_E[U,\bar{\psi},\psi]} \,, \tag{3.1}$$

which can be interpreted as an ensemble average over gauge-field configurations weighted by e^{-S_E}. Discretizing spacetime into a hypercubic lattice of spacing a introduces an ultraviolet regulator while preserving exact gauge invariance: the gauge field $A_\mu(x)$ is replaced by link variables

$$U_\mu(x) = \exp[iagA_\mu(x)] \,, \tag{3.2}$$

living on links between neighboring sites. The simplest gauge action is the Wilson plaquette action [58], constructed from elementary 1×1 Wilson loops (plaquette), which reproduces the Yang–Mills action in the classical continuum limit. Together with a discretized fermion action, this defines a regulated version of QCD in which all ultraviolet divergences are removed by the finite cutoff $1/a$.

Conceptually, lattice QCD realizes Wilson's renormalization-group picture: long-distance physics is insensitive to the microscopic details of the regulator. The lattice spacing plays the role of a short-distance cutoff, while physical observables depend only on the correlation length measured in lattice units. Asymptotic freedom implies that the continuum limit corresponds to tuning the bare coupling toward the ultraviolet fixed point at $g_0 \to 0$, such that the physical correlation length becomes much larger than the lattice spacing. In this regime, different discretizations of the action lie in the same universality class and yield identical continuum predictions after renormalization. The continuum QCD Lagrangian thus emerges as the critical theory describing the long-distance behavior of the lattice system.

Beyond providing a regulator, the lattice formulation makes the theory computationally accessible. Because the Euclidean path integral resembles a partition function of a four-dimensional statistical system, importance-sampling Monte Carlo methods can be employed to generate representative gauge-field configurations. Observables are computed as ensemble averages over these configurations, without reliance on perturbative expansions. Nonperturbative phenomena such as confinement, spontaneous chiral symmetry breaking, and topological fluctuations arise dynamically from the underlying gauge dynamics rather than from model assumptions.

Moreover, the framework is systematically improvable. Discretization effects can be reduced through optimized discretization schemes, such as Symanzik improvement [59, 73], and continuum extrapolation by performing simulations at multiple lattice spacings. Finite-volume effects are controlled by increasing the physical volume and are quantitatively understood through effective field theory. Quark masses can be tuned to their physical values, and renormalization procedures connect lattice operators to continuum schemes. Modern simulations therefore analyze multiple ensembles varying a, L, and quark masses in order to perform controlled extrapolations to the continuum, infinite-volume, and physical-mass limits.

In summary, lattice QCD provides a first-principles, nonperturbative formulation of the strong interaction directly from the QCD Lagrangian. By defining the theory with a gauge-invariant ultraviolet regulator and evaluating the Euclidean path integral numerically, it converts the formal definition of quantum field theory into a controllable statistical system. The framework is systematically improvable through well-defined extrapolations. Crucially, no model assumptions about confinement or chiral symmetry breaking are imposed. In this sense, lattice QCD furnishes a quantitatively reliable bridge between the fundamental gauge theory and observable hadronic physics in the nonperturbative regime.

3.2 Fermion Formulations on the Lattice

While gauge fields admit a geometrically natural discretization in terms of link variables, the formulation of fermions on the lattice is considerably more subtle. The difficulty already arises at the level of the free Dirac field and is rooted in the so-called fermion doubling problem [74]. A naive transcription of the continuum Dirac action to the lattice, obtained by replacing derivatives with symmetric finite differences, leads in four dimensions to sixteen fermionic modes in the continuum limit instead of a single physical species.

To see the origin of the problem, consider the Euclidean lattice action for a free fermion obtained by replacing ∂_μ with the antihermitian symmetric difference operator. In momentum space, the inverse propagator takes the form

$$D(p) = i\sum_\mu \gamma_\mu \tilde{p}_\mu + M\ , \qquad \tilde{p}_\mu = \frac{1}{a}\sin(p_\mu a)\ , \tag{3.3}$$

with p_μ restricted to the Brillouin zone $(-\pi/a, \pi/a]$. While $\tilde{p}_\mu \to p_\mu$ for small momenta, the sine function also vanishes at the edges of the Brillouin zone. As a consequence, the propagator develops additional poles near $p_\mu = \pi/a$ (and permutations thereof). In d space-time dimensions this yields 2^d fermionic excitations in the continuum limit, fifteen of which are lattice artefacts in four dimensions. These extra modes are not removable by simply discarding high-momentum contributions, since they are associated with symmetry transformations of the lattice action that relate the different corners of the Brillouin zone.

The doubling phenomenon is closely tied to fundamental structural requirements. The Nielsen-Ninomiya theorem [75] states that, under the assumptions of locality, translational invariance, and hermiticity, a lattice fermion formulation with exact chiral symmetry necessarily contains fermion doublers. Therefore, eliminating the unwanted species requires relaxing at least one of these conditions. Most practical formulations break chiral symmetry explicitly at nonzero lattice spacing, with the understanding that it is recovered in the continuum limit.

The Wilson formulation [76] achieves this by adding a second-derivative term (the Wilson term) to the naive action. The Wilson-Dirac operator reads

$$D_W = i\sum_{\mu} \gamma_\mu \tilde{p}_\mu + M + \frac{r}{a}\sum_{\mu}\left(1-\cos(p_\mu a)\right) \ , \tag{3.4}$$

where r is the Wilson parameter. The additional term acts as a momentum-dependent mass that vanishes for small momenta but grows as $1/a$ near the edges of the Brillouin zone. Consequently, the doublers acquire masses of order the cutoff and decouple in the continuum limit. The price paid is the explicit breaking of chiral symmetry at finite lattice spacing, which leads to additive mass renormalization and necessitates fine tuning of the bare mass parameter.

Discretization errors in Wilson fermions are generically of $\mathcal{O}(a)$, which can be reduced by Symanzik improvement [59, 73]. In particular, the clover-improved Wilson action (Sheikholeslami-Wohlert improvement) [77] introduces an additional dimension-five operator proportional to $\sigma_{\mu\nu}F_{\mu\nu}$, which cancels $\mathcal{O}(a)$ effects in on-shell quantities at tree level and, with appropriate tuning, beyond tree level. Clover fermions are widely used in large-scale simulations due to their balance between theoretical control and computational cost.

An alternative strategy is provided by staggered (Kogut-Susskind) fermions [78, 79]. Instead of lifting the doubler masses, the staggered approach reduces the number of fermionic degrees of freedom per lattice site by distributing spin components over neighboring sites. Through a spin-diagonalization transformation, the naive Dirac structure is replaced by position-dependent phases, leaving a single Grassmann variable per site. In four dimensions, this procedure reduces the sixteen naive species to four so-called "tastes." The resulting action preserves a remnant $U(1) \times U(1)$ chiral symmetry at vanishing bare mass, which protects against additive mass renormalization. However, because four tastes remain in the continuum limit, practical simulations with $N_f = 2$ or $2 + 1$ flavors require the rooting procedure, in which the fermion determinant is replaced by an appropriate fractional power. Although supported by substantial numerical evidence, the theoretical justification of rooting has been the subject of extensive discussion.

Formulations that preserve exact lattice chiral symmetry at finite a are obtained by implementing the Ginsparg-Wilson relation [80]. Overlap fermions [81] provide an explicit realization of this relation via a nonlocal Dirac operator constructed from the sign function of a Hermitian Wilson operator. Domain-wall fermions [82, 83] introduce an additional fifth dimension, in which chiral modes are localized on opposite boundaries; in the limit of infinite fifth dimension, they approach overlap fermions. These formulations possess an exact lattice-modified chiral symmetry and automatically reproduce the correct axial anomaly. They are therefore particularly well suited for studies of chiral symmetry breaking, topological properties, and flavor-singlet observables. Their principal drawback is the significantly increased computational cost compared to Wilson or staggered fermions.

All of these discretizations represent distinct regularizations of continuum QCD. Provided that the lattice action is local and gauge invariant, universality guarantees convergence to the same continuum theory as $a \to 0$. Nevertheless, the pattern of lattice artefacts, the realization of chiral symmetry, and the computational efficiency differ markedly among formulations. Systematic continuum extrapolations and cross-checks between different fermion actions therefore constitute an essential component of precision lattice QCD. The coexistence of multiple fermion formulations is not merely a historical artifact but a practical and conceptual asset, enabling controlled assessments of discretization effects and strengthening the reliability of lattice results as first-principles determinations of hadronic observables.

3.3 Monte Carlo Algorithms for Lattice Gauge Ensembles

The generation of lattice gauge ensembles amounts to constructing a stochastic process whose stationary distribution coincides with the Euclidean path-integral measure,

$$\pi[U] \propto e^{-S_g[U]} \tag{3.5}$$

for pure gauge theories, or

$$\pi[U] \propto e^{-S_g[U]} \det D[U] \tag{3.6}$$

for theories including dynamical fermions. Here, $S_g[U]$ is the discretized gauge action, and $\det D[U]$ arises from integrating out the fermion fields. The goal is to sample gauge-field configurations U according to $\pi[U]$, enabling the computation of observables as ensemble averages. In practice, this is achieved through Markov Chain Monte Carlo (MCMC) algorithms, which define a transition kernel $K(U \to U')$ such that $\pi[U]$ is invariant under the evolution of the chain. Two fundamental requirements ensure correctness of the sampling procedure: ergodicity and stationarity. Ergodicity guarantees that all physically relevant regions of configuration space can be reached from any initial configuration in a finite number of steps, while stationarity ensures that $\pi[U]$ is preserved under the transition dynamics. A stationary distribution $\pi[U]$ satisfies the global balance condition,

$$\pi[U] = \int \mathcal{D}U' \, \pi[U'] K(U' \to U) \; . \tag{3.7}$$

A sufficient, though not strictly necessary, condition for stationarity is detailed balance,

$$\pi[U] K(U \to U') = \pi[U'] K(U' \to U) \; , \tag{3.8}$$

which implies that the Markov chain is reversible with respect to $\pi[U]$. In addition to satisfying these formal requirements, practical algorithms must minimize autocorrelation times and computational cost, particularly as the lattice spacing decreases and critical slowing down becomes pronounced.

From this viewpoint, different Monte Carlo algorithms correspond to different constructions of the transition kernel. In the Metropolis–Hastings framework, the kernel is built from a proposal distribution $q(U \to U')$ together with an accept–reject step. Choosing the acceptance probability as

$$P_{\text{acc}}(U \to U') = \min\left(1, \frac{\pi[U'] \, q(U' \to U)}{\pi[U] \, q(U \to U')}\right) \tag{3.9}$$

ensures that the resulting kernel satisfies detailed balance. For symmetric proposals, $q(U \to U') = q(U' \to U)$, this reduces to the familiar Metropolis form

$$P_{\text{acc}} = \min\left(1, e^{-[S(U') - S(U)]}\right) \; . \tag{3.10}$$

In early lattice gauge simulations, such proposals often modified a single link or a small set of links at each step. While conceptually simple and straightforward to implement for pure gauge simulations, local update schemes suffer from strong autocorrelations as the correlation length increases, leading to critical slowing down near the continuum limit. With dynamical fermions, the nonlocal dependence of $\det D[U]$ renders naive local Metropolis proposals computationally inefficient.

Overrelaxation updates [84,85] provide a complementary strategy by constructing deterministic, approximately microcanonical transformations that preserve (or nearly preserve) the local action. In the formulation of Ref. [86], the overrelaxed heat-bath algorithm can be continuously interpolated between the conventional stochastic heat bath ($\omega = 1$) and a microcanonical limit ($\omega = 2$), in which the update becomes a deterministic constant-action reflection. Such updates move configurations along (nearly) constant-action surfaces and thereby reduce the diffusive, random-walk behavior characteristic of purely stochastic local updates. However, purely deterministic overrelaxation updates were observed to produce incorrect expectation values due to lack of ergodicity [86]. For this reason, overrelaxation steps must be alternated with conventional stochastic updates (such as heat-bath or Metropolis sweeps) to ensure ergodicity of the Markov chain. Although highly effective for pure gauge theories, overrelaxation alone does not incorporate the nonlocal fermion determinant and therefore does not address the fundamental computational challenges posed by dynamical fermions. It thus does not constitute a complete algorithm in that setting.

One of the dominant methods for simulations with dynamical fermions is Hybrid Monte Carlo (HMC) [66], which introduces an extended phase space in order to construct global proposals. Auxiliary conjugate momenta P are assigned to the gauge fields, and an artificial Hamiltonian is defined as

$$H(P, U) = \frac{1}{2}P^2 + S(U) \,, \tag{3.11}$$

so that the joint distribution in phase space factorizes as $\Pi(P, U) \propto e^{-H(P,U)} = e^{-\frac{1}{2}P^2}e^{-S(U)}$. Because the kinetic term depends only on P while the potential term depends only on U, integrating out the Gaussian momenta reproduces the desired gauge-field distribution $\pi[U] \propto e^{-S(U)}$. Hamilton's equations are then integrated along fictitious molecular-dynamics trajectories using a reversible and symplectic integrator. In the limit of exact integration the Hamiltonian would be conserved and $\Delta H = 0$, implying unit acceptance; in practice, discretization errors lead to small violations of energy conservation, and a final Metropolis accept–reject step with probability $\min(1, e^{-\Delta H})$ restores exact sampling.

The essential insight of HMC is that it resolves the usual tension between sampling efficiency and acceptance rate. In conventional local updates, small changes yield high acceptance but explore configuration space slowly, whereas large updates typically generate large ΔS and are rejected. In contrast, HMC produces global, collective changes in U through deterministic molecular-dynamics evolution that approximately conserves the Hamiltonian. The acceptance rate is therefore governed by the integrator error in ΔH, rather than directly by the magnitude of ΔS, allowing large moves in configuration space while maintaining high acceptance. Combined with stochastic momentum refreshment to ensure ergodicity, this mechanism greatly suppresses random-walk behavior and makes HMC substantially more efficient than purely local algorithms for simulations with dynamical fermions.

An alternative class of algorithms is based on stochastic differential equations, most notably Langevin dynamics [87]. In this approach, the gauge fields are evolved in a fictitious stochastic time t

according to the Langevin equation

$$dU_t = -\nabla S[U_t]\, dt + \sqrt{2}\, dW_t \,, \tag{3.12}$$

where the first term is a deterministic *drift* term proportional to the functional gradient of the action, driving the system toward regions of lower action, and dW_t denotes a Wiener process representing Gaussian white noise. The noise satisfies $\langle dW_t \rangle = 0$ and $\langle dW_t dW_{t'} \rangle = \delta(t - t')dt$, ensuring a balance between deterministic relaxation and stochastic diffusion. The probability density $P(U, t)$ for configurations evolving under this stochastic process obeys the associated Fokker–Planck equation,

$$\frac{\partial P(U,t)}{\partial t} = \nabla \cdot \left(\nabla S[U]\, P(U,t) + \nabla P(U,t) \right), \tag{3.13}$$

whose stationary solution is precisely $P_{\text{eq}}(U) \propto e^{-S(U)}$. Thus, in continuous stochastic time, Langevin evolution samples the desired path-integral measure without the need for a Metropolis accept–reject step. In practice, however, the stochastic differential equation must be discretized in finite time steps, typically leading to systematic step-size errors unless a corrective accept–reject mechanism is introduced, as in the stochastic quantization framework. While conceptually simple and local in structure, and appealing due to the absence of global momentum variables, Langevin-type algorithms are generally less efficient and less robust than HMC for real-action lattice QCD simulations, particularly at fine lattice spacings.

More recently, generalized MCMC strategies have been explored to mitigate critical slowing down and topological freezing in lattice gauge theory. While detailed balance is sufficient for correctness, it is not strictly necessary, and nonreversible Markov chains that preserve global balance have been shown in statistical sampling theory to reduce diffusive behavior by breaking time-reversal symmetry [88]. In the lattice context, algorithmic modifications aimed at improving topological tunneling and long-distance decorrelation include, for example, open boundary conditions [89] and related developments [90]. Although fully developed irreversible formulations for four-dimensional QCD remain under active investigation, such approaches reflect a broader effort to alter the dynamical structure of the Markov process in order to alleviate critical slowing down without changing the target distribution. Another direction introduces learned, invertible transformations of the gauge fields, such as normalizing flows [91, 92], which map configurations to a latent space where the distribution is simpler and proposals are more global. In such approaches, an invertible map $U \mapsto \tilde{U}$ with computable Jacobian is used to reshape the sampling manifold, and a Metropolis correction step ensures exactness of the resulting Markov chain. By effectively preconditioning configuration space and flattening energy barriers, these methods aim to enhance mixing and reduce autocorrelation times. Although promising and actively investigated in lattice field theory contexts, they introduce additional algorithmic and computational complexity and remain under ongoing development.

In summary, lattice ensemble generation can be understood as the construction of an ergodic Markov process with invariant distribution $\pi[U] \propto e^{-S[U]}$. Local Metropolis-type algorithms emphasize simplicity but suffer from critical slowing down; Langevin dynamics offers a stochastic differential perspective with limited practical efficiency; and Hybrid Monte Carlo, through its global Hamiltonian evolution and exact accept-reject correction, has emerged as the standard framework for dynamical fermion simulations. Ongoing algorithmic research continues to address the remaining challenges of scaling toward finer lattice spacings and larger physical volumes.

3.4 Benchmark Calculations: Axial Charge

One example is worth a thousand arguments. In this section, a benchmark calculation is presented to demonstrate the efficacy of lattice QCD and the control of relevant systematics. We focus on the nucleon axial charge g_A, which has historically shown tension with experiment mainly due to excited-state contaminations. This benchmark shows that these systematic effects can be quantified and controlled, yielding a reliable determination of a key single-hadron matrix element.

3.4.1 Nucleon Axial Charge

The nucleon axial charge, g_A, is a fundamental quantity in hadron structure and weak interactions, governing neutron β decay and providing a precision benchmark for lattice QCD. Experimentally, the ratio of the nucleon's axial to vector charges, g_A/g_V, is known at the sub-percent level from polarized ultracold-neutron measurements, including the UCNA result, $1.2772(20)$ [93, 94], and the PERKEO II result, 1.2761^{+14}_{-17} [95]. In the Standard Model, g_V is protected by vector-current conservation and is unity up to small isospin-breaking corrections [96, 97]; consequently, g_A/g_V directly constrains the axial structure of the nucleon.

Historically, lattice determinations of g_A tended to lie below experiment. Recent calculations, however, have reached a level of precision and systematic control sufficient for inclusion in the latest FLAG review [98]. The current status of lattice results and FLAG averages for the isovector axial charge g_A^{u-d} in $2+1$ and $2+1+1$ flavor simulations is shown in Fig. 3.1. As shown in the figure, there remains some tension in the recent lattice calculations. It is now broadly understood that excited-state contamination in nucleon three-point functions is a leading source of the lattice calculations. The analysis of Ref. [100] was therefore designed to quantify this contamination and to establish a robust

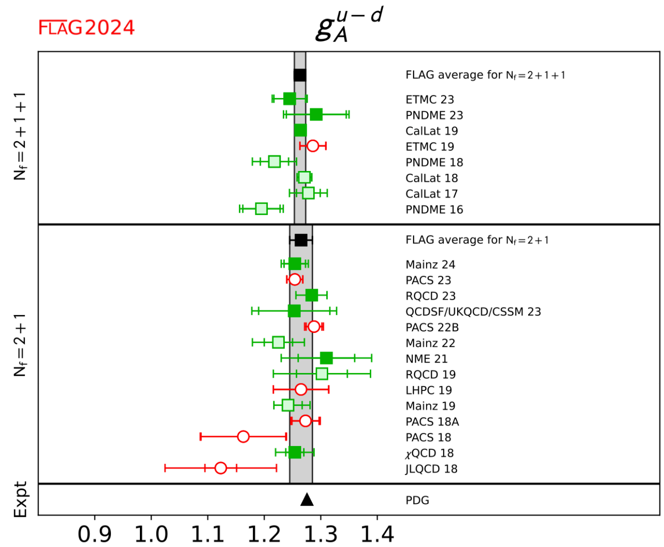

Figure 3.1: Lattice results and FLAG averages for the isovector axial charge g_A^{u-d} for 2 + 1 and 2 + 1 + 1 flavour calculations from Ref [98]. Also shown is the experimental result as quoted in the PDG [99].

extraction strategy for g_A.

The calculation was performed on an ensemble with lattice spacing $a \approx 0.09$ fm and pion mass $m_\pi \approx 310$ MeV, using Möbius domain-wall valence fermions on a Highly Improved Staggered Quark (HISQ) sea. The approximate chiral symmetry of the valence action implies that the axial and vector renormalization constants are equal up to small corrections, allowing the renormalized axial charge to be obtained from the ratio of bare matrix elements. This reduces renormalization-related systematics and isolates excited-state effects as the dominant uncertainty.

A central feature of this calculation is the use of a large set of source–sink separations, $t_{\text{sep}}/a \in \{2, 3, \ldots, 14\}$, corresponding to physical separations between approximately 0.18 and 1.22 fm. Conventional nucleon matrix-element calculations typically employ only a few values of t_{sep} concentrated near 1 fm or larger, where statistical noise becomes severe. In contrast, the present strategy exploits precise short- and intermediate-time data combined with controlled multi-state fits.

The lattice QCD calculation of g_A and g_V relies on the analysis of two-point and three-point correlation functions. The two-point function, $C_2(t)$, encodes the propagation of a nucleon state from source to sink, defined as

$$C_2\left(t_{\text{sep}}\right) = \textstyle\sum_{\vec{x}} \langle \Omega | N\left(t_{\text{sep}}, \vec{x}\right) N^{\dagger}(0, \vec{0}) | \Omega \rangle \tag{3.14}$$

$$= \textstyle\sum_{n=0}^{\infty} \left|z_n\right|^2 e^{-E_n t_{\text{sep}}} = \left|z_0\right|^2 e^{-E_0 t_{\text{sep}}} \left[1 + \sum_{n=1}^{\infty} \left|r_n\right|^2 e^{-\Delta_{n0} t_{\text{sep}}}\right] \; . \tag{3.15}$$

In this expression, we assume that the overlap factors, $z_n = \langle \Omega | N | n \rangle$, used to create ($N^{\dagger}$) and annihilate ($N$) the states with quantum numbers of the nucleon from the vacuum ($|\Omega\rangle$) are conjugate to each other. [1] In the last line, we have defined the energy splitting and ratio of overlap factors:

$$\Delta_{mn} = E_m - E_n \; , \quad r_n = \frac{z_n}{z_0} \, . \tag{3.16}$$

The three-point function, $C_\Gamma(t_{\text{sep}}, \tau)$, includes an insertion of the axial or vector current at an intermediate time τ. In the limit of zero momentum and zero-momentum transfer, it is defined as

$$\begin{aligned} C_\Gamma(t_{\text{sep}}, \tau) &= \sum_{\vec{y}, \vec{x}} \langle \Omega | N(t_{\text{sep}}, \vec{y}) \, j_\Gamma(\tau, \vec{x}) \, N^{\dagger}(0, \vec{0}) | \Omega \rangle && (3.17) \\ &= \sum_n |z_n|^2 \, g^\Gamma_{nn} \, e^{-E_n t_{\text{sep}}} + 2 \sum_{n<m} z_n z^\dagger_m \, g^\Gamma_{nm} \, e^{-(E_n + \frac{\Delta_{mn}}{2}) t_{\text{sep}}} && (3.18) \\ &\quad \times \cosh\left[\Delta_{mn}\left(\tau - \frac{t_{\text{sep}}}{2}\right)\right] \; , && (3.19) \end{aligned}$$

where the matrix elements of interest are given by

$$g^\Gamma_{mn} = \langle m | j_\Gamma | n \rangle \, . \tag{3.20}$$

[1] We are using the non-relativistic normalization $\langle n | m \rangle = \delta_{nm}$ and $\mathbf{1} \equiv |\Omega\rangle\langle\Omega| + \sum_{n=0}^{\infty} |n\rangle\langle n|$. We assume that the contributions from the thermal terms that are exponentially suppressed from the full temporal extent are sufficiently suppressed that they can be ignored.

In this limit, the correlation functions are real, so $z_n^\dagger = z_n$ and $g_{nm}^\Gamma = g_{mn}^\Gamma$.

The axial and vector matrix elements can be extracted from the ratio

$$R_\Gamma(t_{\rm sep}, \tau) = \frac{C_\Gamma(t_{\rm sep}, \tau)}{C_2(t_{\rm sep})} \ , \tag{3.21}$$

where $\Gamma = A_3$ or V_4 denotes the axial or vector current, C_Γ is the three-point function with current insertion at time τ, and C_2 is the nucleon two-point function. In the limit of large $t_{\rm sep}$ and τ away from the source and sink, this ratio approaches the desired matrix element g_A or g_V. Physically, deviations from the asymptotic value arise from intermediate excited states, most prominently $N\pi$ scattering states and their interactions. A key observation of Ref. [100] is that transition matrix elements connecting the ground state to excited states dominate the contamination. These contributions decay parametrically as $e^{-\Delta t_{\rm sep}/2}$, whereas diagonal excited-state contributions scale as $e^{-\Delta t_{\rm sep}}$, as shown in the Eq. (10) of Ref. [100]. As a consequence, moderately increasing $t_{\rm sep}$ does not eliminate the leading systematic.

To further suppress excited states, the analysis also employs Feynman–Hellmann (FH) correlators constructed from summed insertion correlation functions. The FH correlator is defined as

$$\mathrm{FH}_\Gamma\left(t_{\rm sep}, \tau_c, dt\right) = \sum_{\tau=\tau_c}^{t_{\rm sep}-\tau_c} \frac{R_\Gamma(t_{\rm sep}+dt,\tau) - R_\Gamma(t_{\rm sep},\tau)}{dt} = \frac{1}{dt}\left[\frac{\Sigma_\Gamma(t_{\rm sep}+dt,\tau_c)}{C_2(t_{\rm sep}+dt)} - \frac{\Sigma_\Gamma(t_{\rm sep},\tau_c)}{C_2(t_{\rm sep})}\right] \ , \tag{3.22}$$

$$\mathrm{FH}_\Gamma\left(t_{\rm sep}, \tau_c\right) \equiv \mathrm{FH}_\Gamma\left(t_{\rm sep}, \tau_c, dt = 1\right) \ . \tag{3.23}$$

In this formulation, the dominant transition contributions are suppressed by the full energy gap, as shown in the Eq. (12) of Ref. [100], resulting in significantly improved ground-state isolation at shorter separations. Detailed decompositions of excited-state contamination in ratio and FH correlators are presented in Fig. 3.2. In the left panel, we see that the scattering excited state contribution (the upper

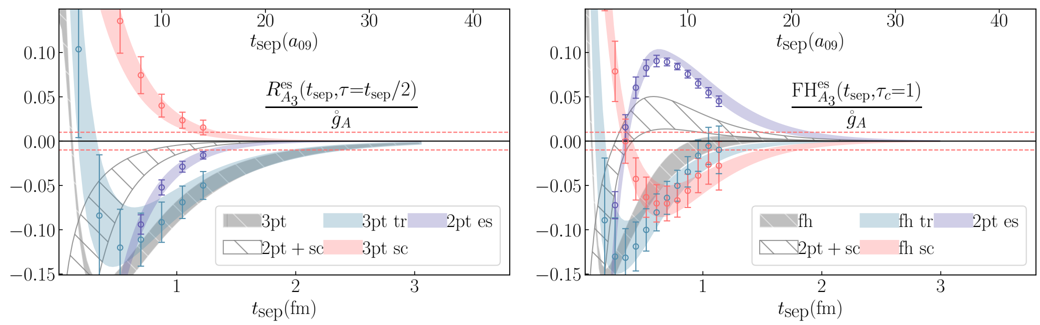

Figure 3.2: Ratio of excited state contributions to the ground state, g_A, for $R_{A_3}(t_{\rm sep},\tau)$ (left, 3pt) and $\mathrm{FH}_{A_3}(t_{\rm sep},\tau_c=1)$ (right, fh). Each panel contains three types of data points: the n-to-n scattering (sc) excited states, $C^{\rm sc}_\Gamma/(g_A C_2)$, the n-to-m transition (tr) excited states, $C^{\rm tr}_\Gamma/(g_A C_2)$, and the excited states from the two-point (2pt es) function, $-C^{\rm es}_2/C_2$ (and the corresponding Feynman-Hellmann versions of these ratios in the right figure). The decomposition of the numerical data into these three classes of excited state contamination depends upon our posterior reconstruction of the correlation functions (2pt, 3pt and FH). Because we are looking down the middle of R_{A_3} (along $\tau = t_{\rm sep}/2$), we only show data at even $t_{\rm sep}$s in the left panel. In addition to plotting the corresponding posterior contribution from each class of excited state, we plot two hatched bands. The white hatched band is the sum of the 3pt sc and 2pt es contributions, which are observed to largely cancel. The gray hatched band is the sum of all excited state contributions (the full posterior distribution, minus the ground state, normalized by the ground state). The 3pt sc contribution is positive for $R_{A_3}(t_{\rm sep},\tau)$ and it becomes negative for $\mathrm{FH}_{A_3}(t_{\rm sep},\tau_c=1)$ at $t_{\rm sep} > 0.5$ fm. The opposite is true for the 2pt es contribution. The sum of the scattering and two-point excited state contributions is opposite in sign to the transition excited state contributions, unlike for $R_{A_3}(t_{\rm sep},\tau)$ in which they are the same sign, leading to an even stronger suppression of the total excited state contributions beyond the expected suppression, as described in the text. The horizontal red dashed lines represent the threshold of a 1% contribution of excited state contamination to the ground state value.

(red) band/data) and the contribution from the excited states coming from the two-point function (the middle (purple) band/data) are roughly equal and opposite in sign. The hatched curve, which is mostly white with gray hash lines, is the sum of these two contributions, and it lies between them. We observe that the sum of these two sources of excited state contributions decay to a 1% correction (the horizontal dashed (red) lines) at $t_{\rm sep} \approx 1$ fm. In contrast, the transition (3pt tr) excited states depicted by the lowest (blue) band/data do not decay to the 1% level until $t_{\rm sep} \approx 2.2$ fm. In the right panel, we show the same excited state breakdown for the Feynman-Hellmann correlation function. The sign of the scattering and two-point excited state contributions both change compared to $R_{A_3}(t_{\rm sep}, \tau = t_{\rm sep}/2)$. It is also interesting to note that the magnitude of the transition excited state contributions at $t_{\rm sep} \approx 1$ fm go from $\approx 7.5\%$ for $R_{A_3}(t_{\rm sep}, \tau = t_{\rm sep}/2)$ to $\approx 2.5\%$ for $FH_{A_3}(t_{\rm sep}, \tau_c = 1)$. Besides, for $R_{A_3}(t_{\rm sep}, \tau = t_{\rm sep}/2)$, the sum of the scattering and two-point excited state contributions is the same sign as the sum of transition excited states, while for $\mathrm{FH}_{A_3}(t_{\rm sep}, \tau_c = 1)$ it is the opposite sign at intermediate and large $t_{\rm sep}$. Thus, the total excited state contamination decays to the 1% level for the FH correlation function at $t_{\rm sep} \approx 1$ fm while this does not happen until $t_{\rm sep} > 2$ fm for $R_{A_3}(t_{\rm sep}, \tau = t_{\rm sep}/2)$, which shows the significant improvement of the FH construction in suppressing excited states.

The extraction of g_A is performed through a fully correlated joint analysis of the two-point function, the axial and vector three-point ratios, and the corresponding FH correlators. Rather than determining the excited-state spectrum from the two-point function alone, all correlation functions are fitted simultaneously. This approach is essential because local three-quark interpolating operators do not uniquely isolate individual excited states; fitting correlators independently can therefore introduce biases in the ground-state matrix element. The joint fit results as well as the data points are presented in Fig. 3.3. In the left panel (the "Psychedelic Moose"), we plot the numerical results for the ratio of the three-point function generated with the $A_3 = \bar{q}\gamma_3\gamma_5\tau_3 q$ current, divided by the two-point function at the given value of $t_{\rm sep}$. The fit quality is good and visually one can see that the model accurately describes the numerical results over the full range of $t_{\rm sep}$ and τ used in the analysis. The horizontal (gray) band is the value of the ground state matrix element $\mathring{g}_A$. In the right panel we plot the $\mathrm{FH}_{A_3}(t_{\rm sep}, \tau_c = 1)$ data that are used in the joint fit as well as the resulting posterior distribution of this correlation function.

An important aspect of the analysis is the demonstration of stability under systematic variations. The extracted value of g_A remains consistent when varying the number of excited states included in

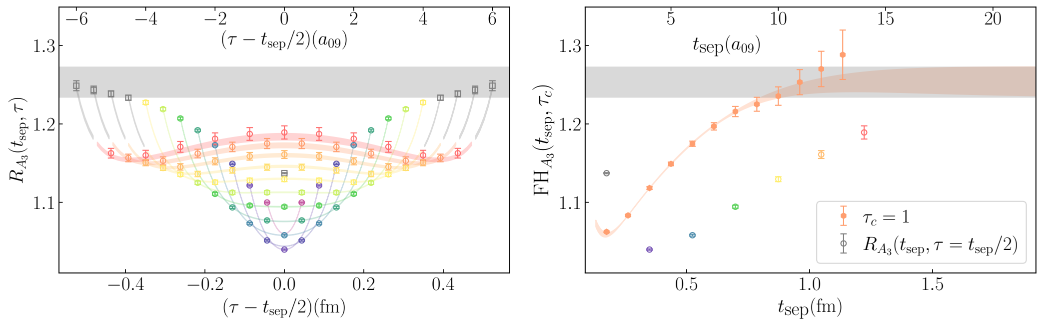

Figure 3.3: Left: (Psychedelic Moose/Water Buffalo plot) We plot the numerical results of $R_{A_3}(t_{\rm sep}, \tau)$ for $t_{\rm sep} = 2$ (the single gray point in the middle) to $t_{\rm sep} = 14$, the top (red) dataset. In addition, we plot the resulting posterior description of the correlation function from our 5-state fit as the (correspondingly colored) fit bands. The (gray) squares in the upper left/right of the "moose antlers" are not included in the analysis, as indicated by the break in the fit band from the inner region. The horizontal (gray) band is the ground state matrix element, $\mathring{g}_A$. Right: The $\mathrm{FH}_{A_3}(t_{\rm sep}, \tau_c = 1)$ numerical data (filled orange symbols) is plotted along with the posterior description of the correlation function from the joint analysis. As a comparison, for even values of $t_{\rm sep}$, we plot $R_{A_3}(t_{\rm sep}, \tau = t_{\rm sep}/2)$ with open (colored) symbols (the numcerical data in the middle of the "moose"). The time-axis is converted from lattice units (top) to fm (bottom) using our scale-setting [101].

the fit, modifying the prior model for energy gaps, truncating early- or late-time data, or restricting the analysis to subsets of source–sink separations. This stability indicates that the excited-state uncertainty is quantitatively captured within the fit.

Beyond the determination of g_A on this ensemble, this study establishes a methodological framework for nucleon structure calculations. It shows that reliable control of excited states requires a sufficiently large lever arm in $t_{\rm sep}$ combined with joint multi-correlator fits, and that FH correlators offer a practical avenue for improved ground-state isolation.

Chapter 4: Systematics Controlled Calculations of Partonic Physics from Lattice QCD

4.1 Large Momentum Effective Theory

If we think of the operator definition of PDFs, it is a light-cone correlation function with light-like separation. However, any real-time evolution is not directly accessible in lattice QCD due to the sign problem. More specifically, if we try to do the Wick rotation on a light-like separation, we will end up with a null separation in Euclidean space, which is not sufficient to define a nontrivial correlation function. Therefore, the early stage applications of lattice QCD to partonic physics were limited to the calculations of low moments of PDFs [102], which can be expressed as local operators. However, the calculation of higher moments is hindered by the power-divergent mixing with lower-dimensional operators, making it impractical to access the full x-dependence of PDFs.

A few years ago, a general approach to calculate x-dependent parton distributions based on Feynman's original idea about partons: They are the infinite-momentum limit of static properties of the proton at large momentum, and therefore are intrinsically Euclidean quantities accessible through lattice QCD [6]. This approach is known as LaMET [68–71], since a rigorous connection between the infinite-momentum frame partons and quarks and gluons at a finite momentum requires the matching of the UV modes with large momentum in effective field theory and systematic power counting. The key idea of LaMET is to construct a class of Euclidean correlation functions, called quasi-distributions, which are defined with spatially separated operators and can be computed on the lattice. These quasi-distributions depend on the hadron momentum P^z and approach the light-cone distributions in the infinite-momentum limit. By performing a perturbative matching calculation, one can relate the quasi-distributions to the physical PDFs, allowing for a systematic extraction of partonic physics from lattice

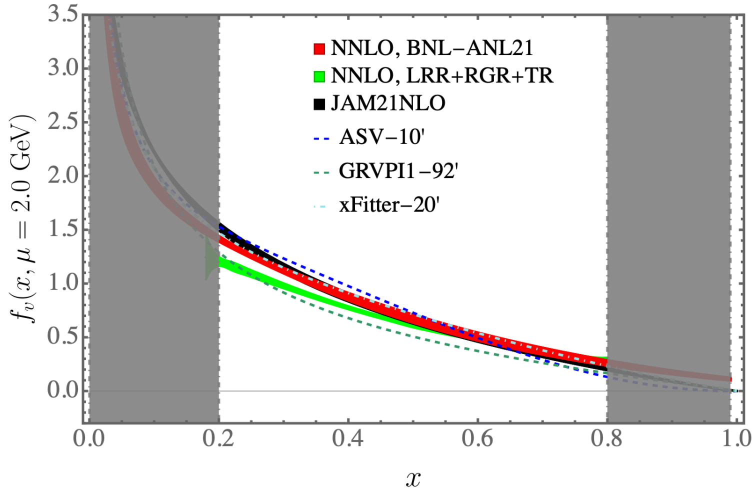

Figure 4.1: Illustration of a pion valence quark PDF extracted within the LaMET framework via calculations on a lattice with spacing $a = 0.04$ fm and pion mass 300 MeV at momentum $P^z = 1.94$ GeV using NNLO matching with LRR, RGR, and TR.

QCD calculations at finite momentum. So far, LaMET has been widely applied to calculate various partonic observables, including PDFs, GPDs, TMDs, as well as the light-cone wave functions [70, 103–262].

In this chapter, we will briefly review the recent developments within LaMET, with a focus on the novel Coulomb gauge formulation of quasi-distributions within LaMET framework. We will also present some example calculations of pion and nucleon partonic distributions using LaMET, including both collinear and TMD structures.

4.1.1 Recent Developments in LaMET

Since its original proposal, the LaMET has undergone substantial developments. These developments have significantly improved the precision from theoretical side. Firstly, several nonperturbative renormalization strategies were proposed, including regularization-independent momentum-subtraction (RI/MOM) schemes [117, 120, 121, 126] and ratio-based schemes [166, 255, 263, 264], as well as the hybrid renormalization framework [169], which leads to the later self-renormalization approach [189].

In parallel, advances in perturbative matching have played a central role. The introduction of leading renormalon resummation (LRR) [219, 265] clarified ambiguities in the matching kernel and eliminated spurious linear power corrections in the LaMET expansion. Higher-order matching calculations [164, 251, 264, 266] and the implementation of renormalization-group and threshold resummations (RGR and TR) [220, 237, 258] further enabled systematic resummation of large logarithms arising in the limits $x \to 0$ and $x \to 1$. As a result, the perturbative matching kernel is now under quantitative control across a wide kinematic range.

With these theoretical ingredients in place, the remaining power corrections suppressed by inverse powers of the hadron momentum can be estimated through controlled infinite-momentum extrapolation. Consequently, modern LaMET calculations achieve well-defined theoretical uncertainties for $x \in [x_{\min}, x_{\max}]$, rendering the framework suitable for precision phenomenology. An illustrative example is shown in Fig. 4.1, where the pion valence quark PDF is calculated on a lattice with spacing $a = 0.04$ fm and pion mass 300 MeV at momentum $P^z = 1.94$ GeV using next-to-next-to-leading order (NNLO) matching supplemented by LRR, RGR, and TR. The lattice prediction exhibits quantitative agreement with global analyses in the intermediate region $0.4 < x < 0.6$, despite relying on entirely independent theoretical and methodological inputs.

4.1.2 Quasi Distributions in Coulomb Gauge

It should be noted that the operator definition of quasi-distributions in LaMET framework is not unique. Instead, all operators projecting out the same light-front physics form a universality class [70, 267]. While can be proven that all operators in the same universality class will yield the same light-front physics in the infinite-momentum limit, they can have different UV structures at finite momentum, and therefore different renormalization and power corrections. In particular, the original formulation of LaMET is based on gauge-invariant nonlocal operators with Wilson lines, which are known to have complicated renormalization properties due to the presence of linear divergences associated with the Wilson lines. A novel formulation of quasi-distributions in Coulomb gauge has been proposed recently [217, 224], which avoids the use of Wilson lines and therefore has a much simpler renormalization structure.

In this formulation, the quasi-distributions are defined as nonlocal quark or gluon correlators in Coulomb gauge without Wilson lines, which can be expressed gauge-invariantly as the correlators of Coulomb-gauge-dressed fields. Thanks to the absence of Wilson lines, the renormalization of these operators is multiplicative and does not involve linear divergences, making it more straightforward to implement on the lattice. The perturbative matching between the Coulomb gauge quasi-distributions and the light-cone distributions has been derived using Soft Collinear Effective Theory (SCET) [268–271] and verified at one-loop order.

4.1.2.1 Gribov Copies

After the proposal of the Coulomb gauge formulation of quasi-distributions, there are concerns on whether the CG fixing process introduces extra systematic uncertainties that cannot be controlled. The main concern is about the effects from Gribov copies [272, 273], which are multiple gauge configurations that satisfy the same gauge-fixing condition.

To understand the origin of these ambiguities more precisely, it is useful to revisit the general framework of gauge fixing in non-Abelian gauge theories. In the path-integral quantization of Yang–Mills theory, the naive functional integration over gauge fields is ill-defined because gauge-equivalent configurations are overcounted. The Faddeev–Popov procedure resolves this redundancy by inserting a gauge-fixing condition together with the associated determinant, thereby isolating one representative from each gauge orbit. Implicit in this construction is the assumption that the gauge condition intersects each gauge orbit uniquely and that the corresponding Faddeev–Popov determinant remains positive. In the context of perturbation theory, one expands around small gauge fields, then the Faddeev–Popov operator is dominated by the Laplacian term (take Landau gauge as an example), which has no zero eigenvalues. Consequently, no nontrivial zero modes arise, and the Gribov copies are irrelevant. However, this assumption is not automatically guaranteed beyond perturbation theory. In particular, the existence of zero modes of the Faddeev–Popov operator is connected to the presence of multiple solutions to the gauge-fixing condition, i.e. Gribov copies [274].

On the lattice, gauge fixing is implemented numerically, typically by minimizing a gauge-fixing functional. In the case of Coulomb gauge, the condition is imposed independently on each time slice. The minimization procedure generally finds a local minimum of the gauge-fixing functional with a finite tolerance, rather than the absolute minimum. Different local minima correspond to different Gribov copies within the same gauge orbit. The numerical effects introduced by Gribov copies can be categorized into two primary types: lattice Gribov noise and measurement distortion [275]. Lattice Gribov noise arises from residual gauge freedoms and is related to the distribution of the target quantity across all Gribov copies. This noise is indistinguishable from the statistical uncertainties that arise from variations between lattice configurations [276, 277]. Consequently, the predominant concern regarding Gribov copies is the resultant measurement distortion. The distortion is related to the choice of the representative within various Gribov copies in gauge-fixing, i.e., the gauge-fixing strategy. A biased strategy can result in deviations from the true value, even though the Gribov noise remains indistinguishable from the statistical uncertainties. For example, selecting the copy with the maximal value of the target quantity as the representative for each configuration will clearly produce a distorted measurement. Therefore, it is necessary to quantify measurement distortion, for which a standard approach within the literature is to take various gauge-fixing strategies for comparative analysis. The resultant discrepancies among these strategies give an estimation of the measurement distortion.

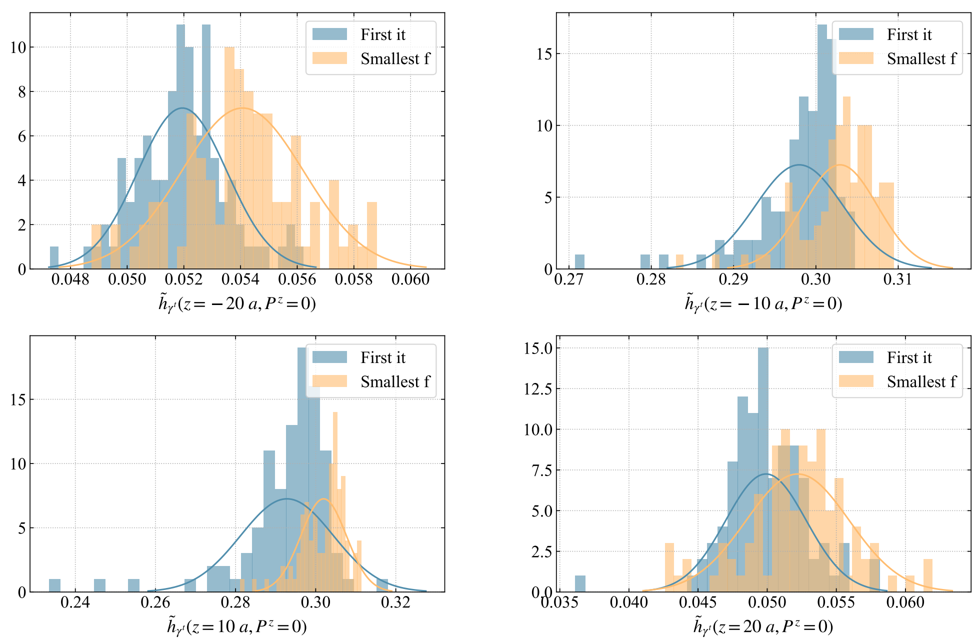

Figure 4.2: Histograms of the matrix element $\tilde{h}_{\gamma^t}(z, P^z = 0)$ from the bootstrap samples of the 100 configurations, measured on the two different sets of Gribov copies.

Our recent study [232] has investigated this issue by performing a systematic analysis of the Gribov copy effects on the lattice calculation of quark propagator and quasi-distribution matrix elements in Coulomb gauge. Considering the computational cost of the $3 + 1$ dimensional QCD simulation on the lattice, we study two strategies in this work:

1. **First instance** (“First it”): choosing the first copy obtained in the first Gribov region;

2. **Smallest functional** (“Smallest f”): choosing the copy with the smallest functional value among

all the obtained copies in the first Gribov region.

Using these two strategies, we should obtain two different sets of configurations after fixing the CG. The measurement distortion introduced by Gribov copies is estimated via the comparison between these two sets. As shown in Fig. 4.2, we select four representative separations $z = \{-20, -10, 10, 20\}a$ and plot histograms of bootstrap samples for the quasi-distribution matrix elements. No statistically significant difference is observed between the two strategies, indicating that any measurement distortion from Gribov copies is smaller than the statistical errors. The study finds that, although Gribov copies can introduce measurable distortions, these are negligible relative to the current statistical precision of the quasi-distribution matrix elements. It is consistent with the argument in Ref. [217] and supports the effectiveness of the CG method in the calculation of parton physics. If future calculations achieve substantially smaller statistical uncertainties, however, measurement distortions from Gribov copies will need to be examined more carefully.

4.1.3 General Analysis Procedure

This chapter presents several lattice calculations of partonic observables using the LaMET framework, including the pion DA, nucleon PDF, pion TMDs, and nucleon TMDs. The general analysis workflow common to these studies is summarized below.

4.1.3.1 Operator Definition

In lattice QCD calculations, hadron states are projected out by the interpolating operators. The choice of the interpolating operator is not unique, and different operators can have different overlaps with the ground state and excited states. Therefore, it is important to choose an operator that has a good overlap with the ground state and suppresses the excited state contamination.

A standard choice for pion interpolating operator is the pseudoscalar operator

$$\pi_s(\vec{x},t) = \overline{d}_s(\vec{x},t)\gamma_5 u_s(\vec{x},t) \ , \tag{4.1}$$

and for nucleon, a standard interpolating operator is

$$N^s_\alpha(\vec{x},t) = \epsilon_{abc}\, u^s_{a\alpha}(\vec{x},t) \left(u^s_b(\vec{x},t)^T\, C\gamma_5\, d^s_c(\vec{x},t) \right) \ , \tag{4.2}$$

where $C = \gamma^t\gamma^y$ is the charge conjugation operator, the subscript a, b, c are the color indices, α is the spin index, and the superscript s indicates whether the quark field is smeared ($s = S$) or not ($s = P$).

Although the above interpolating operators have good overlap with static pion and nucleon states, they are not necessarily optimal for highly boosted hadrons, which are important in LaMET framework. In the large-momentum limit, hadrons are more naturally described in terms of light-cone dynamics, where the dominant contributions arise from the "plus" components of the quark fields.

For a hadron boosted in the z-direction, one can decompose the Dirac field into light-cone components,

$$\psi = \psi_+ + \psi_- \ , \qquad \psi_\pm = \frac{1}{2}\gamma^\mp\gamma^\pm\psi \ , \tag{4.3}$$

where $\gamma^\pm = \frac{1}{\sqrt{2}}(\gamma^t \pm i\gamma^z)$. In the large P_z limit, the ψ_+ component dominates the leading Fock state of the hadron, while ψ_- is suppressed by powers of $M/(E + P_z)$.

As a result, interpolating operators constructed from the plus components have enhanced overlap with the ground state of a highly boosted hadron. In a recent work [278], new interpolating operators were proposed and referred to as *kinematically-enhanced* interpolating operators. The central idea is to construct operators aligned with the boost direction so that their overlap with the hadron ground state acquires a kinematic enhancement at large momentum.

For the pion, instead of the standard pseudoscalar operator $\bar{d}\gamma_5 u$, one may consider the axial-vector operator $\bar{d}\gamma^\mu\gamma_5 u$ with μ chosen parallel to the boost direction. In the large-P_z limit, this operator couples predominantly to the leading light-cone Fock component built from the plus spinor fields. Its overlap with the boosted ground state is enhanced by a factor proportional to P_μ, while the variance of the correlator does not receive a corresponding kinematic enhancement. Consequently, the signal-to-noise ratio is parametrically improved at large momentum.

A similar construction applies to the nucleon case, where interpolators involving $\gamma^\mu\gamma_5$ or γ^μ structures aligned with the boost direction lead to analogous kinematic enhancement for highly boosted states.

4.1.3.2 Asymptotic Extrapolation

One step in LaMET analysis is to perform a Fourier transform of the quasi-distribution matrix elements from coordinate space to momentum space, which formally requires an infinite range of spatial separations z. However, the lattice data for the quasi-distribution matrix elements are only available at a finite range of spatial separations z. Theoretically, lattice data should decay exponentially at large separations, implying that contributions from the long-distance tail to FT should be negligible. While in practice, while the values are statistically consistent with zero, the statistical uncertainties increase (or persist at a constant level for CG formalism) at large separations, which leads to non-physical fluctuations of the uncertainty band in a discrete Fourier transform. Therefore, in many cases, an asymptotic extrapolation is performed to extend the lattice data to large separations.

Many efforts have been made to control the systematic uncertainties associated with the Fourier transform in the literature, such as Refs. [169, 188, 279, 280]. In some earlier projects, a commonly used extrapolation form is

$$\tilde{h}(\lambda = zP^z) \quad = \left[\frac{c_1}{(-i\lambda)^{n_1}} + e^{i\lambda}\frac{c_2}{(i\lambda)^{n_2}}\right] e^{-\lambda/\lambda_0} \;, \tag{4.4}$$

where the algebraic terms in the square bracket account for a power law behavior $x^a(1-x)^b$ in the momentum space, and the exponential term comes from the expectation that at finite momentum the correlation function has a finite correlation length (denoted as λ_0), which becomes infinite when approaching light-cone. In the very recent work [279], more sophisticated derivations and results on the asymptotic behavior of the quasi-distribution matrix elements at large separations are presented. For example, for nucleon quark quasi-PDFs in GI formalism, the asymptotic behavior is found to be

$$\tilde{h}(\lambda = zP^z) = \left[A_2 \cdot e^{i\mathrm{sign}(\lambda)\phi_2} + \frac{A_2'}{|\lambda|} \cdot e^{i\mathrm{sign}(\lambda)\phi_2'}\right] e^{-m|\lambda|} \;. \tag{4.5}$$

4.1.3.3 Matching

According to LaMET [70, 71], when $P^z \gg \Lambda_{\mathrm{QCD}}$ the PDF can be obtained from the quasi-PDF through a perturbative matching [136, 217],

$$f_\Gamma^{\overline{\mathrm{MS}}}(x,\mu) \quad = \int_{-\infty}^{+\infty} \frac{\mathrm{d}y}{|y|}\, C_{\tilde{\Gamma}}^R\left(\frac{x}{y}, \frac{\mu}{|y|P^z}, z_s P^z\right) \tilde{f}_{\tilde{\Gamma}}^R(y, P^z, z_s) + \mathcal{O}\left(\frac{\Lambda_{\mathrm{QCD}}^2}{x^2P_z^2}, \frac{\Lambda_{\mathrm{QCD}}^2}{(1-x)^2P_z^2}\right) \;, \tag{4.6}$$

where Γ is the gamma matrix depending on the polarization, $C_{\tilde{\Gamma}}^R$ is the matching coefficient corresponding to the quasi-PDF $\tilde{f}_{\tilde{\Gamma}}^R(y, P^z, z_s)$ renormalized in a renormalization scheme labelled R, and the power corrections are suppressed by the active and spectator parton momenta. After resumming the linear renormalon [265], the power corrections in Eq. (4.6) start from the quadratic order.

In GI case, the matching coefficient in $\overline{\mathrm{MS}}$ scheme is derived at NNLO in the literature [164,264].

In CG case, the $\overline{\mathrm{MS}}$ matching coefficient is obtained as a series in the strong coupling α_s,

$$C^{\overline{\mathrm{MS}}}_{\tilde{\Gamma}}(\xi, \frac{\mu}{p^z}) = \delta(\xi - 1) + \frac{\alpha_s C_F}{2\pi} C^{\overline{\mathrm{MS}}}_{\tilde{\Gamma}(1)}(\xi, \frac{\mu}{p^z}) + \mathcal{O}(\alpha_s^2)\,, \tag{4.7}$$

where $C_F = 4/3$. At NLO the matching kernels are given by:

$$\begin{aligned} C^{\overline{\mathrm{MS}}}_{\gamma^t(1)}(\xi, \frac{\mu}{p^z}) &= C^{\overline{\mathrm{MS}}}_{\gamma^t\gamma_5(1)}(\xi, \frac{\mu}{p^z}) \\ &= C_r^{(1)}(\xi, \frac{\mu}{p^z}) + \frac{1}{2|1-\xi|} + \frac{1}{2}\delta(1-\xi)\left[1 + \ln\frac{4p_z^2}{\mu^2} - \int_0^2 d\xi' \frac{1}{|1-\xi'|}\right]\,, \end{aligned} \tag{4.8}$$

$$C^{\overline{\mathrm{MS}}}_{\gamma^z(1)}(\xi, \frac{\mu}{p^z}) = C^{\overline{\mathrm{MS}}}_{\gamma^z\gamma_5(1)}(\xi, \frac{\mu}{p^z}) = C^{\overline{\mathrm{MS}}}_{\gamma^t(1)}(\xi, \frac{\mu}{p^z}) + 2(1-\xi)^{[0,1]}_{+(1)} + \delta(1-\xi)\,, \tag{4.9}$$

where

$$\begin{aligned} C_r^{(1)}(\xi, \frac{\mu}{p^z}) &= \left[\tfrac{1+\xi^2}{1-\xi}\ln\tfrac{4p_z^2}{\mu^2} + \xi - 1\right]^{[0,1]}_{+(1)} + \Bigg\{\tfrac{1+\xi^2}{1-\xi}\left[\mathrm{sgn}(\xi)\ln|\xi| + \mathrm{sgn}(1-\xi)\ln|1-\xi|\right] \\ &\qquad + \mathrm{sgn}(\xi) + \tfrac{3\xi-1}{\xi-1}\tfrac{\tan^{-1}\left(\sqrt{1-2\xi}/|\xi|\right)}{\sqrt{1-2\xi}} - \tfrac{3}{2|1-\xi|}\Bigg\}^{(-\infty,\infty)}_{+(1)} \end{aligned} \tag{4.10}$$

satisfies particle number conservation. Here the plus functions on a domain D are defined as

$$[g(x)]^D_{+(x_0)} = g(x) - \delta(x - x_0)\int_D dx'\, g(x')\,. \tag{4.11}$$

Note that $C_r^{(1)}$ is analytical at $\xi = 1/2$ despite its form.

For the transversity PDF case,

$$C^{\overline{\rm MS}}_{\gamma^t\gamma^\alpha_\perp\gamma_5(1)}(\xi,\frac{\mu}{p^z}) \;\; = C^{\overline{\rm MS}}_{\gamma^z\gamma^\alpha_\perp\gamma_5(1)}(\xi,\tfrac{\mu}{p^z}) = C^{\perp(1)}_r(\xi,\tfrac{\mu}{p^z})\,, \tag{4.12}$$

where

$$\begin{aligned} C^{\perp(1)}_r(\xi,\frac{\mu}{p^z}) \;\; &= \left[\tfrac{2\xi}{1-\xi}\ln\tfrac{4p_z^2}{\mu^2}\right]^{[0,1]}_{+(1)} + \left\{\tfrac{2\xi}{1-\xi}\Big[\mathrm{sgn}(\xi)\ln|\xi| + \mathrm{sgn}(1-\xi)\ln|1-\xi|\Big]\right. \\ &\qquad \left. + \tfrac{3\xi-1}{\xi-1}\tfrac{\tan^{-1}\left(\sqrt{1-2\xi}/|\xi|\right)}{\sqrt{1-2\xi}} - \tfrac{1}{|1-\xi|}\right\}^{(-\infty,\infty)}_{+(1)}\,. \end{aligned} \tag{4.13}$$

In the TMD case, it has been shown that quasi TMDPDFs have the same collinear degrees of freedoms as TMDPDFs [281]. Their differences from soft modes can be attributed to the intrinsic soft function and different rapidity scales. Also, contributions from highly off-shell modes are local [282]. Thus, TMDPDFs $f\left(x,b_\perp;\mu,\zeta\right)$ are connected to quasi TMDPDFs $\tilde{f}_\Gamma\left(x,b_\perp,P^z;\mu\right)$ via a multiplicative factorization [199, 281]:

$$\begin{aligned} \sqrt{S_I\left(b_\perp;\mu\right)}\cdot\tilde{f}_\Gamma\left(x,b_\perp,P^z;\mu\right) \;\; &= f\left(x,b_\perp;\mu,\zeta\right)H_f\left(x,P^z;\mu\right)\exp\left[\tfrac{1}{2}\ln\tfrac{(2xP^z)^2}{\zeta}\gamma^{\overline{\rm MS}}\left(b_\perp;\mu\right)\right] \\ &\quad +\mathcal{O}\left(\tfrac{\Lambda_{\rm QCD}}{xP^z},\tfrac{1}{b_\perp(xP^z)}\right)\,, \end{aligned} \tag{4.14}$$

where $S_I\left(b_\perp;\mu\right)$ denotes the intrinsic soft function that has been calculated on the lattice in Refs. [171, 190, 221, 222, 246] and $\gamma^{\overline{\rm MS}}\left(b_\perp;\mu\right)$ denotes the Collins-Soper kernel.

In GI case, the hard kernel H_f as a function of $\zeta_z/\mu^2 = \left(2xP^z\right)^2/\mu^2$, has been perturbatively determined up to next-to-leading order (NLO) [70, 137, 281, 282]

$$H^{(1)}\left(\frac{\zeta_z}{\mu^2}\right) = \frac{\alpha_s C_F}{2\pi}\left(-2+\frac{\pi^2}{12}+\ln\frac{\zeta_z}{\mu^2}-\frac{1}{2}\ln^2\frac{\zeta_z}{\mu^2}\right), \tag{4.15}$$

as well as next-to-next-to-leading order (NNLO), calculated recently [220, 283]

$$H^{(2)}\left(\frac{\zeta_z}{\mu^2}\right) = \alpha_s^2\left[c_2-\frac{1}{2}\left(\gamma_C^{(2)}-\beta_0 c_1\right)\ln\frac{\zeta_z}{\mu^2}-\frac{1}{4}\left(\Gamma_{\text{cusp}}^{(2)}-\frac{\beta_0 C_F}{2\pi}\right)\ln^2\frac{\zeta_z}{\mu^2}-\frac{\beta_0 C_F}{24\pi}\ln^3\frac{\zeta_z}{\mu^2}\right], \tag{4.16}$$

where $\zeta_z = (2xP^z)^2$ and $c_2 = 0.0725C_F^2 - 0.0840C_F C_A + 0.1453C_F n_f/2$.

In CG case, the hard kernel $H_f(x, P^z; \mu) = |C_{\text{TMD}}(xP^z; \mu)|^2$, where the coefficient is calculated up to NLO as [224]

$$C_{\text{TMD}}(xP^z, \mu) = 1+\frac{\alpha_s C_F}{4\pi}\left[-\frac{L_p^2}{2}-3L_p-12+\frac{7\pi^2}{12}\right], \tag{4.17}$$

in which $L_p = \ln\frac{\mu^2}{(2xP^z)^2}$ and $C_F = 4/3$.

4.2 Calculation of Pion DAs

As the pseudo Nambu-Goldstone boson of the spontaneous chiral symmetry breaking in QCD, the pion plays a special role in hadron physics. The pion distribution amplitude (DA) is a fundamental nonperturbative function that describes the momentum fraction distribution of the quark and antiquark inside the pion. It is an essential input for understanding exclusive pion processes, including those that probe the pion electromagnetic form factor and the pion-photon transition form factor [19]. The calculation of the pion DA from first principles using lattice QCD has been a long-standing challenge

Ensemble	a(fm)	$L^3\times\,T$	$c_{\rm SW}$	$m_{u/d}$	m_s
a06m130	0.057	$96^3\times192$	1.03493	-0.0439	-0.0191
a09m130	0.088	$64^3\times$ 96	1.04239	-0.0580	-0.0174
a12m130	0.121	$48^3\times$ 64	1.05088	-0.0785	-0.0191

Table 4.1: Details of the simulation setup. The light and strange quark mass (both valence and sea quark) of the clover action are tuned such that m_π=140 MeV and m_{η_s}=670 MeV.

due to the need to access light-cone correlations. However, with the development of LaMET, it has become possible to calculate the pion DA on the lattice by constructing appropriate quasi-distributions and performing the perturbative matching to extract the physical DA.

The calculations in this project [203] are done using the clover valence fermion on three ensembles with 2+1+1 flavors of HISQ sea quarks generated by the MILC Collaboration [284, 285]. The valence quark masses are tuned to $m_\pi = 140$ MeV and $m_{\eta_s} = 670$ MeV, and the three lattice spacings are 0.057, 0.088 and 0.121 fm. The detailed lattice setup is summarized in Table. 4.1. We use three different pion momenta, $P_z = \{1.29, 1.72, 2.15\}$ GeV, the corresponding Lorentz boost factors γ for the pion at the physical mass are $\gamma = \{9.21, 12.29, 15.36\}$, to study the convergence of the LaMET expansion and to perform the extrapolation to the infinite-momentum limit.

The leading-twist light-cone DA of a pseudoscalar meson is defined as Eq. (2.10). To extract this quantity, we calculate the following quasi correlation on the lattice with momentum $\vec{P}$ along the z-direction [128]:

$$C_2^m(z,\vec{P},t) = \int d^3y e^{-i\vec{P}\cdot\vec{y}} \times \langle 0|\mathcal{O}_{\Gamma_1}(z;\vec{y},t)\bar{\psi}_2(0,0)\Gamma_2\psi_1(0,0)|0\rangle\ , \tag{4.18}$$

where $\mathcal{O}_{\Gamma_1}(z;\vec{y},t) \equiv \bar{\psi}_1(\vec{y},t)\Gamma_1 U(\vec{y},\vec{y}-z\hat{z})\psi_2(\vec{y}-z\hat{z},t)$ is the quasi operator with $\hat{z}$ being the unit vector in the z-direction, $U(\vec{x},\vec{x}-\vec{z})$ is the spatial Wilson line connecting lattice sites $\vec{x}$ and $\vec{x}-\vec{z}$, $\psi_2\Gamma_2\psi_1$ is the interpolating field of the meson m, and $\Gamma_{1,2}$ are chosen as $\Gamma_1 = \gamma_z\gamma_5$, $\Gamma_2 = \gamma_5$ for the pseudoscalar meson. The ground-state matrix elements can be extracted from the following two-state fit formula:

$$\frac{C_2^m(z,\vec{P},t)}{C_2^m(z=0,\vec{P},t)} = \frac{H_m^{\mathrm{B}}(z)(1+c_m(z)e^{-\Delta E t})}{(1+c_m(0)e^{-\Delta E t})}\ , \tag{4.19}$$

where $H_m^{\mathrm{B}}(z)$ is the normalized ground-state matrix element, c_m and ΔE are free parameters accounting for (one or more) excited state contamination, which are exponentially suppressed in the large time limit. Based on the comparison of one- and two-state fits, we use the one-state fit results in the analysis below with $t_{min} = 0.72, 0.54, 0.42$ fm (for $P_z = 1.29, 1.72, 2.15$ GeV) which is large enough to eliminate the excited-state contamination.

The bare quasi correlation calculated above contains both linear and logarithmic ultraviolet (UV) divergences which have to be removed by renormalization. On the lattice, the numerical subtraction of linear divergences is extremely delicate. In particular, such divergences may not be fully removed if the RI/MOM renormalization scheme is used [180]. Thus, we adopt the self renormalization proposed in Ref. [189], which amounts to fitting the bare quasi correlation and subtracting the relevant UV divergences. To be more precise, one fits the bare quasi correlation at given hadron momentum and multiple lattice spacings with a perturbative-QCD-dictated parametrization that contains a linear divergence, a logarithmic divergence, and discretization effects. After removing all the UV divergences and discretization effects, one is left with the renormalized quasi correlation encoding the intrinsic non-perturbative physics.

As suggested in Ref. [189], the UV divergences in the quasi correlator can be determined by using, e.g., the pion PDF matrix elements $\mathcal{M}(z) \equiv \langle\pi|\mathcal{O}_{\gamma_t}|\pi\rangle$ in the rest frame at multiple lattice spacings, and fitting the bare data $\mathcal{M}^{\mathrm{B}}$ to the following form [189]

$$\mathcal{M}^{\mathrm{B}}(z,a) = Z^{\mathrm{self}}(z,a)\mathcal{M}^{\mathrm{R}}(z), \tag{4.20}$$

with the renormalization factor parametrized as [189]

$$\begin{aligned} Z^{\mathrm{self}}(z,a) \equiv \quad & \exp\Big\{\frac{kz}{a\ln[a\Lambda_{\mathrm{QCD}}]} + m_0 z + f(z)a \\ & + \frac{3C_F}{b_0}\ln\left[\frac{\ln[1/(a\Lambda_{\mathrm{QCD}})]}{\ln[\mu/\Lambda_{\mathrm{QCD}}]}\right] + \ln\left[1+\frac{d}{\ln(a\Lambda_{\mathrm{QCD}})}\right]\Big\}\,, \end{aligned} \tag{4.21}$$

where the first term in the curly bracket is the linear divergence, m_0 denotes a finite mass contribution arising from renormalon ambiguity, etc., and $f(z)a$ accounts for the discretization effects (The $\mathcal{O}(a)$ correction here arises from the mixed action effect in using clover valence fermions on HISQ sea ones). The last two terms come from the resummation of leading and sub-leading logarithmic divergences, which only affect the overall normalization at different lattice spacings. To partially account for higher-order perturbative effects as well as remaining lattice artifacts, we also treat d and Λ_{QCD} as fitting parameters [189]. The renormalized matrix element is required to be equal to the continuum perturbative $\overline{\mathrm{MS}}$ result at short distances (chosen to be $z \in z_s = [0.06, 0.18]$ fm as defined in [169]),

$$\mathcal{M}^{\mathrm{R}}(z)|_{z\in z_s} = \mathcal{M}^{\overline{\mathrm{MS}},\,1\text{-loop}}(z) \equiv 1 + \frac{\alpha_s^{\overline{\mathrm{MS}}} C_F}{4\pi}\left[3\ln\frac{z^2\mu^2}{4e^{-2\gamma_E}} + 5\right] + \mathcal{O}(\alpha_s^{2,\overline{\mathrm{MS}}})\,, \tag{4.22}$$

which helps the determination of m_0 and d. In the calculation we use the $\overline{\mathrm{MS}}$ renormalization scale $\mu = 2$ GeV and $\Lambda_{\overline{\mathrm{MS}}} = 0.24$ GeV.

In the present case, we follow the same strategy as above, except that the renormalized matrix element in the $\overline{\mathrm{MS}}$ scheme,

$$H_m^{\overline{\mathrm{MS}}}(z) = H_m^{\mathrm{B}}(z,a)/\tilde{Z}^{\mathrm{self}}(z,a) \tag{4.23}$$

is now required to be matched to the continuum perturbative $\overline{\mathrm{MS}}$ result of the normalized quasi-DA matrix element at short distances in the rest frame, which reads at one-loop

$$H_m^{\overline{\mathrm{MS}},\ 1\text{−loop}}(z) \equiv 1 + \frac{\alpha_s^{\overline{\mathrm{MS}}} C_F}{4\pi}\left[3\ln\frac{z^2\mu^2}{4e^{-2\gamma_E}} + 7\right] . \tag{4.24}$$

$\tilde{Z}^{\mathrm{self}}$ turns out to be the same as Z^{self} except for the value of d.

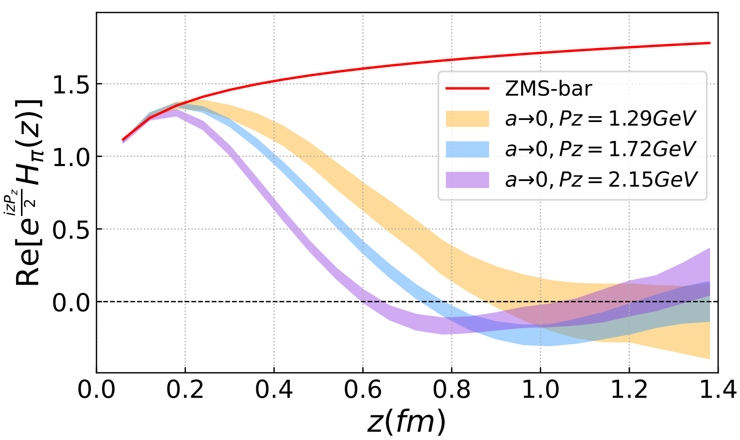

Figure 4.3: Comparison of self-renormalized quasi correlation $\mathrm{Re}[e^{\frac{izP_z}{2}} H_m^{\overline{\mathrm{MS}}}(z)]$ of the pion with different momenta (bands), and the perturbative result in the $\overline{\mathrm{MS}}$ scheme $H_m^{\overline{\mathrm{MS}},\ 1\text{−loop}}(z)$ (red curve).

In Fig. 4.3, we show a comparison between the self-renormalized quasi correlations $\mathrm{Re}[e^{\frac{izP_z}{2}} H_m^{\overline{\mathrm{MS}}}(z)]$ (after linear $\mathcal{O}(a)$ continuum extrapolation and phase rotation) with the perturbative one-loop result $H_m^{\overline{\mathrm{MS}},\ 1\text{−loop}}(z)$. As can be seen from the figure, all quasi correlations agree well with the perturbative result for small z, indicating a mild P_z dependence in that region. Besides, from the figure, we can also see that the uncertainty of the renormalized quasi correlation grows rapidly at large distance. As discussed before, an asymptotic extrapolation is needed to avoid non-physical fluctuations of the uncertainty band in a discrete Fourier transform. The extrapolation form in Eq. (4.4) is adopted in this work. Note that the Lorentz boost factor γ for the pion at the physical point is very large $\{9.21, 12.29, 15.36\}$, and thus the correlation length is very large. We therefore drop the $e^{-\lambda/\lambda_0}$ factor, and directly perform a polynomial extrapolation as suggested in [169].

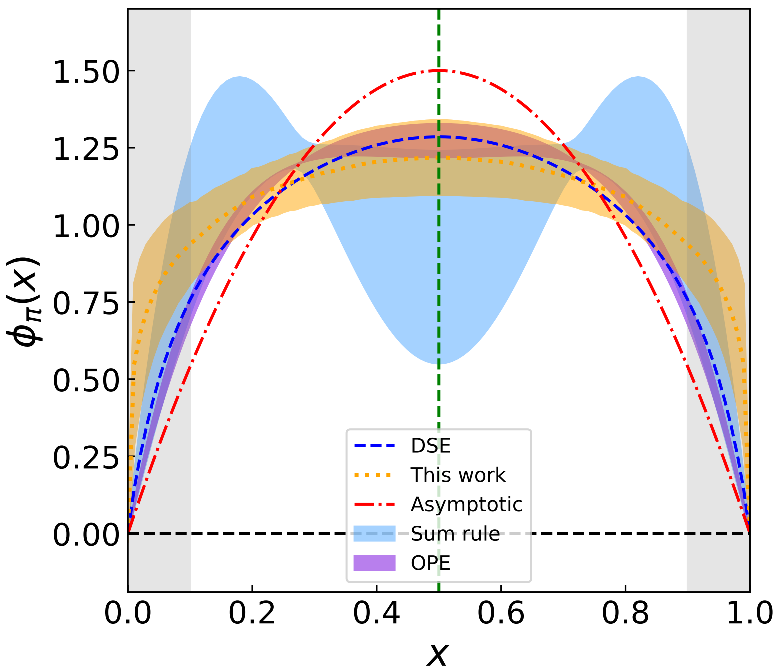

Figure 4.4: DAs of π meson, extrapolated to the continuum $(a \to 0)$ and infinite momentum limit $(P_z \to \infty)$. For the kaon, x is the momentum fraction carried by the light quark.

With the results for $P_z = 1.29, 1.72, 2.15$ GeV above, we can perform an extrapolation to $P_z \to \infty$ using the functional form:

$$\phi(x, P_z) = \phi(x, P_z \to \infty) + \frac{c_2(x)}{P_z^2} + \mathcal{O}\Big(\frac{1}{P_z^4}\Big). \tag{4.25}$$

The final results of the pion DAs are given in Fig. 4.4, where systematic uncertainties from renormalization scale μ dependence, large λ extrapolation, continuum and infinite momentum extrapolation have been taken into account. While the systematic uncertainty from continuum extrapolation dominates for the kaon, the statistical (larger than for the kaon due to the lighter quark mass) and continuum extrapolation uncertainties are comparable and dominate for the pion. Therefore, we expect that the kaon DA uncertainty can get significantly suppressed when data at smaller lattice spacings are available which enable us to do a more reliable continuum extrapolation, whereas for the pion, besides data at smaller lattice spacings, we also need increasing statistics to reduce the statistical uncertainty. In the endpoint region which cannot be reliably predicted by LaMET, we adopt a phenomenological $x^a(1-x)^b$ extrapolation (taken as $0 < x < 0.1$ & $0.9 < x < 1$). The unreliable region is expected to shrink if a larger P_z can be reached in future calculations. For comparison, we also plot the asymptotic form $6x(1-x)$ and results from QCD sum rules [286], Dyson-Schwinger equations (DSE) [287], and reconstructed from moments calculations (OPE) [288]. As can be seen from the figure, the π DAs deviate significantly from the asymptotic form, but are close to the results from DSE and OPE calculations. The shape of π DA is much broader than the asymptotic form, manifesting the impact of DCSB at such a low scale. A similar behavior has also been observed in a recent analysis using QCD sum rules with nonlocal condensates [289].

In summary, this is a state-of-the-art lattice calculation of the pion DA at the physical point. The renormalization is done in the hybrid scheme with self renormalization proposed recently. Based on the results at physical light and strange quark masses with three lattice spacings and momenta, we perform an extrapolation to the continuum and infinite momentum limit. The final results exhibit a significant deviation from the asymptotic form, while they are close to the DSE and OPE results, especially in the middle x region where our method is reliable. However, there are still some significant differences in the endpoint regions. This could be due to missing higher-power/ high-order corrections in LaMET which can be improved in future calculations, or due to effects of higher moments ignored in the OPE and DSE calculations. A more accurate determination of the endpoint behavior of the DAs would be important towards a better understanding of quantities like the pion-photon transition form factor.

4.3 Calculation of Nucleon PDFs

Compared with the pion DA, which encodes the longitudinal momentum sharing between a valence quark and antiquark in an exclusive meson channel at small transverse separation, the parton distribution function (PDF) characterizes the probability density of finding a quark or gluon carrying a given longitudinal momentum fraction inside a fast-moving hadron in inclusive processes. In other words, the DA is a light-cone wave-function–like object relevant for exclusive reactions and involves a projection onto Fock states, whereas the PDF is an inclusive density that integrates over transverse dynamics and sums over all allowed final states.

Moving from the pion to the nucleon further increases both theoretical and computational complexity. While the pion is a comparatively simple two-body bound state with a relatively clean spectral structure, the nucleon is a three-valence-quark system embedded in a dense spectrum of excited states with significant sea-quark and gluon contributions. In lattice QCD calculations, this richer excitation spectrum leads to substantially more severe excited-state contamination in nucleon correlation functions. The energy gaps above the ground-state nucleon are typically smaller, and the state multiplicity is higher, making it considerably more challenging to isolate the ground state within feasible Euclidean time separations.

In addition, nucleon correlators generally suffer from a poorer signal-to-noise ratio, which decays exponentially with Euclidean time in a way governed by the difference between the nucleon and multi-pion energies. This statistical degradation limits the temporal window available for reliable ground-state extraction and exacerbates systematic uncertainties. The presence of more intricate spin and flavor structures in baryonic operators, as well as enhanced operator mixing under renormalization, further complicates the analysis.

Despite these theoretical and numerical challenges, nucleon PDFs remain among the most fundamental nonperturbative quantities in QCD. Owing to their universality in QCD factorization, PDFs extracted from a subset of processes can be used consistently as nonperturbative inputs for cross-section predictions across a broad class of hard-scattering reactions. Over recent decades, high-energy scattering experiments at HERA, COMPASS, SLAC, Fermilab, and the LHC have delivered state-of-the-art constraints on PDFs, enabling global analyses of the collinear structure of both unpolarized and polarized nucleons, with percent-level uncertainties in well-constrained kinematic regions [31–33,290,291]. Looking ahead, the future Electron-Ion Collider [7,8] will probe PDFs and other partonic observables with unprecedented precision, opening new opportunities to map the multidimensional structure of hadrons.

In our recent work [103], we present the first comprehensive benchmark of the CG method for nucleon PDFs through a lattice QCD calculation of the isovector unpolarized, helicity, and transversity distributions. We perform a numerical lattice QCD calculation on a 2+1 flavor gauge ensemble generated by the HotQCD Collaboration [292] with Highly Improved Staggered Quarks (HISQ) [284]. The ensemble has a lattice spacing of $a = 0.0601(5)$ fm and a volume of $L_s^3 \times L_t = 48^3 \times 64$. We used tadpole-improved clover Wilson valence fermions on a hypercubic (HYP) smeared [293] gauge background. We set the clover coefficient $c_{\rm sw} = u_0^{-3/4}$, in which u_0 is the average plaquette after HYP smearing; the parameter is set as $c_{\rm sw} = 1.0336$ for both time and spatial directions. For the masses of the valence quarks, we used hopping parameter $\kappa = 0.12623$ for the light quarks u and d, corresponding to a valence pion mass $m_\pi = 300$ MeV. By boosting the nucleon along an off-axis direction $\vec{n} = (1, 1, 0)$, we are able to reach momenta of $P^z = 2.43$ and 3.04 GeV at a pion mass of $m_\pi = 300$ MeV, which are among the largest boosts achieved in nucleon PDF calculations at a comparable pion mass. Nevertheless, the CG method faces similar systematic uncertainties as the GI approach, including excited-state contamination and power-suppressed corrections. A detailed analysis is carried out to address excited-state contamination in the extraction of matrix elements, where we find that imaginary parts of matrix elements are significantly more sensitive to excited-state contamination, consistent with previous observations for GI operators [156]. Besides excited-state contamination, we also perform a systematic investigation of other sources of uncertainty in the CG framework, including the choice of Dirac gamma structure in quasi-PDF operators, the renormalization schemes, uncertainties associated with the Fourier transform, and power-suppressed corrections in the LaMET factorization. For the real parts of the quasi-PDF matrix elements, we utilize a hybrid scheme [169] to renormalize the matrix elements. To renormalize the imaginary part of the quasi-PDF matrix elements, we argue that the hybrid scheme is not optimal and instead propose a static version of the RI′-MOM scheme [294], where we find that the systematic uncertainties are under control. While the CG formulation does not eliminate the intrinsic challenges of nucleon PDF calculations, such as excited-state contamination due to the dense excited-state spectra, our results demonstrate reasonable agreement with existing lattice and phenomenological studies. This work thus validates the CG method as a promising approach for precision studies of the nucleon partonic structure and provides a solid foundation for its future extensions, including the computation of TMDs.

4.3.1 Theoretical Framework

The quasi PDFs in the CG are defined as [217]

$$
\begin{aligned}
\tilde{f}_{\tilde{\Gamma}}(x, P^z; \mu) &= P^z \int \frac{dz}{2\pi} e^{iz(xP^z)} \tilde{h}_{\tilde{\Gamma}}(z, P^z, \mu) \,, \qquad (4.26)\\
\tilde{h}_{\tilde{\Gamma}}(z, P^z; \mu) &\equiv \frac{1}{2N_{\tilde{\Gamma}}} \langle PS| \, \bar{\psi}(z)\tilde{\Gamma}\psi(0)\Big|_{\vec{\nabla}\cdot\vec{A}=0} |PS\rangle \,, \qquad (4.27)
\end{aligned}
$$

where the normalization factor $N_{\tilde{\Gamma}}$ is given by:

$$
N_{\tilde{\Gamma}} = \begin{cases} P^t : & \tilde{\Gamma} \in \{\gamma^t, \gamma^z\gamma_5, \gamma^t\gamma_\perp^\alpha\gamma_5\} \,, \\ P^z : & \tilde{\Gamma} \in \{\gamma^z, \gamma^t\gamma_5, \gamma^z\gamma_\perp^\alpha\gamma_5\} \,. \end{cases} \qquad (4.28)
$$

The Dirac matrices $\tilde{\Gamma} = \{\gamma^t/\gamma^z, \gamma^t\gamma_5/\gamma^z\gamma_5, \gamma^t\gamma_\perp^\alpha\gamma_5/\gamma^z\gamma_\perp^\alpha\gamma_5\}$ in the quasi-PDF corresponds to $\Gamma = \{\gamma^+, \gamma^+\gamma_5, \gamma^+\gamma_\perp^\alpha\gamma_5\}$ in the PDF.

In the LaMET framework, several nonperturbative renormalization schemes are used in practice to renormalize the gauge-invariant quasi-PDF matrix element, including ratio-type schemes [117, 120, 121, 126, 255, 263, 264], regularization independent momentum subtraction schemes [117, 120, 121, 126], and the hybrid scheme [169, 189]. The main advantage of the hybrid scheme is that discretization effects are cancelled by taking appropriate ratios at short separations, while also avoiding introducing any nonperturbative artifacts in the long-distance matrix elements.

As shown in Ref. [217, 244], due to the absence of Wilson line the CG correlator is free from the linear divergence and the associated renormalon [225]. As a result, the renormalization of the nonlocal

CG operator is purely multiplicative and can be defined as

$$\bar{\psi}_0(z)\tilde{\Gamma}\psi_0(0) = Z_\psi^R(a)\left[\bar{\psi}_R(z)\tilde{\Gamma}\psi_R(0)\right] , \tag{4.29}$$

where $Z_\psi^R(a)$ is the renormalization factor of the quark wave-function into some renormalization scheme R as a function of the lattice spacing a. ψ_0 and ψ_R represent the bare and renormalized quark fields, respectively.

In the standard definition of the hybrid scheme [188, 202, 236], the renormalized quasi-PDF matrix element is defined by the following ratios:

$$\tilde{h}_{\tilde{\Gamma}}^{\mathrm{hyb.}}(z, P^z; z_s) = N\frac{\tilde{h}_{\tilde{\Gamma}}^0(z,P^z;a)}{\tilde{h}_{\tilde{\Gamma}}^0(z,0;a)}\theta(z_s - |z|) + N\frac{\tilde{h}_{\tilde{\Gamma}}^0(z,P^z;a)}{\tilde{h}_{\tilde{\Gamma}}^0(z_s,0;a)}\theta(|z| - z_s) , \tag{4.30}$$

where $\tilde{h}_{\tilde{\Gamma}}^0/\tilde{h}_{\tilde{\Gamma}}^{\mathrm{hyb.}}$ denotes the bare/hybrid-renormalized matrix element respectively, the normalization factor is defined by $N = \tilde{h}_{\tilde{\Gamma}}^0(0,0;a)/\tilde{h}_{\tilde{\Gamma}}^0(0,P^z;a)$, and z_s ($a \ll z_s \ll \Lambda_{\mathrm{QCD}}^{-1}$) denotes the point that separates the short- and long-distance regimes. Since the matrix elements are usually strongly correlated at neighboring lattice sites, this normalization can reduce their statistical uncertainties at small z. However, a drawback of this treatment is that the discretization errors of $z = 0$ matrix elements are unnecessarily propagated to all $z \neq 0$, where such effects are expected to be suppressed as powers of $a/|z|$. An ideal solution would be a z-dependent normalization that accurately removes the discretization effects mostly at small z, which might be achieved with lattice perturbation theory. However, it is beyond the scope of this work. After all, the discretization effects can be systematically eliminated through continuum extrapolation, so this normalization may still be favored for the purpose of reducing statistical uncertainties.

On a single lattice ensemble, there is no clear preference for applying or omitting the above normalization, since the discretization effects cannot be quantified. However, since the $z = 0$ matrix elements are expected to be independent of P^z due to Lorentz covariance, a significant deviation of N from unity signals sizable lattice artifacts; in this case, a blind normalization would rescale the quasi-PDF by the same factor. Therefore, in this work we adopt a hybrid scheme without the normalization at $z = 0$, namely,

$$\tilde{h}^{\mathrm{hyb.}}_{\tilde{\Gamma}}(z, P^z; z_s) = \frac{\tilde{h}^0_{\tilde{\Gamma}}(z, P^z; a)}{\tilde{h}^0_{\tilde{\Gamma}}(z, 0; a)}\theta(z_s - |z|) + \frac{\tilde{h}^0_{\tilde{\Gamma}}(z, P^z; a)}{\tilde{h}^0_{\tilde{\Gamma}}(z_s, 0; a)}\theta(|z| - z_s)\ . \tag{4.31}$$

As a result, the discretization effects at $z \sim a$ remain uncorrected and propagate into momentum space as artifacts of order $(P^z a)^n$. Although this treatment spoils the normalization of the quasi-PDF and PDF, the effect vanishes in the continuum limit. The hybrid renormalization is applied to the real part of the quasi-PDF matrix elements with the separation point chosen as $z_s = 3\sqrt{2}a$, where the factor of $\sqrt{2}$ originates from the off-axis momentum setup.

For the imaginary part of the quasi-PDF matrix elements, however, it is likely that the ratios employed in the hybrid scheme do *not* remove the leading lattice artifacts for $z \leq z_s$, as the zero-momentum matrix elements in the denominator of the ratio are purely real. To address this issue, we instead employ a modification of the RI′-MOM [294] scheme for the quark wave function renormalization factor Z_R in Eq. 4.29. Because the Coulomb gauge is not a complete gauge fixing due to the residual gauge freedom in the timelike direction, such a scheme based on the quark propagator either requires introducing *additional* gauge-fixing conditions in the timelike direction [295, 296], or one must only use spatial correlation functions. Choosing the latter route leads to a Static RI′-MOM scheme (SRI′-MOM) which we define below. Given the CG static propagator:

$$S(\vec{k}) = \int \mathrm{d}^3\vec{x}\, e^{i\vec{x}\cdot\vec{k}} \left\langle \psi(\vec{0}, t_E)\overline{\psi}(\vec{x}, t_E) \right\rangle , \tag{4.32}$$

computed from position-space propagators at a single Euclidean time t_E, we can define a renormalization constant such that the renormalized static propagator is equal to the tree-level value at $\vec{k}$:

$$\psi_0 = Z^{1/2}_{\mathrm{SRI}'}(|\vec{k}|)\psi_{\mathrm{SRI}'}, \qquad Z_{\mathrm{SRI}'}(|\vec{k}|) = \lim_{m\to 0} \frac{\mathrm{Tr}[\vec{k}\!\!\!/ S_0(\vec{k})]}{\mathrm{Tr}[\vec{k}\!\!\!/ S_{\mathrm{tree}}(\vec{k})]} , \tag{4.33}$$

where the chiral limit is taken in order to define a mass-independent renormalization scheme to match to $\overline{\mathrm{MS}}$. Using leading-logarithmic running and $O(\alpha_S)$ matching between SRI′-MOM and $\overline{\mathrm{MS}}$, we find that the renormalization constant for the ensemble used in this work is:

$$Z_{\overline{\mathrm{MS}}}(a, \mu = 2\mathrm{GeV}) = 0.915(2) . \tag{4.34}$$

As mentioned above, the PDF can be obtained from the CG quasi-PDF through a perturbative matching, as shown in Eq. (4.6). In the CG method, due to the absence of linear renormalon, the leading renormalon resummation (LRR) is not required.

The $\overline{\mathrm{MS}}$ matching coefficient is given in Eq. (4.7), which can be converted to the hybrid scheme as follows:

$$C^{\mathrm{hyb.}}_{(1)}\left(\xi, \frac{\mu}{p^z}, z_s p^z\right) = C^{(1)}_r\left(\xi, \frac{\mu}{p^z}\right) + \delta C^{\mathrm{hyb.}}_{(1)}\left(\xi, z_s p^z\right) , \tag{4.35}$$

where

$$\delta C^{\rm hyb.}_{\gamma^t(1)}(\xi, z_s p^z) = \delta C^{\rm hyb.}_{\gamma^z(1)}(\xi, z_s p^z) = \delta C^{\rm hyb.}_{\gamma^t\gamma_5(1)}(\xi, z_s p^z) = \delta C^{\rm hyb.}_{\gamma^z\gamma_5(1)}(\xi, z_s p^z)$$
$$= \frac{1}{2}\left[\frac{1}{|1-\xi|} - \frac{2{\rm Si}[(1-\xi)z_s p^z]}{\pi(1-\xi)}\right]^{(-\infty,\infty)}_{+(1)}, \tag{4.36}$$
$$\delta C^{\rm hyb.}_{\gamma^t\gamma^\alpha_\perp\gamma_5(1)}(\xi, z_s p^z) = \delta C^{\rm hyb.}_{\gamma^z\gamma^\alpha_\perp\gamma_5(1)}(\xi, z_s p^z) = 0\,. \tag{4.37}$$

The matching coefficient includes the logarithm $\ln\frac{\mu^2}{(2yP^z)^2}$ which follows the DGLAP evolution equation. However, since y is an integration variable, the matching would be ill defined for all x if we naively resum such logarithms. Instead, the physically meaningful logarithms should be given by the external scales after the convolution integral. For Eq. (4.6), they are identified as $\ln\frac{\mu^2}{(2xP^z)^2}$ [258], which indicates that the resummation becomes significant at small x. When $\alpha_s(2xP^z) \gtrsim 1$, the matching loses perturbative control, which provides an estimate of the smallest $x_{\rm min}$ that LaMET can reach.

The procedure for resumming small-x logarithm is as follows: first, for each x we match the quasi-PDF to the PDF at the initial scale $\mu_0(x) = 2\kappa x P^z$ where $\kappa \sim 1$. Then, we evolve the PDF $f(x, \mu_0(x))$ to a fixed $\overline{\rm MS}$ scale $\mu = 2$ GeV for $x \in [x_{\rm min}, 1]$. For matching accuracy at NLO, the two-loop DGLAP evolution kernel [297] is needed, which makes the resummation at next-to-leading logarithmic (NLL) order. Since the DGLAP evolution is closed for $x \in [x_{\rm min}, 1]$, such an evolution kernel can be calculated uniquely following the strategies in Refs. [231, 258].

Apart from the small-x logarithms, the matching coefficient also contains the singular terms $\ln(1-\xi)/(1-\xi)$ and $1/(1-\xi)$ [187], which become important as $\xi \to 1$. These are the standard threshold logarithms that also appear in the matching kernel for the GI quasi-PDFs [220, 237]. The factorization and resummation of such logarithms have been derived for the GI quasi-PDFs [220, 237] and for the quasi generalized parton distributions [298], and the same methodology can be applied to the CG quasi distributions, although this has not yet been carried out in the literature. According to Ref. [237], the relevant external scale is $2(1-x)P^z$, implying that threshold resummation becomes significant—and eventually loses perturbative control—at large x. For this reason, we postpone its implementation in the present work and instead estimate the maximum value $x_{\max}$ allowed by LaMET by imposing the condition $\alpha_s(2(1-x_{\max})P^z) \sim 1$.

4.3.2 Numerical Results

The two-point correlator with hadron momentum $\vec{P}$ is defined as

$$C^{ss'}_{\text{2pt}}\left(t_{\text{sep}}\right) = \sum_{\vec{x}} e^{-i\vec{P}\cdot\vec{x}}\, \mathcal{P}^{\text{2pt}}_{\alpha\beta} \langle N^s_\alpha(\vec{x}, t_{\text{sep}}) \bar{N}^{s'}_\beta(\vec{0}, 0)\rangle \ , \tag{4.38}$$

where the nucleon interpolating operator N is defined as Eq. (4.2), the source is positioned at the origin for simplicity, and the positive parity is projected using $\mathcal{P}^{\text{2pt}} = \frac{1}{2}(1+\gamma^t)$. In this work, we use a Gaussian-smeared source and sink ($s = s' = S$), following the setup in Ref. [299]. The superscript SS will be removed in the following text for simplicity. By inserting a complete set of states, the two-point correlator can be expressed as a sum over energy eigenstates:

$$C_{\text{2pt}}\left(t_{\text{sep}}\right) = \sum_{n=0}^{n_s-1} \frac{\left|z_n\right|^2}{2E_n} \left(e^{-E_n t_{\text{sep}}} + e^{-E_n(L_t - t_{\text{sep}})}\right) \ . \tag{4.39}$$

The overlap amplitude $z_n = \langle\Omega|N_\alpha \mathcal{P}^{\text{2pt}}_{\alpha\beta}|n\rangle$ quantifies the projection of the nucleon interpolator onto the nth energy eigenstate, where $|\Omega\rangle$ denotes the QCD vacuum state and n_s denotes the number of excited states with the same quantum numbers as the nucleon, which are considered within the fit function.

Fit	n_s	$t_{\rm sep}$ range	τ range
$C_{\rm 2pt}(t_{\rm sep})$	2	$t_{\rm sep} \in [4, 12]$	/
$R(t_{\rm sep}, \tau)$	2	$t_{\rm sep} \in [8, 12]$	$\tau \in [3, t_{\rm sep} - 3]$
$\mathrm{FH}(t_{\rm sep}, \tau_{\rm cut} = 3)$	1	$t_{\rm sep} \in [8, 12]$	$\tau \in [3, t_{\rm sep} - 3]$

Table 4.2: Collection of ground state fit settings. #state = 1 means that only one ground state is included in the fit functions.

The bare quasi-PDFs can be extracted from the three-point function with hadron momentum $\vec{P}$

$$C_{\rm 3pt}(t_{\rm sep}, \tau) = \sum_{\vec{x}, \vec{z}_0} e^{-i\vec{P}\cdot\vec{x}}\, \mathcal{P}^{\rm 3pt}_{\alpha\beta} \left\langle N_\alpha(\vec{x}, t_{\rm sep})\, \mathcal{O}_{\tilde{\Gamma}}(\vec{z}_0, z, \tau)\, \bar{N}_\beta(\vec{0}, 0) \right\rangle \,, \tag{4.40}$$

where nucleon polarization projection operator $\mathcal{P}^{\rm 3pt}$ is chosen to be $\frac{1}{2}(1+\gamma_t)$, $\frac{1}{2}(1+\gamma_t)(-i\gamma_5\gamma_z)$, and $\frac{1}{2}(1+\gamma_t)(-i\gamma_5\gamma_\perp)$ for the unpolarized, helicity and transversity quasi-PDF respectively. Throughout, we take the hadron momentum direction as $\hat{z}$ and the transverse spin direction as $\hat{x}$. The insertion operator is given as isovector bilinear quark operators without Wilson line

$$\mathcal{O}_{\tilde{\Gamma}}(\vec{z}_0, z, \tau) = \bar{u}(\vec{z}_0 + z\hat{z}, \tau)\tilde{\Gamma}u(\vec{z}_0, \tau) - \bar{d}(\vec{z}_0 + z\hat{z}, \tau)\tilde{\Gamma}d(\vec{z}_0, \tau) \,, \tag{4.41}$$

where the corresponding Dirac structures $\tilde{\Gamma}$ appropriate for the three channels are γ^t (unpolarized), $i\gamma^z\gamma^5$ (helicity) and $-i\gamma^z\gamma^y$ (transversity). In terms of the spectral decomposition, it can be expanded as

$$C_{\rm 3pt}\left(t_{\rm sep}, \tau\right) = \sum_{n,m=0}^{n_s-1} \frac{z_n^\dagger O_{nm} z_m}{4E_n E_m} e^{-E_n(t_{\rm sep}-\tau)} e^{-E_m\tau} \,, \tag{4.42}$$

where $O_{nm} = \langle n|\mathcal{O}_{\tilde{\Gamma}}|m\rangle$ are the matrix elements of the insertion operator. The ground state matrix element O_{00}, determined by our selection of $\tilde{\Gamma}$, is related to the bare quasi-PDFs $\tilde{h}^0_{\tilde{\Gamma}}$ through the relationship $O_{00} = 2E_0\tilde{h}^0_{\tilde{\Gamma}}$.

We employ the Feynman–Hellmann (FH) method [100, 300] to suppress excited-state contamination in the extraction of the ground-state matrix elements. The analysis is carried out using a fully correlated Bayesian least-squares procedure. We first fit the two-point correlator $C_{2\text{pt}}$ from which the ground- and excited-state energies and amplitudes (E_0, E_1, z_0, z_1) are obtained. The posterior distributions of these parameters are then used as Bayesian priors in a subsequent joint fit of the ratio

$$R\left(t_{\text{sep}}, \tau\right) = \frac{C_{3\text{pt}}\left(t_{\text{sep}}, \tau\right)}{C_{2\text{pt}}\left(t_{\text{sep}}\right)}\,, \tag{4.43}$$

together with the FH correlators. The FH correlator is defined as

$$S\left(t_{\text{sep}}, \tau_{\text{cut}}\right) \equiv \textstyle\sum_{t=\tau_{\text{cut}}}^{t=t_{\text{sep}}-\tau_{\text{cut}}} R\left(t_{\text{sep}}, \tau\right) \tag{4.44}$$

$$\text{FH}\left(t_{\text{sep}}, \tau_{\text{cut}}, d\tau\right) \equiv \tfrac{S(t_{\text{sep}}+d\tau, \tau_{\text{cut}})-S(t_{\text{sep}}, \tau_{\text{cut}})}{d\tau}\,, \tag{4.45}$$

where $d\tau$ is set to unit. In the large-source–sink separation limit $t_{\text{sep}} \to \infty$, both R and FH converge asymptotically to the same bare quasi-PDF matrix element, $\tilde{h}^0_{\tilde{\Gamma}}$. The entire chained procedure—first fitting $C_{2\text{pt}}$ and then performing the joint fit of R and FH with the two-point posteriors as priors—is repeated on each jackknife sample to ensure robust and reliable uncertainty estimates. The detailed setting can be found in Table 4.2.

As discussed above, we employ the hybrid scheme as Eq. (4.31) to renormalize the real part of the quasi-PDF matrix elements, with the separation point is chosen as $z_s = 3\sqrt{2}a$; while the SRI′-MOM scheme as Eq. (4.33) is used for the imaginary part. In addition, to preserve the antisymmetry

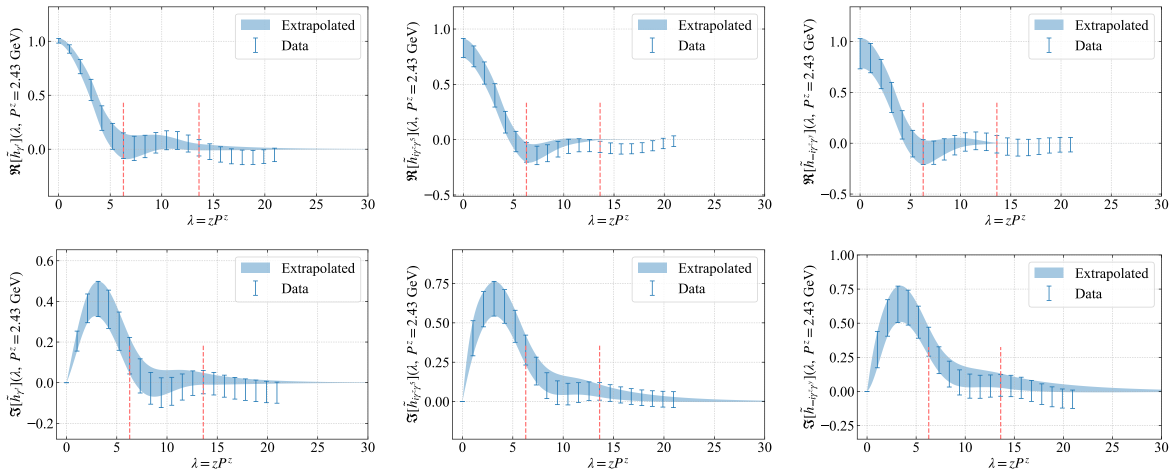

Figure 4.5: Extrapolation of the renormalized quasi-PDF matrix elements in coordinate space as functions of $\lambda = zP^z$. The three columns from left to right correspond to the unpolarized, helicity, and transversity channels, while the upper (lower) row shows the results of real (imaginary) part at $P^z = 2.43$ GeV. The data points are shown with error bars, and the shaded bands represent the extrapolation using the asymptotic form. The two red dashed vertical lines indicate the sub-asymptotic region used in the fit.

of the imaginary part, its local contribution at $z = 0$ is explicitly set to zero. The renormalized matrix elements in the coordinate space are shown in Fig. 4.5 for three polarization channels.

The quasi-PDF matrix elements decrease rapidly as a function of $\lambda = zP^z$, approaching zero for $\lambda \gtrsim 5$. Nevertheless, as discussed above, an asymptotic extrapolation remains necessary to suppress unphysical fluctuations in the uncertainty band. A sub-asymptotic form for nucleon PDFs in GI formulations was derived in Ref. [279], as given in Eq. (4.5). In this form, the large-z behavior is characterized by an exponential suppression controlled by the lightest intermediate state, accompanied by a power-law prefactor and a phase factor that depends on the sign of λ.

While the exponential decay structure is robust, the precise power-law behavior depends on the details of the theoretical formulation. In particular, for the CG formalism, the corresponding algebraic dependence has not yet been fully understood from first principles. Motivated by these considerations, we adopt the following parameterization for the sub-asymptotic behavior in the CG,

$$\tilde{h}(\lambda) = \left[A_2 \cdot e^{i\text{sign}(\lambda)\phi_2} + \frac{A_2'}{|\lambda|} \cdot e^{i\text{sign}(\lambda)\phi_2'}\right] \frac{e^{-m|\lambda|}}{|\lambda|^n}, \tag{4.46}$$

where the overall factor $1/|\lambda|^n$ is introduced to parametrize the unknown algebraic decay. This form is used to fit the data in the sub-asymptotic region $0.5 \text{ fm} \leq z \leq 1.1 \text{ fm}$, as indicated by the two red dashed lines in Fig. 4.5. In the fit function above, there are six parameters: A_2, A_2', ϕ_2, ϕ_2', m and n, where m is set to be larger than 0.1 GeV / P^z [188]. To get a smooth change of the error band, the extrapolated results are given in the form as

$$\tilde{h}^{\text{ext}} = w \cdot \tilde{h}^{\text{data}} + (1-w) \cdot \tilde{h}^{\text{fit}} , \tag{4.47}$$

where w is a weight function that transitions linearly from 1 to 0 within the fit range, $\tilde{h}^{\text{data}}$ is the renormalized quasi-distribution, and $\tilde{h}^{\text{fit}}$ refers to the fit results of the asymptotic fit. After applying this extrapolation, the uncertainty bands smoothly converge to zero, mitigating non-physical fluctuations. The model uncertainty of this asymptotic extrapolation is bounded from above and can be improved systematically [188, 280].

4.3.2.1 Unpolarized PDF

As discussed above, the real and imaginary parts of the matrix elements are renormalized in different schemes, and their contributions correspond to the valence and full quark channels of the nucleon PDFs, respectively. Therefore, these two contributions can be independently compared with existing results in the literature. After performing the large-momentum expansion, the matched results

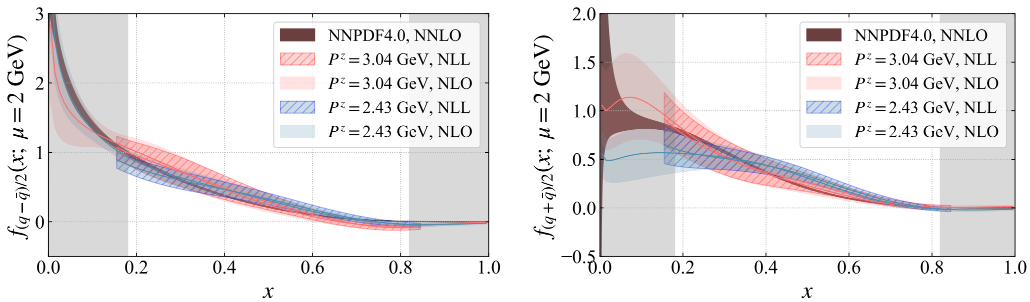

Figure 4.6: Unpolarized quark PDFs of the nucleon at momenta $P^z = 2.43$ and 3.04 GeV, extracted from lattice QCD and compared with the NNLO phenomenological results from NNPDF4.0 [33]. The NLO and NLL labels indicate the perturbative accuracy of matching coefficients. The left and right panels show the valence contribution $(q - \bar{q})/2$ and the full distribution $(q + \bar{q})/2$, respectively. Results obtained using matching coefficients at NLO and NLL are shown to illustrate the perturbative convergence. The shaded gray regions ($x < 0.18$ or $x > 0.82$) denote kinematic domains where the strong coupling $\alpha_s(\mu = 2xP^z)$ or $\alpha_s(\mu = 2(1-x)P^z)$ at NLO exceeds 0.5, indicating a poor convergence of perturbation theory.

from the real and imaginary parts are shown in the left and right panels of Fig. 4.6, respectively. The unpolarized nucleon PDFs at hadron momenta $P^z = 2.43$ and 3.04 GeV are presented together with the NNLO phenomenological determination from NNPDF4.0 [33]. By employing matching coefficients at different perturbative accuracies, we further obtain results at next-to-leading order (NLO) as well as at next-to-leading logarithmic (NLL) accuracy with DGLAP evolution. The shaded gray regions in Fig. 4.6 indicate kinematic domains where the strong coupling $\alpha_s(\mu = 2xP^z)$ evaluated at NLO exceeds 0.5. In these regions, the perturbative expansion underlying the matching procedure is expected to have poor convergence, and the corresponding lattice results should therefore be interpreted with caution.

The results in Fig. 4.6 show that the lattice calculations at hadron momentum $P^z = 3.04$ GeV are in good agreement with the determination of NNPDFs4.0 for both the valence and the full components of the unpolarized PDFs. The comparison between the NLO and NLL cases further indicates a stable perturbative behavior and supports the convergence of the matching expansion. In the right panel,

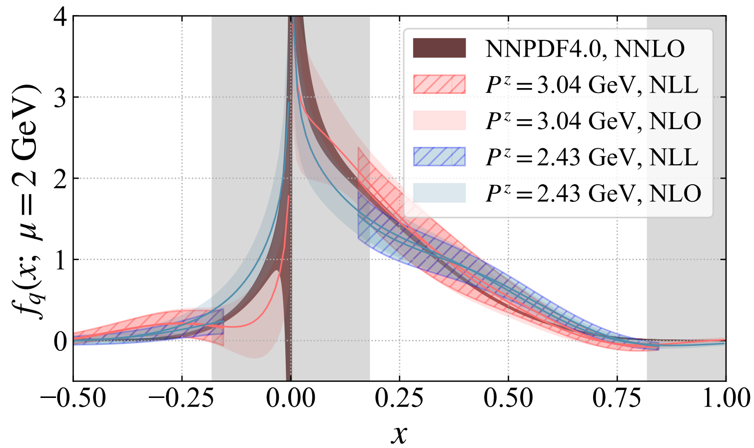

Figure 4.7: Unpolarized quark PDFs of the nucleon at momenta $P^z = 2.43$ and $P^z = 3.04$ GeV. The lattice results are shown together with the NNLO phenomenological determination from NNPDF4.0 for comparison.

corresponding to $(q+\bar{q})/2$, the difference between the two lattice momenta is more pronounced than in the $(q-\bar{q})/2$ channel, which is likely caused by residual excited-state contamination in the extraction of the bare quasi-matrix elements. By combining the valence and sea contributions shown in Fig. 4.6, we present the resulting unpolarized nucleon PDFs in Fig. 4.7, together with the results of NNPDF4.0 [33] for comparison.

4.3.2.2 Helicity PDF

Similar to the unpolarized case, the full $(q+\bar{q})/2$ and valence $(q-\bar{q})/2$ components of the helicity PDFs are shown in the left and right panels of Fig. 4.8, respectively. Note that the full helicity distribution is obtained from the real part of the matrix elements, since the helicity quasi-distribution is even under charge conjugation. The helicity nucleon PDFs at hadron momenta $P^z = 2.43$ and 3.04 GeV are compared with phenomenological determinations, including the NLO results from NNPDFpol1.1 [301], the NNLO results from NNPDFpol2.0 [291], as well as the BDSSV24 [302], MAPPDFpol1.0 [303], and JAMpol25 [304] analyses. All phenomenological results are normalized

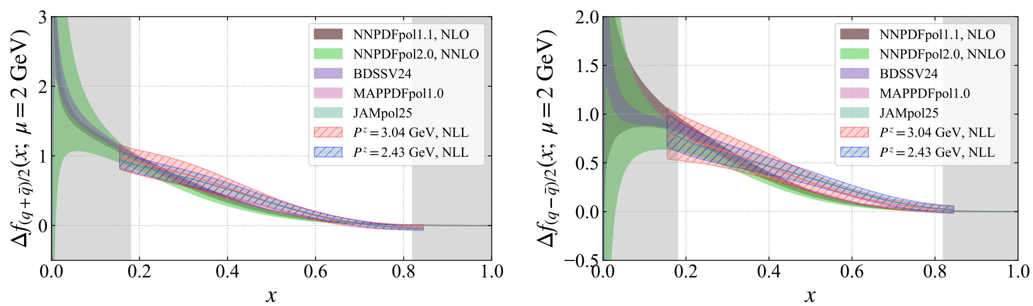

Figure 4.8: Helicity quark PDFs of the nucleon at momenta $P^z = 2.43$ and $P^z = 3.04$ GeV, compared with phenomenological determinations. The left and right panels show the full $(q + \bar{q})/2$ and valence $(q - \bar{q})/2$ components, respectively. The lattice results are presented at NLL accuracy, while the phenomenological bands include the NLO NNPDFpol1.1 [301], NNLO NNPDFpol2.0 [291], BDSSV24 [302], MAPPDFpol1.0 [303], and JAMpol25 [304] analyses.

by their respective axial charges g_A, while the valence $(q - \bar{q})/2$ components of the lattice results are normalized using the g_A value from JAMpol25 [304].

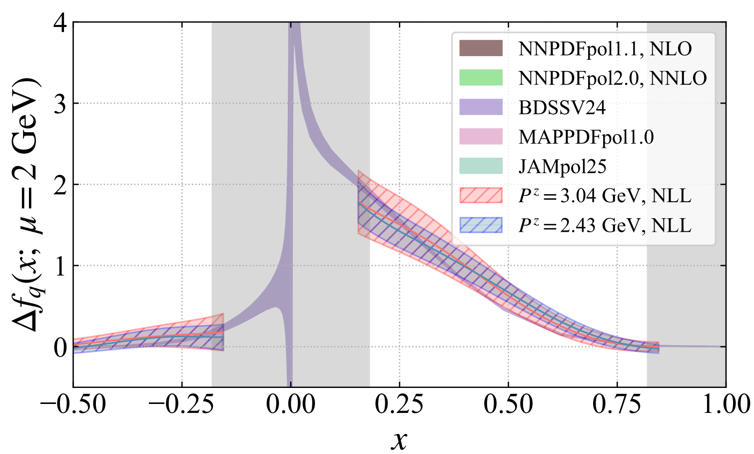

Figure 4.9: Helicity quark PDFs of the nucleon at momenta $P^z = 2.43$ and $P^z = 3.04$ GeV. The lattice-extracted distributions are shown together with phenomenological determinations for comparison.

From the plots in Fig. 4.8, we find that our results at $P^z = 3.04$ GeV are consistent with the phenomenological ones within one standard deviation for both the full and valence parts. Moreover, as in the unpolarized case, the discrepancy between the two momenta is more pronounced in the contribution from the imaginary part of the matrix elements. Combining the full and valence parts, the helicity nucleon PDFs are presented in Fig. 4.9.

4.3.2.3 Transversity PDF

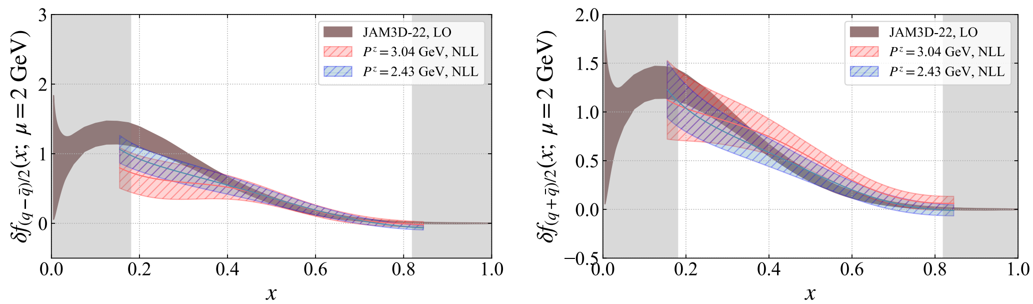

Figure 4.10: Transversity quark PDFs of the nucleon at momenta $P^z = 2.43$ and $P^z = 3.04$ GeV, compared with phenomenological and lattice determinations. The left and right panels show the valence $(q - \bar{q})/2$ and full $(q + \bar{q})/2$ components, respectively. The lattice results at NLL accuracy are compared with the leading-order phenomenological extraction from JAM3D-22 [305].

In analogy with the unpolarized and helicity cases, we also present our results for the transversity PDFs. The valence $(q - \bar{q})/2$ and full $(q + \bar{q})/2$ contributions are shown in the left and right panels of Fig. 4.10, respectively. The transversity nucleon PDFs at hadron momenta $P^z = 2.43$ and 3.04 GeV are compared with the phenomenological extraction from JAM3D-22 [305] at leading order (LO). The JAM3D-22 distributions are normalized by their own tensor charge g_T, and the full $(q + \bar{q})/2$ contribution of the lattice results are normalized using the same value to ensure a consistent comparison.

As illustrated in Fig. 4.10, we observe that our lattice results show good agreement with the JAM3D-22 phenomenological extractions in the moderate-x region. The level of consistency is even more evident in the comparison of the full part of the transversity PDFs. Similar to the unpolarized and helicity cases, the discrepancy between the two momenta is more pronounced in the contribution from the imaginary part of the matrix elements, underlining once again the need to carefully control

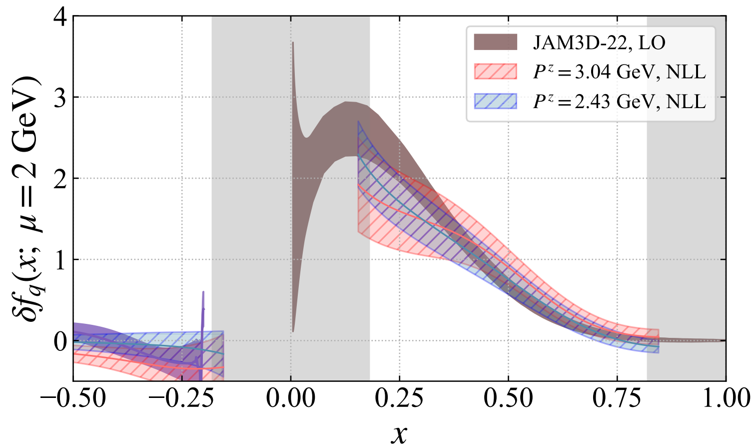

Figure 4.11: Transversity quark PDFs of the nucleon at momenta $P^z = 2.43$ and $P^z = 3.04$ GeV, shown together with phenomenological results for comparison. The lattice-extracted PDFs are compared with the JAM3D-22 extraction at LO.

excited-state contamination in this sector. The combined valence and full parts of the transversity nucleon PDFs are presented in Fig. 4.11. After all, since the phenomenological extraction of the transversity PDF still carries substantial uncertainties, caution is required in our comparison.

4.3.3 Conclusion

In this work, we have carried out the first lattice QCD calculation of isovector nucleon PDFs using the CG method, including all three polarization channels. Employing an off-axis momentum setup, we achieved large hadron momenta up to $P^z > 3$ GeV, which allows for a controlled application of the LaMET framework. A systematic treatment of excited-state contamination was performed, and we demonstrated that the CG method produces nucleon PDFs in the moderate-x region that are broadly consistent with phenomenological determinations. These findings validate the CG method as a promising tool for extracting precision-controlled nucleon parton distributions from lattice QCD.

More specifically, at the largest momentum in this work, $P^z = 3.04$ GeV, we observed an encouraging agreement between our lattice results and the existing phenomenological extractions for all three polarizations of nucleon PDFs: unpolarized, helicity, and transversity. The consistency is more evident in the contributions from the real part of the matrix elements. In contrast, the contributions from the imaginary part of the matrix elements display larger discrepancies across different hadron momenta. This suggests that the imaginary sector is more sensitive to residual excited-state contamination and may require even more careful analyses in future studies.

The renormalization structure of the Coulomb-gauge quasi-PDF operators is significantly simpler than that of the GI operator due to the lack of the Wilson line (and hence, lack of linear divergence and associated renormalon). At short separations $z \sim a$, the operator $\bar{\psi}_0(z)\tilde{\Gamma}\psi_0(0)$ picks up large discretization effects. The hybrid scheme should remove the leading artifacts for the real part of the matrix elements, which, however, does not necessarily work for the imaginary part. Therefore, we propose to renormalize the latter using the SRI$'$-MOM scheme, which is a simple method for renormalizing the Coulomb-gauge quark fields. To further understand the lattice artifacts, a full study of these schemes is planned by studying the quark mass and lattice spacing dependence of the renormalization constants.

Together, these observations highlight both the promise and the challenges of extracting partonic information from the lattice at large momentum. The path forward is clear. With access to greater computational resources, it will become possible to carry out simulations at multiple lattice spacings, thereby enabling controlled continuum extrapolations that systematically reduce discretization effects. In parallel, accumulating higher statistics for correlators at large source–sink separations will improve the stability of the extracted observables and enhance the overall robustness of the final results. Further progress can be achieved by adopting more refined and effective strategies for treating excited-state contamination, such as the Lanczos algorithm [306–308] and the current-inspired Generalized Eigenvalue Problem (GEVP) [309, 310], which is expected to be particularly important for reliably resolving the imaginary components of matrix elements. Finally, the application of new lattice technologies aimed at improving the signal-to-noise ratio (SNR), such as kinematically enhanced interpolating operators [278] and more optimized smearing schemes, will be essential to extend these studies to higher momenta and increased precision. Combining these improvements will place the CG method on a firmer foundation and allow precise lattice determinations of nucleon PDFs that can stand alongside phenomenological and experimental results. In this way, our present work provides not only a first demonstration of the CG method for nucleon PDFs, but also a roadmap for systematically improving the approach toward a future of quantitatively precise lattice QCD parton distributions.

4.4 Calculation of Pion TMDs

In high-energy scatterings, transverse-momentum-dependent parton distribution functions (TMDPDFs) provide a fundamental description of the transverse momentum and polarization degrees of freedom of quarks and gluons within hadrons [49]. These distributions play a crucial role in understanding the intricate dynamics of quark-gluon interactions and the phenomenon of color confinement. Knowledge of TMDPDFs is essential for predicting observables in multi-scale, non-inclusive high-energy processes, such as semi-inclusive deep-inelastic scattering (SIDIS) and Drell-Yan (DY), based on QCD factorization. Experiments at facilities including the Jefferson Lab 12 GeV Program [311], the EIC [7, 8], and the LHC [312] rely heavily on accurate knowledge of the TMDPDFs.

Significant progress has been made in the phenomenological parameterizations of TMDPDFs, particularly for the nucleon, through global analyses of experimental data [313–333]. While these studies have significantly improved our understanding of the transverse-momentum structure of quarks and gluons within the nucleon, they remain in their early stages due to the limited availability of experimental data sensitive to the non-perturbative region. Compared to nucleons, much less is known about the transverse momentum structure of the pion [334–336]. As the lightest QCD bound state, the pion plays a fundamental role in hadron structure and non-perturbative QCD. Its TMDPDFs provide crucial insights into the internal dynamics of quarks and gluons in a strongly interacting relativistic system, with direct implications for hadronization processes and the underlying mechanisms of non-perturbative QCD. However, the scarcity of experimental data on pion TMDPDFs, combined with the inherent nature of the parameterization dependence of global analyses, leaves significant gaps in our non-perturbative first-principles understanding. Lattice QCD calculations provide a natural approach to gain insight into non-perturbative TMD structures of the pion starting from first-principles QCD.

In our recent work [246], we present the first lattice QCD calculation of the unpolarized valence-quark pion TMDPDF using the CG method. We compute the quasi-TMDPDF and quasi-TMDWF matrix elements of the pion in CG. These matrix elements enable the extraction of the CS kernel, the intrinsic soft function, the TMDWF, and ultimately the unpolarized valence TMDPDF of the pion. Furthermore, we improve the perturbative accuracy of the matching procedure, particularly in the extraction of the intrinsic soft function, by resumming logarithmic terms in the Sudakov kernel, an aspect that has not been accounted for in previous lattice QCD studies. Finally, the comparison of our results with global fits of the experimental data shows encouraging consistency, demonstrating the bright potential of the CG approach for TMD physics from lattice QCD.

4.4.1 Theoretical Framework

As proposed in Ref. [224], the light-cone TMDPDF can be derived from the CG quasi-TMD beam function defined as

$$\tilde{f}_{\Gamma}(x, b_{\perp}, P^z; \mu) = P^z \int \frac{dz}{2\pi} e^{iz(xP^z)} \tilde{h}_{\Gamma}(z, b_{\perp}, P^z; \mu) , \tag{4.48}$$

where the matrix elements are given by

$$\tilde{h}_{\Gamma}(z, b_{\perp}, P^z; \mu) = \frac{1}{2P_{\Gamma}} \langle P| \, \bar{q}(z, b_{\perp}) \Gamma q(0)|_{\vec{\nabla}\cdot\vec{A}=0} \, |P\rangle , \tag{4.49}$$

where P_{Γ} is the kinetic factor for normalization. In this work, we choose $\Gamma = \gamma^t$ for quasi-TMD beam function and $P_{\Gamma} = P^t$. The parameter μ represents the $\overline{\mathrm{MS}}$ renormalization scale, and the space-like separation between the quark and the antiquark is denoted as $\vec{b} \equiv (z, \vec{b}_{\perp})$. The hadron state carries momentum $P = (P^t, 0, 0, P^z)$ and satisfies the relativistic normalization condition $\langle P|P\rangle = 2P^t(2\pi)^3\delta^{(3)}(\vec{0})$. When the hadron moves with a large momentum P^z, the quasi-TMD beam function factorizes as

$$\tilde{f}_{\Gamma}(x, b_{\perp}, P^z; \mu) = H_f \left(x, P^z; \mu\right) B(x, b_{\perp}, xP^+; \mu, \nu) \times S_C^0(b_{\perp}; \mu, \nu) , \tag{4.50}$$

where $B(x, b_{\perp}, xP^+; \mu, \nu)$ is the light-cone beam function with zero-bin subtraction, following the procedure outlined in Refs. [52, 337]. The term $S_C^0(b_{\perp}; \mu, \nu)$ represents the zero-bin contribution, and the operator definitions of both functions are detailed in Ref. [224]. The parameters μ and ν correspond to the renormalization scales associated with the ultraviolet (UV) and rapidity divergences, respectively. The function $H_f\left(x, P^z; \mu\right) = |C_{\rm TMD}(xP^z;\mu)|^2$ represents the hard kernel that matches the QCD quark field operator to Soft-Collinear Effective Theory (SCET) [177].

The light-cone TMDPDF can be defined from the light-cone beam function B and the soft function S as

$$f(x, b_\perp; \mu, \zeta) = B(x, b_\perp, xP^+; \mu, \nu) S(b_\perp, y_n; \mu, \nu) \,, \tag{4.51}$$

where the dependence on the rapidity renormalization scale ν cancels out on the right-hand side, leaving only a dependence on the Collins-Soper scale $\zeta \equiv 2(xP^+)^2 e^{-2y_n}$. Combining Eqs. (4.50) and (4.51), the factorization formula connecting the quasi-TMDPDF and the light-cone TMDPDF is given by [134, 224, 281, 282, 338]

$$\sqrt{S_I\left(b_\perp, y_n; \mu\right)} \cdot \tilde{f}_\Gamma(x, b_\perp, P^z; \mu) = f(x, b_\perp; \mu, \zeta) H_f\left(x, P^z; \mu\right) + \text{p.c.} \,, \tag{4.52}$$

where p.c. indicates the power corrections, and the intrinsic soft function S_I is defined as

$$S_I\left(b_\perp, y_n; \mu\right) \equiv \left(\frac{S(b_\perp, y_n; \mu, \nu)}{S_C^0(b_\perp; \mu, \nu)}\right)^2 \,. \tag{4.53}$$

Here S_C^0 is the CG quasi-soft function, as defined in Ref. [224], which exactly cancels the rapidity divergences in S. The above factorization formula is of the same form as that for the GI quasi-TMDPDFs [134, 281, 282, 338], while the method to calculate S_I was first proposed in Ref. [338], which enables a complete determination of the TMDs from the lattice.

The intrinsic soft function satisfies the evolution equation

$$\frac{d}{dy_n} \ln S_I\left(b_\perp, y_n; \mu\right) = -2\gamma^{\overline{\rm MS}}(b_\perp; \mu) \,, \tag{4.54}$$

where $\gamma^{\overline{\mathrm{MS}}}(b_\perp;\mu)$ is the Collins-Soper kernel. Therefore, the intrinsic soft function at any y_n satisfies

$$S_I\left(b_\perp, y_n;\mu\right) = S_I\left(b_\perp, y_n = 0;\mu\right) \cdot \exp(-2y_n \cdot \gamma^{\overline{\mathrm{MS}}}(b_\perp;\mu)) \, . \tag{4.55}$$

One can redefine the intrinsic soft function as $S_I\left(b_\perp;\mu\right) \equiv S_I\left(b_\perp, y_n = 0;\mu\right)$, allowing us to rewrite Eq. (4.52) in a more explicit form, as shown in Eq. (4.14). This factorization framework provides a systematic approach to relating quasi-TMDPDFs extracted from lattice QCD to their light-cone counterparts, ensuring a well-defined separation between perturbative and non-perturbative contributions. However, power corrections of order $\mathcal{O}\left(\frac{\Lambda_{\mathrm{QCD}}}{xP^z}, \frac{1}{b_\perp xP^z}\right)$ [225] introduce systematic uncertainties, particularly in the small-x region. Despite these limitations, the framework offers valuable first-principles constraints in the large-x regime, where power corrections are expected to be relatively small.

Similar to the quasi-TMD beam function, the light-cone TMDWF can be derived from the CG quasi-TMDWF, which is defined as

$$\tilde{\phi}_\Gamma(x, b_\perp, P^z;\mu) = P^z \int \frac{dz}{2\pi} e^{iz(xP^z)} \tilde{\varphi}_\Gamma(z, b_\perp, P^z;\mu) \, , \tag{4.56}$$

with the quasi-TMDWF matrix elements given by

$$\tilde{\varphi}_\Gamma(z, b_\perp, P^z;\mu) = \frac{e^{-izP^z/2}}{f_\pi P_\Gamma} \langle\Omega| \, \bar{q}(z/2, b_\perp/2)\Gamma q(-z/2, -b_\perp/2)|_{\vec{\nabla}\cdot\vec{A}=0} \, |P\rangle \, , \tag{4.57}$$

which corresponds to a pion-to-vacuum matrix element, the $|\Omega\rangle$ represents the QCD vacuum state, the f_π is the pion decay constant, and P_Γ is the kinetic factor for normalization. In this work, we choose $\Gamma = \gamma^z\gamma^5$ for quasi-TMDWFs with nonzero momenta and $P_\Gamma = iP^z$, while for zero-momentum quasi-TMDWFs, we set $\Gamma = \gamma^t\gamma^5$ and $P_\Gamma = P^t$.

In the large-momentum limit, the quasi-TMDWF can be matched to the light-cone TMDWF via the factorization formula

$$\sqrt{S_I\left(b_\perp, y_n; \mu\right)} \cdot \tilde{\phi}_\Gamma\left(x, b_\perp, P^z; \mu\right) = \phi(x, b_\perp, y_n; \mu, \zeta, \bar{\zeta}) H_\phi\left(x, \bar{x}, P^z; \mu\right) + \text{p.c.}\,, \tag{4.58}$$

with $\bar{x} = 1 - x$, $\bar{\zeta} \equiv 2(\bar{x}P^+)^2 e^{-2y_n}$, and the hard kernel

$$H_\phi(x, \bar{x}, P^z; \mu) = C_{\text{TMD}}(xP^z; \mu) \cdot C_{\text{TMD}}(\bar{x}P^z; \mu)\,, \tag{4.59}$$

calculated up to NLO in Ref. [224].

An essential component of quasi-TMD factorization is the CS kernel $\gamma^{\overline{\text{MS}}}(b_\perp; \mu)$, which governs the rapidity evolution of TMDs and is crucial to achieve consistent matching between quasi-TMDPDFs and light-cone TMDPDFs. The CS kernel can be extracted from the ratios of quasi-TMD matrix elements computed at different hadron momenta [106, 134, 338]. In this work, we employ the ratio of CG quasi-TMDWFs [25, 338] as

$$\begin{aligned}
\gamma^{\overline{\text{MS}}}(b_\perp; \mu) \quad &= \gamma^{\overline{\text{MS}}}(b_\perp, P_1, P_2; \mu) + \text{p.c.} \\
&= \frac{1}{\ln\left(P_2/P_1\right)} \ln \frac{H_\phi\left(x, \bar{x}, P_1; \mu\right) \tilde{\phi}_{\gamma^z\gamma^5}\left(x, b_\perp, P_2; \mu\right)}{H_\phi\left(x, \bar{x}, P_2; \mu\right) \tilde{\phi}_{\gamma^z\gamma^5}\left(x, b_\perp, P_1; \mu\right)} \\
&\quad + \mathcal{O}\left(\frac{\Lambda_{\text{QCD}}}{xP^z}, \frac{1}{b_\perp\left(xP^z\right)}, \frac{\Lambda_{\text{QCD}}}{\bar{x}P^z}, \frac{1}{b_\perp\left(\bar{x}P^z\right)}\right)\,.
\end{aligned} \tag{4.60}$$

This method offers a key advantage over direct extractions from quasi-TMDPDFs, as quasi-TMDWFs naturally peak around $x = 0.5$, a region where power corrections are significantly suppressed. On the other hand, the quasi-TMDPDFs usually decrease quickly at moderate to large x, making the extraction of the CS kernel more susceptible to power corrections.

Another essential component of the factorization of quasi-TMDs is the intrinsic soft function, which accounts for soft gluon radiation in the process. Its calculation is particularly challenging for two reasons. First, as a non-perturbative quantity, it cannot be computed using standard perturbative techniques. Second, because the soft function involves correlations along two light-cone directions, it cannot be directly simulated on the lattice, even with a large-momentum boost.

A practical approach to determining the intrinsic soft function, as proposed in Refs. [338], is to use the TMD factorization of a pseudoscalar light-meson form factor, which is defined as

$$F\left(b_{\perp}, P_1, P_2, \Gamma, \Gamma'\right) \equiv -4N_c \frac{\langle P_2|\, \bar{q}\left(b_{\perp}\right) \Gamma q\left(b_{\perp}\right) \bar{q}(0)\Gamma' q(0)\,|P_1\rangle}{f_{\pi}^2(P_1 \cdot P_2)}\ , \tag{4.61}$$

where $|P_1\rangle$ and $|P_2\rangle$ denoting the pion states with momentum P_1 and P_2. The coefficient $-4N_c = -12$ is a normalization factor [198]. This form factor can be extracted using the following ratio

$$R_F\left(b_{\perp}, P_1, P_2, \Gamma, \Gamma'\right) \equiv -4N_c \frac{\langle P_2|\, \bar{q}\left(b_{\perp}\right) \Gamma q\left(b_{\perp}\right) \bar{q}(0)\Gamma' q(0)\,|P_1\rangle}{\langle 0|\bar{q}(0)\gamma^{\mu}\gamma^5 q(0)\,|P_1\rangle \,\langle P_2|\, \bar{q}(0)\gamma_{\mu}\gamma^5 q(0)|0\rangle}\ . \tag{4.62}$$

To extract the leading-twist contribution, one can choose the Dirac matrices as

$$\Gamma = \Gamma' \in \{I, \gamma^5, \gamma^{\perp}, \gamma^{\perp}\gamma^5\} \tag{4.63}$$

as suggested in Refs. [198, 222]. In this work, we choose $\Gamma = \Gamma' \in \{\gamma^{\perp}, \gamma^{\perp}\gamma^5\}$. The denominator of the ratio R_F consists of two-point functions used to normalize the form factor, and the index μ should be summed. The momenta in the external pion states are taken as $P_1 = (P^t, 0, 0, P^z)$ and $P_2 = (P^t, 0, 0, -P^z)$, moving back-to-back to achieve a large momentum transfer $Q^2 = (2P^z)^2$. In this kinematic regime, the form factor can be factorized in terms of the TMDWF $\phi(x, b_\perp, y_n; \mu, \zeta)$ [25, 198, 338] as

$$F(b_\perp, P^z) = \int dx_1 dx_2 H_F(x_1, x_2, P^z; \mu)\phi^\dagger(x_1, b_\perp, y_n; \mu, \zeta_1, \bar{\zeta}_1)\phi(x_2, b_\perp, -y_n; \mu, \zeta_2, \bar{\zeta}_2) \ . \quad (4.64)$$

where H_F is the hard kernel encapsulating the short-distance dynamics of the process. The y_n-dependence cancels between ϕ and $\phi^\dagger$, and the hard kernel can be expressed as the product of two Sudakov kernels [198]:

$$H_F(x_1, x_2, P^z; \mu) = C_{\rm Sud}(x_1, x_2, P^z; \mu) \cdot C_{\rm Sud}(\bar{x}_1, \bar{x}_2, P^z; \mu) \ , \quad (4.65)$$

where $C_{\rm Sud}$ represents the Sudakov kernel, which has been computed at one-loop order in the literature [198, 339].

By combining the factorization of the quasi-TMDWF in Eq. (4.58) with the form factor factorization in Eq. (4.64), the intrinsic soft function at $y_n = 0$ can be extracted as

$$S_I(b_\perp; \mu) = \frac{F(b_\perp, P^z)}{\int dx_1 dx_2 H_F(x_1, x_2, P^z; \mu)\tilde{\Phi}^\dagger(x_1)\tilde{\Phi}(x_2)} \ , \quad (4.66)$$

where the reduced quasi-TMDWF $\tilde{\Phi}(x)$, is defined as

$$\tilde{\Phi}(x) \equiv \frac{\tilde{\phi}_\Gamma\left(x, b_\perp, P^z; \mu\right)}{H_\phi\left(x, \bar{x}, P^z; \mu\right)} \ . \quad (4.67)$$

Using different renormalization schemes for quasi-TMDWF, one can get different intrinsic soft functions, which can be perturbatively converted to each other at small $b_\perp$. However, the scheme dependence will eventually cancel between the renormalized quasi-beam and intrinsic soft functions.

4.4.2 Numerical Results

We perform a numerical lattice QCD calculation using a $2+1$-flavor gauge ensemble generated by the HotQCD Collaboration [292]. The ensemble employs the Highly Improved Staggered Quark (HISQ) action [284] with a lattice spacing of $a = 0.06$ fm and a volume of $L_s^3 \times L_t = 48^3 \times 64$. The valence sector is treated using tadpole-improved clover Wilson fermions on a hypercubic (HYP) smeared [293] gauge background. The clover coefficient is set to $c_{\rm sw} = u_0^{-3/4}$, where u_0 is the average plaquette after HYP smearing. For this ensemble, we use $c_{\rm sw} = 1.0336$ in both the time and spatial directions. The valence quark masses are tuned to yield a valence pion mass of 300 MeV, with a corresponding hopping parameter of $\kappa \approx 0.12623$.

The key requirement for the factorization of quasi-TMDs is a sufficiently large hadron momentum. To achieve a higher boost factor, we take advantage of the three-dimensional rotational symmetry of the CG approach and adopt off-axis momenta for the quasi-TMD, choosing momentum directions along $\vec{n} = (n^x, n^y, 0)$. The hadron momenta on the lattice are given by $P^z = \frac{2\pi|\vec{n}|}{L_s a}$. In this study, we consider $n^x = n^y \in \{3, 4, 5\}$, which allows us to reach a maximum hadron momentum of $P^z = 3.04$ GeV for the quasi-TMD calculations. To optimize the signal-to-noise ratio and enhance overlap with large-momentum hadron ground states, we employ boosted Gaussian smearing [340], using the same setup as in Ref. [217]. To extract the ground-state contribution, we compute quasi-TMD three-point functions for multiple source-sink separations, choosing $t_{\rm sep}/a = 6, 8, 10, 12$. Calculations are performed on 553 gauge configurations and we apply the All-Mode Averaging (AMA) technique [341] to further improve the signal. The stopping criteria for the sloppy and exact inversions are set to 10^{-4}

and 10^{-10}, respectively, aligning with the settings in Ref. [167]. For the smaller separation $t_{\text{sep}}/a = 6$, we compute 1 exact source and 32 sloppy sources per configuration, while for the larger separations $t_{\text{sep}}/a = 8, 10, 12$, we compute 4 exact sources and 128 sloppy sources.

In addition, we compute the large-momentum pion form factor following the strategy outlined in Ref. [171]. Specifically, we evaluate the two-current three-point function

$$C_{\text{F}}(t_{\text{sep}}, \tau) = \left\langle \pi_{\text{w}}(-\vec{P}, t_{\text{sep}})\, j_u(b_\perp, \tau)\, j_d(0, \tau)\, \pi_{\text{w}}^\dagger(\vec{P}, 0) \right\rangle, \tag{4.68}$$

where the two currents $j_u(b_\perp, \tau) \equiv \bar{u}(b_\perp, \tau)\Gamma u(b_\perp, \tau)$ and $j_d(0, \tau) \equiv \bar{d}(0, \tau)\Gamma d(0, \tau)$ are separated by a transverse distance $b_\perp$, inserted at the same time slice τ. The Dirac structures are chosen as $\Gamma \in \{\gamma^\perp, \gamma^\perp\gamma^5\}$ to extract the leading-twist contribution [198, 222]. Unlike the two-point and quasi-TMD correlator calculations, where Gaussian-smeared sources are used, here we adopt wall sources π_{w} for the pions. Given the significant computational demands of this calculation and the universality of the soft factor, insensitive to the choice of hadron species, we choose to employ a heavier valence pion mass of 670 MeV. Consequently, our analysis of the pion form factor is limited to a subset of 100 gauge configurations using wall sources in $L_t = 64$ time slices. We compute the form factor at a single momentum $\vec{n} = (0, 0, 6)$, corresponding to $P^z = 2.58$ GeV. To ensure consistency in the extraction of the intrinsic soft function, we also compute the quasi-TMDWF using the same valence pion mass. For quasi-TMDWF, we use momenta in the z-direction with $\vec{n} = (0, 0, n^z)$ and $n^z \in \{8, 9, 10\}$, reaching a maximum momentum of $P^z = 4.30$ GeV. This part of the calculation is performed on the same 553 HotQCD gauge configurations used in the main quasi-TMD analysis, with 1 exact source and 32 sloppy sources per configuration.

To extract the ground-state quasi-TMD matrix elements, it is essential to first determine the energy spectrum associated with the pion interpolator $\pi^\dagger$. The two-point correlator is defined as

$$C_{\rm 2pt}^{ss'}(t_{\rm sep}, P^z) = \langle \pi_{s'}(P^z, t_{\rm sep}) \pi_s^\dagger(P^z, 0) \rangle \ , \tag{4.69}$$

where $\pi_s^\dagger$ and $\pi_{s'}$ represent the pion source and sink defined from Eq. (4.1) as

$$\pi_s(\vec{P}, t_{\rm sep}) = \sum_{\vec{x}} \pi_s(\vec{x}, t_{\rm sep}) e^{-i\vec{P}\cdot\vec{x}} \ . \tag{4.70}$$

In this work, we use a Gaussian-smeared source and sink ($s = s' = S$), following the setup in Ref. [299]. The superscript SS will be removed in the following text for simplicity.

By inserting a complete set of states, the two-point correlator can be expressed as a sum over energy eigenstates:

$$C_{\rm 2pt}(t_{\rm sep}) = \sum_{n=0}^{N_{\rm s}-1} \frac{|z_n|^2}{2E_n} \left(e^{-E_n t_{\rm sep}} + e^{-E_n (L_t - t_{\rm sep})} \right) \ . \tag{4.71}$$

The overlap amplitude $z_n = \langle n|\pi^\dagger|\Omega\rangle$ quantifies the projection of the pion interpolator onto the nth energy eigenstate, and N_s denotes the number of excited states with the same quantum numbers as the pion, which are considered within the fit function.

The quasi-TMD beam function is extracted from the three-point correlator

$$C_{\tilde{h}}(t_{\rm sep}, \tau) = \left\langle \pi_s(\vec{P}, t_{\rm sep}) \bar{q}(z, b_\perp, \tau) \gamma^t q(\vec{0}, \tau) \pi_s^\dagger(\vec{P}, 0) \right\rangle \ , \tag{4.72}$$

which can be expressed in terms of a spectral decomposition:

$$C_{\tilde{h}}(t_{\rm sep}, \tau) = \sum_{n,m=0}^{N_s-1} \frac{z_n^\dagger O_{nm} z_m}{4E_n E_m} e^{-E_n(t_{\rm sep}-\tau)} e^{-E_m \tau} \ , \tag{4.73}$$

where $O_{nm} = \langle n|\bar{q}(z, b_\perp)\gamma^t q(0)|m\rangle$ is proportional to the matrix elements of the quasi-TMD operator. To extract the ground-state matrix element $O_{00} = 2E_0\tilde{h}_{\gamma^t}$, we employ a two-state chained fit using the two-point correlator $C_{\rm 2pt}$ and the ratio

$$R_{\tilde{h}}\left(t_{\rm sep}, \tau\right) = \frac{C_{\tilde{h}}\left(t_{\rm sep}, \tau\right)}{C_{\rm 2pt}\left(t_{\rm sep}\right)} \ . \tag{4.74}$$

In the limit $t_{\rm sep} \to \infty$, the ratio $R_{\tilde{h}}$ asymptotically converges to the bare quasi-TMD matrix element $\tilde{h}_{\gamma^t}$. The chained fit procedure first fits the two-point correlator $C_{\rm 2pt}$, then uses the posterior distributions of E_0, E_1, z_0, and z_1 as priors in the subsequent fit of $R_{\tilde{h}}$.

As discussed in Ref. [217], the absence of Wilson lines in the CG correlator eliminates linear divergences, allowing the renormalized operator to be defined as

$$\bar{q}_0(z, b_\perp)\Gamma q_0(0) = Z_q(a)\left[\bar{q}(z, b_\perp)\Gamma q(0)\right] \ , \tag{4.75}$$

where Z_q is the CG quark wave function renormalization factor. This renormalization is independent of both the external hadron states and the spatial separation of the quark bilinear operator. Consequently, an appropriate ratio can be constructed to cancel out the renormalization factor.

Following the approach in Ref. [228], we renormalize the quasi-TMD matrix elements using the ratio

$$\tilde{h}_{\gamma^t}(z, b_\perp, P^z; \mu) = \frac{\tilde{h}^0_{\gamma^t}(z, b_\perp, P^z; a)}{\tilde{\varphi}^0_{\gamma^t\gamma^5}(z=0, b_\perp, P^z=0; a)} \ , \tag{4.76}$$

and similarly for the quasi-TMDWF, it is renormalized as

$$\tilde{\varphi}_{\gamma^z\gamma^5}(z, b_\perp, P^z; \mu) = \frac{\tilde{\varphi}^0_{\gamma^z\gamma^5}(z, b_\perp, P^z; a)}{\tilde{\varphi}^0_{\gamma^t\gamma^5}(z = 0, b_\perp, P^z = 0; a)} \,. \tag{4.77}$$

Here, $\tilde{\varphi}^0$ denotes the bare quasi-TMDWF matrix elements defined in Eq. (4.57).

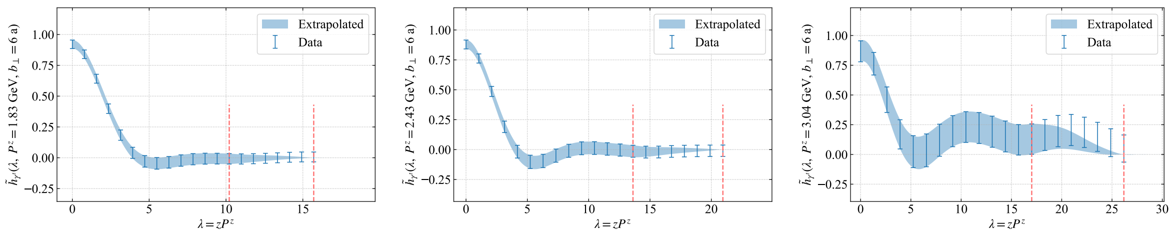

Figure 4.12: The extrapolation of the renormalized quasi-TMD in coordinate space for various hadron momenta is shown in three panels. From left to right, they correspond to hadron momenta $P^z = 1.83$, 2.43 and 3.04 GeV, respectively. The regions between the two red dashed lines indicate the extrapolation range. For different hadron momenta, the same starting point of $z = 0.78$ fm is chosen for the extrapolation.

The renormalized matrix elements in the coordinate space are presented in Fig. 4.12. As demonstrated in Ref. [188], the extracted quasi-distributions in momentum space within the moderate x region remain largely insensitive to the choice of extrapolation strategy. Therefore, we apply a non-fit extrapolation method to smooth the uncertainties of the renormalized quasi-TMD matrix elements at long distances. The extrapolation is performed using the following form:

$$\tilde{h}^{\text{ext}} = w \cdot \tilde{h} + (1 - w) \cdot 0 \,, \tag{4.78}$$

where w is a weight function that transitions linearly from 1 to 0 within the range $\lambda \in [z_{\text{ext}}P^z, z_{\text{max}}P^z]$, the $z_{\text{ext}} = 0.78$ fm is the starting point of extrapolation and $z_{\text{max}} = 1.2$ fm is the largest longitudinal separation of quasi-TMD. It is expected that $z > z_{\text{ext}}$ is large enough to see the exponential decay behavior of quasi-TMD.

4.4.2.1 The Collins-Soper Kernel

The CS kernel can be extracted from the ratio of quasi-TMDWFs with different momenta, as described in Eq. (4.60). For the matching procedure, we use the fixed-order one-loop results for the CG matching coefficient $C_{\rm TMD}$ and the corresponding hard kernel H_ϕ, as provided in Ref. [224]. Furthermore, we account for large logarithmic terms in the hard kernel by applying renormalization group evolution to improve the accuracy of the perturbative matching up to NLL.

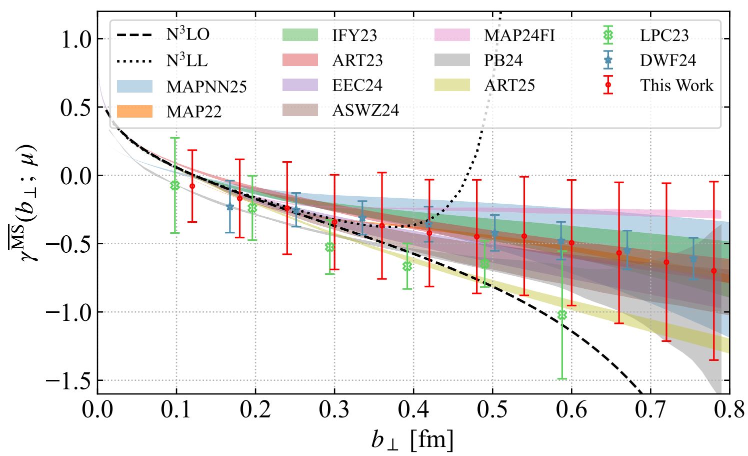

Figure 4.13: The CS kernel at $\mu = 2$ GeV in this work is presented using red points with error bars. The 3-loop perturbative results [342, 343] for the CS kernel are denoted as N^{3}LO and N^{3}LL in the figure. The CS kernels from recent phenomenological parameterizations of experimental data are shown from MAP22 [326], IFY23 [327], ART23 [330], MAP24FI [332], EEC24 [344], PB24 [345], MAPNN25 [346] and ART25 [333]. In addition, some recent lattice calculations are presented from LPC23 [222], ASWZ24 [226] and DWF24 [228].

After incorporating both statistical and systematic uncertainties, Fig. 4.13 presents our final results for the CS kernel at $\mu = 2$ GeV. In the small $b_\perp$ region, our results align well with the three-loop perturbative calculations from Refs. [342, 343], labeled N^{3}LO and N^{3}LL. Moreover, our calculation remains reliable in the large $b_\perp$ region, where perturbative methods break down.

To further contextualize our findings, we compare them with CS kernels extracted from recent global fits of experimental data, including MAP22 [326], IFY23 [327], ART23 [330], MAP24FI [332], EEC24 [344], PB24 [345], MAPNN25 [346] and ART25 [333]. Additionally, we include recent lattice QCD results from LPC23 [222], ASWZ24 [226], and DWF24 [228].

A particularly important observation is the agreement between our calculation and the DWF24 result from chirally symmetric domain-wall fermion configurations, both of which employ the CG framework despite using different lattice actions. Moreover, it is observable that the two most recent global analysis results (MAP24FI and ART25) exhibit a deviation from their prior results in different directions. Nevertheless, both remain consistent with this work due to the large uncertainty in our CS kernel, primarily stemming from power corrections at such a heavy pion mass. In our future study, the precision of our CS kernel could be refined by adopting a smaller valence pion mass.

4.4.2.2 Intrinsic Soft Function

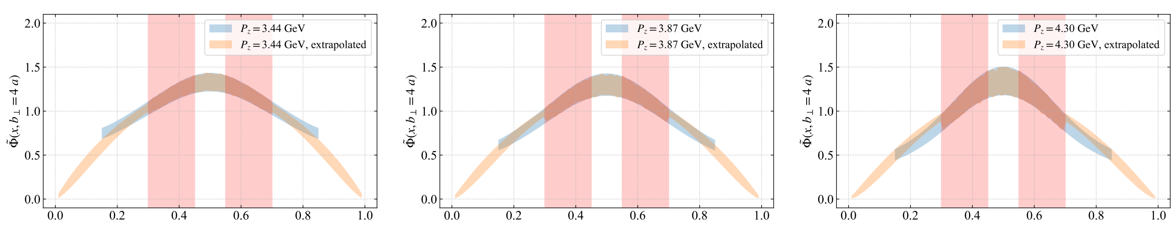

Figure 4.14: Extrapolations of reduced quasi-TMDWF using Eq. (4.79). From left to right, the three panels correspond to hadron momenta $P^z = 3.44$, 3.87 and 4.30 GeV, respectively. All of three cases are evolved to $P_{\text{FF}} = 2.58$ GeV using the rapidity evolution of quasi-TMDWF. The shaded red bands indicate the regions constitute the input for the extrapolation fit.

To extract the intrinsic soft function using Eq. (4.66), we integrate the quasi-TMDWF over x_1 and x_2. However, near the endpoint regions ($x \approx 0$ and $x \approx 1$), quasi-distributions suffer from significant power corrections. To mitigate this issue, we extrapolate the reduced quasi-TMDWF $\tilde{\Phi}$

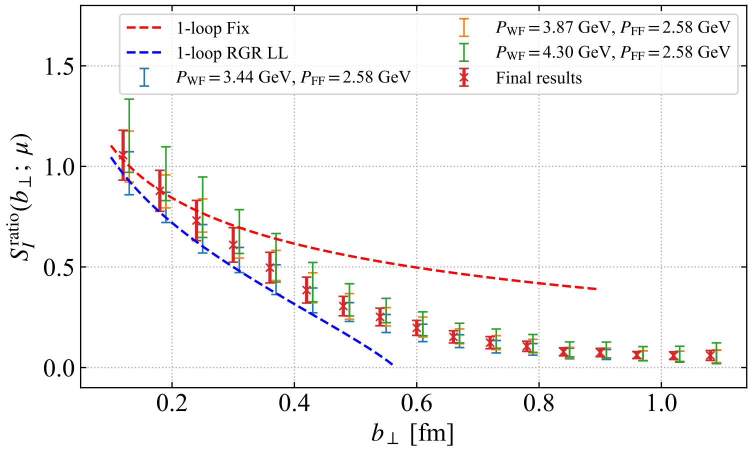

Figure 4.15: Intrinsic soft function in CG at $\mu = 2$ GeV in the ratio scheme, for different hadron momentum pairs of the quasi-TMDWF, $P_{\rm WF}$, and form factor, $P_{\rm FF}$. The final results are shown as red points with error bars (see text for details). The corresponding one-loop fixed-order perturbative results are the red dashed line, and the one-loop RG-resummed (RGR) perturbative results are the blue dashed line.

using the functional form

$$\tilde{\Phi}(x) = cx^n(1-x)^n \,, \tag{4.79}$$

where c and n are fitting parameters. A comparison of the reduced quasi-TMDWF before and after extrapolation is shown in Fig. 4.14, where the three panels correspond to hadron momenta $P^z = 3.44$, 3.87 and 4.30 GeV, respectively. All three cases are evolved to $P_{\rm FF} = 2.58$ GeV using the rapidity evolution of quasi-TMDWF. The red shaded bands indicate the regions that constitute the input for the extrapolation fit, with the selected fit range of $x \in [0.3, 0.45]$ and $x \in [0.55, 0.7]$, where the results of three different momenta show consistency. The extrapolation form derived from the fitting process is applied to the endpoint regions with $x < 0.3$ or $x > 0.7$. Fig. 4.14 shows that the extrapolated results outside the fit range are in alignment with the observed trends. Furthermore, since the Sudakov kernel diverges near the endpoints ($x \approx 0$) after the RG resummation, the integral in Eq. (4.66) is restricted to the range $x_1, x_2 \in [0.05, 0.95]$. This cutoff has a small impact on the results of the intrinsic soft function, since the integrand in the denominator of Eq. (4.66) converges near the endpoint regions.

To assess systematic effects of power corrections, we performed integration with three different momentum pairs, formed by the momentum of the form factor $P_{\rm FF} = 2.58$ GeV and three momenta of the quasi-TMDWF $P_{\rm WF} = 3.44$, 3.87 and 4.30 GeV. In order to bridge the gap between $P_{\rm FF}$ and $P_{\rm WF}$, the quasi-TMDWF is evolved to $P_{\rm FF}$ by solving the rapidity evolution equation. The extracted intrinsic soft function, according to Eq. (4.66), is shown in Fig. 4.15, where comparisons across different momentum pairs indicate that power corrections have only a minor impact. Note that the superscript "ratio" of $S_I^{\rm ratio}$ indicates the ratio scheme in the renormalization of the quasi-TMDWF.

For the final results of $S_I^{\rm ratio}$, the central values and statistical uncertainties are determined from a correlated average over the three momentum pairs formed by $P_{\rm FF}$ and $P_{\rm WF}$. The systematic uncertainties are determined by the spread of the mean values between these three momentum pairs, quantified as half the absolute difference between the maximum and minimum values. The extracted intrinsic soft function at $\mu = 2$ GeV, shown as red points with error bars in Fig. 4.15, is compared to perturbative predictions of fixed order and RG resummed (RGR) at leading logarithmic (LL) precision. In particular, at small $b_\perp$, our lattice results exhibit reasonable agreement with perturbative predictions, highlighting the robustness of our approach. In the large $b_\perp$ region, where perturbation theory becomes unreliable, our lattice results are expected to provide a more accurate description of the intrinsic soft function, offering valuable non-perturbative insights.

4.4.2.3 Pion TMDWF

By integrating the CG kernel, intrinsic soft function, and the renormalized quasi-TMDWF, the light-cone TMDWF can be extracted by employing the factorization formula found in Eq. (4.58).

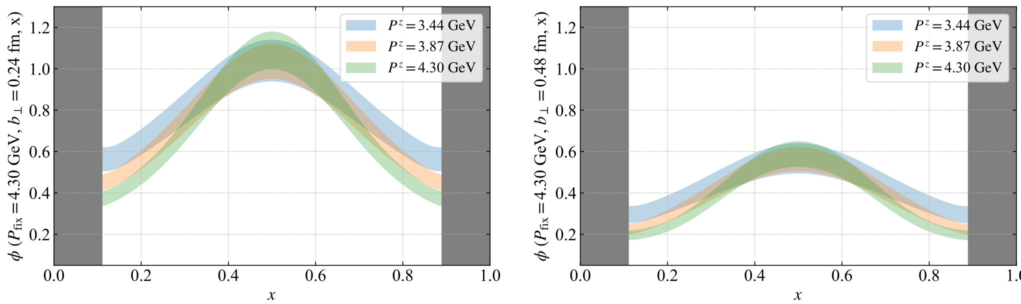

Figure 4.16: The light-cone pion TMDWF at $y_n = 0$, $P_{\rm fix} = 4.30$ GeV and $\mu = 2$ GeV of different hadron momenta, P^z, as the functions of momentum fraction, x, and for two transverse separations $b_\perp = 4a$ (upper panel) and $b_\perp = 8a$ (lower panel). The shaded gray bands ($x < 0.11$ and $x > 0.89$) indicate the endpoint regions where the estimated combined systematics are larger than 30%.

Choosing scales $y_n = 0$, $\zeta = (2xP_{\rm fix})^2$, and $\bar\zeta = (2\bar x P_{\rm fix})^2$, the factorization formula can be rewritten in a more explicit form:

$$\sqrt{S_I(b_\perp;\mu)} \cdot \tilde\phi_\Gamma(x, b_\perp, P^z;\mu) = \phi\left(x, b_\perp; \mu, \zeta, \bar\zeta\right) H_\phi(x, \bar x, P^z; \mu) \tag{4.80}$$

$$\times \exp\left[\ln\left(\frac{P^z}{P_{\rm fix}}\right)\gamma^{\overline{\rm MS}}(b_\perp;\mu)\right] + \mathcal{O}\left(\frac{\Lambda_{\rm QCD}}{xP^z}, \frac{1}{b_\perp(xP^z)}, \frac{\Lambda_{\rm QCD}}{\bar x P^z}, \frac{1}{b_\perp(\bar x P^z)}\right) \tag{4.81}$$

where the CS scale is evolved from $\zeta_0 = (2xP^z)^2$ and $\bar\zeta_0$ to $\zeta = (2xP_{\rm fix})^2$ and $\bar\zeta$ using the CS kernel extracted from quasi-TMDWFs.

The final results for the TMDWF at $y_n = 0$, $P_{\rm fix} = 4.30$ GeV, and $\mu = 2$ GeV are shown in Fig. 4.16 as a function of the momentum fraction x for three different hadron momenta. Two representative transverse separations, $b_\perp = 4a$ (upper panel) and $b_\perp = 8a$ (lower panel), are selected for illustration. As can be seen, the variation between different momenta remains mild in the moderate x region, demonstrating the validity of power expansion in large P^z within the quasi-TMD framework, where power corrections are small. The shaded gray bands ($x < 0.11$ and $x > 0.89$) indicate the endpoint regions where the estimated combined systematics are greater than 30%. The combined systematics include the variation of scales and power corrections, which are estimated by the variation of the initial scale μ_0 of the RGR procedure by a factor of $\sqrt{2}$ and the spread in the central values of the TMDWFs with different momenta, respectively.

4.4.2.4 Pion TMDPDF

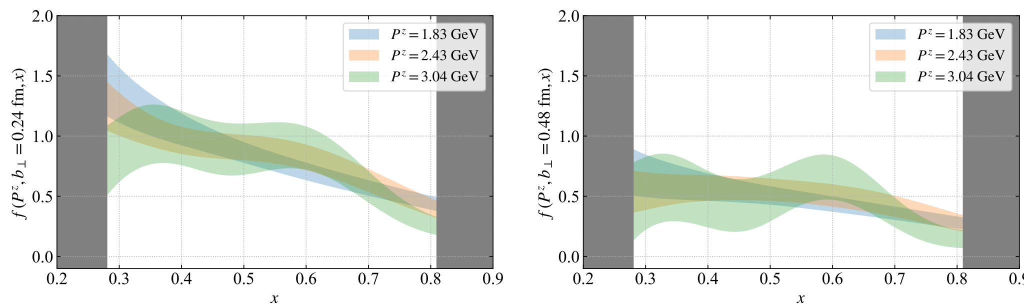

Figure 4.17: The unpolarized light-cone pion TMDPDF at $\zeta = \mu^2 = 4\ \mathrm{GeV}^2$ of different hadron momenta are shown as the function of momentum fraction, x, and for transverse separations, $b_\perp = 4a$ (upper panel) and $b_\perp = 8a$ (lower panel). The shaded gray bands ($x < 0.28$ and $x > 0.81$) indicate the endpoint regions where the estimated combined systematics are larger than 30%.

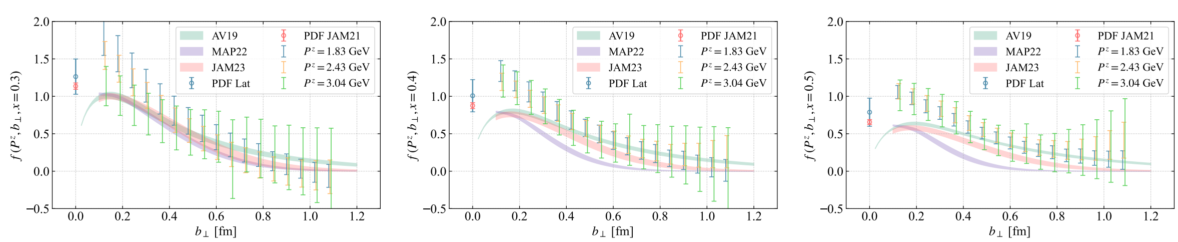

Figure 4.18: Unpolarized light-cone TMDPDF of pion at $x = 0.3$, $x = 0.4$ and $x = 0.5$ in $b_\perp$ space up to $b_\perp \gtrsim 1$ fm, and for different hadron momenta, P^z. The lattice results are illustrated alongside the phenomenological results from AV19 [334], MAP22 [335] and JAM23 [336]. In addition, the collinear PDF from lattice calculation in CG (PDF Lat) [217] and global fit (PDF JAM21) [347] are plotted for comparison.

The combination of the CG kernel, intrinsic soft function, and renormalized quasi-TMD beam functions enables the extraction of the light-cone TMDPDF using the factorization formula in Eq. (4.14). The final results for the TMDPDF at $\zeta = \mu^2 = 4\ \mathrm{GeV}^2$ are shown in Fig. 4.17 as a function of the momentum fraction x for three different hadron momenta. Two representative transverse separations, $b_\perp = 4a$ (upper panel) and $b_\perp = 8a$ (lower panel), are selected for illustration. As can be seen, the variation between different momenta remains mild in the moderate x region, demonstrating the validity of power expansion in large P^z within the quasi-TMD framework, where power corrections are small. The shaded gray bands ($x < 0.28$ and $x > 0.81$) indicate the endpoint regions where the estimated combined systematics are greater than 30%.

In Fig. 4.18, we examine the $b_\perp$ dependence of the TMDPDF at $x = 0.3$, $x = 0.4$, and $x = 0.5$. The results for different hadron momenta exhibit good consistency, which is improved as the x value approaches $x = 0.5$, thereby confirming our expectation of the behavior of power corrections. For comparison, we include recent global fits of the pion TMDPDF from AV19 [334], MAP22 [335] and JAM23 [336]. Although slight deviations in magnitude are observed, the lattice results exhibit a qualitative agreement with global fits, especially at the smaller x values, where the experimental data give a better constraint.

Since the collinear PDF and the $b_\perp \to 0$ limit of the unpolarized TMDPDF differ only by a perturbative expansion in $b_\perp$ [330], we also compare our results to collinear PDFs. Specifically, we include the collinear PDF extracted from lattice calculations in CG (PDF Lat) [217] and the global fit result from JAM21 (PDF JAM21) [347]. As shown in the plots, the collinear PDFs closely match the TMDPDF in the small $b_\perp$ region, supporting the expected theoretical behavior.

An important advantage of the CG method is the absence of linear divergence, which allows us to probe large $b_\perp$ regions with a reasonable SNR. At $b_\perp \gtrsim 1$ fm, the TMDPDF smoothly decays to values consistent with zero, facilitating a feasible Fourier transform into the $k_\perp$ space. In this study, we combined our lattice result for the collinear pion PDF (PDF Lat) with the pion TMDPDF to interpolate the distribution at short distances around $b_\perp \approx a$, and subsequently applied a Gaussian

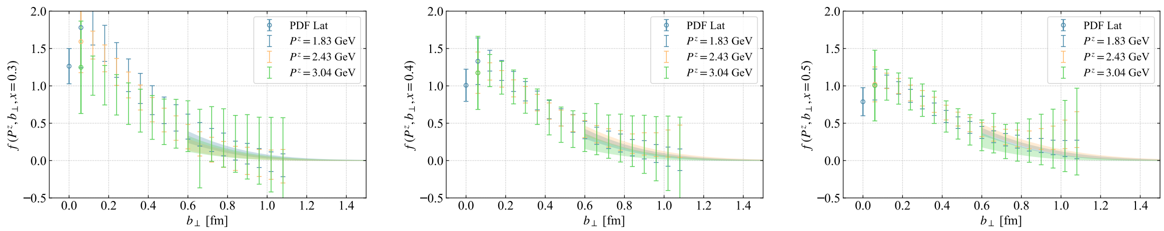

Figure 4.19: Unpolarized light-cone TMDPDF of pion at $x = 0.3$, $x = 0.4$ and $x = 0.5$ in $b_\perp$ space after extrapolation to $b_\perp = 3$ fm. The collinear PDF from lattice calculation in CG (PDF Lat) [217] is plotted at $b_\perp = 0$ fm. The point at $b_\perp = 0.06$ fm is determined by cubic interpolation.

form to extrapolate the large $b_\perp$ regime up to 3 fm, the extrapolation form is

$$f(b_\perp) = Ae^{-m \cdot b_\perp^2} , \tag{4.82}$$

where A and m are fit parameters. This extrapolation form is inspired by the confinement in the transverse plane, while the model dependence of the results will be investigated in future works with improved data precision. The extrapolated unpolarized light-cone TMDPDF of three momenta are plotted in Fig. 4.19.

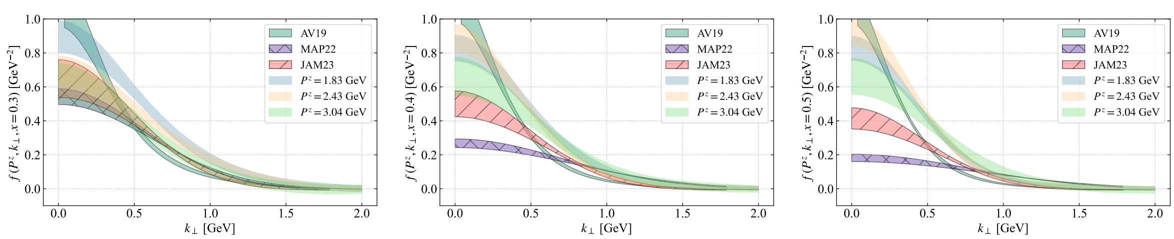

Figure 4.20: Unpolarized light-cone TMDPDF of pion at $x = 0.3$, $x = 0.4$ and $x = 0.5$ in $k_\perp$ space. The lattice results are illustrated alongside the phenomenological results from AV19 [334], MAP22 [335] and JAM23 [336].

After extrapolation, we performed a discretized Fourier transform, with the results presented in Fig. 4.20. Although the extrapolation in $b_\perp$ introduces minor model dependencies, the resulting distributions exhibit qualitative consistency with global fits, especially at the smaller x values, where the experimental data give a better constraint. This demonstrates the potential of lattice QCD to explore the transverse momentum structure of hadrons from first principles. In future studies with higher precision, the extrapolation modeling uncertainties, as well as other systematic effects, will be investigated more comprehensively to further improve the robustness of such calculations.

4.4.3 Conclusion

In this work, we have presented the first lattice QCD calculation of the pion unpolarized valence-quark TMDPDF within the framework of LaMET. The calculation was performed on a 2+1 flavor QCD ensemble with a lattice spacing of $a = 0.06$ fm and a pion mass of 300 MeV. By leveraging high statistics, off-axis momenta, a well-tuned boosted Gaussian smeared pion source, and the novel CG approach, we attained significant hadron momenta reaching 3 GeV. We computed the quasi-TMD beam function, extracted the CS kernel from quasi-TMDWF, and derived the intrinsic soft function. To enhance perturbative accuracy, we performed RGR of the Sudakov kernel at NLL order throughout our analysis. These advancements collectively enabled the determination of both the pion TMDWF and TMDPDF from the first principles.

Our results of the CS kernel and intrinsic soft function demonstrate consistency with perturbative calculations in the small-$b_\perp$ region, where perturbative methods remain valid. Additionally, the extracted CS kernel agrees with both previous lattice QCD studies and phenomenological fits of the experimental data.

By combining the renormalized quasi-TMDPDF, the CS kernel, and the intrinsic soft function, we obtained the pion valence TMDPDF across a range of $b_\perp$. Our results extend beyond $b_\perp \gtrsim 1$ fm, allowing for a feasible Fourier transform into momentum space. The extracted TMDPDFs in the $b_\perp$

and $k_\perp$ space exhibit qualitative agreement with phenomenological fits. This agreement demonstrates the potential of lattice QCD to validate and complement global analyses.

The outcome of this study highlights the efficacy of the CG quasi-TMD approach in probing the transverse momentum structure of hadrons. Our results provide a valuable first-principles complement to global fits, especially in the non-perturbative region where experimental data remain scarce. Future improvements, such as the incorporation of finer lattice spacings, larger momenta, and improved control over systematic uncertainties, will further enhance the precision of lattice QCD determinations of TMDPDFs. Additionally, extending this methodology to other hadrons, including the nucleon, will provide deeper insights into the partonic structure of QCD bound states.

4.5 Calculation of Nucleon TMDs

In contrast to the pion, which serves as the lightest QCD bound state and a key probe of chiral dynamics, the nucleon represents a far more intricate composite system. As a spin-$1/2$ baryon with a rich flavor and spin structure, the nucleon encodes essential information about how quark and gluon degrees of freedom generate the mass, spin, and spatial structure of visible matter. While pion TMDs offer a comparatively clean environment to investigate transverse-momentum dynamics in a relativistic bound state, nucleon TMDs introduce additional layers of complexity associated with spin correlations, flavor asymmetries, and multi-parton interactions.

The TMDs of the nucleon therefore provide a multidimensional characterization of its internal structure, extending beyond the one-dimensional collinear picture to include correlations between transverse momentum and spin. In particular, spin-dependent TMDs encode nontrivial QCD dynamics such as spin-orbit correlations and time-reversal-odd effects, which have no direct analogue in the spinless pion. Consequently, nucleon TMDs play a central role in interpreting semi-inclusive deep-inelastic scattering and Drell–Yan measurements, where polarized observables probe subtle aspects of QCD dynamics [7, 10, 51, 348].

Despite the extensive experimental program devoted to nucleon TMDs, their nonperturbative structure remains incompletely constrained. The Collins–Soper kernel governing rapidity evolution has been computed perturbatively to high orders [349, 350], but the intrinsic transverse-momentum dependence at hadronic scales must be modeled or extracted from global fits. Although phenomenological analyses incorporate large datasets from low-p_T SIDIS and Drell–Yan processes [319, 322, 325, 326, 330], noticeable discrepancies among different extractions indicate substantial systematic uncertainties. Compared with the pion case, where experimental information is scarce, the challenge for the nucleon lies less in data availability and more in disentangling correlated spin, flavor, and kinematic effects within global analyses.

In our 2022 study [204], we computed the quasi-TMDPDF matrix elements of a boosted nucleon within the GI framework. A systematic renormalization program was carried out by combining Wilson-loop subtractions with short-distance hadron matrix elements [208]. The perturbative matching to light-cone TMDPDFs was performed including both one- and two-loop contributions, together with renormalization-group resummation of large logarithms. Our results constitute the first first-principles determination of the flavor isovector nucleon TMD momentum distributions in this framework. The extracted TMDPDFs show qualitative agreement with phenomenological extractions, demonstrating the robustness of the CG approach and its potential for a unified lattice description of pion and nucleon TMD structures.

4.5.1 Theoretical Framework

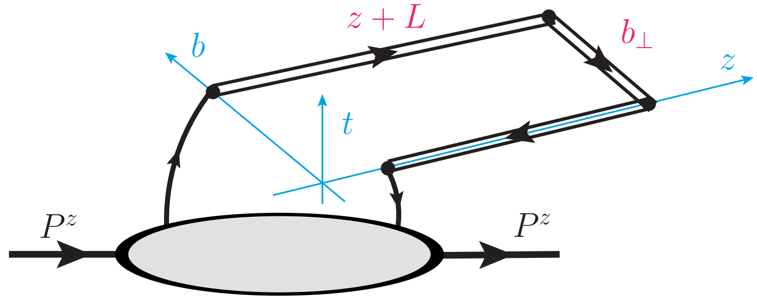

Figure 4.21: Illustration of the quasi TMDPDFs in GI formalism.

According to LaMET, the correlations with modes traveling along the light-cone can be extracted from distributions in a fast-moving nucleon through large-momentum expansion. On the lattice, the equal-time quasi TMDPDFs are constructed as

$$\tilde{f}_\Gamma(x, b_\perp, P^z, \mu) \equiv \lim_{\substack{a\to 0\\ L\to\infty}} \int \frac{dz}{2\pi} e^{-iz(xP^z)} \frac{\tilde{h}^0_\Gamma(z, b_\perp, P^z, a, L)}{\sqrt{Z_E(2L+z, b_\perp, a)} Z_O(1/a, \mu, \Gamma)}, \tag{4.83}$$

where a denotes the lattice spacing. $\Gamma = \gamma^t$ or γ^z is the Dirac matrix that can be projected onto γ^+ in the large momentum limit. The $\tilde{h}^0_\Gamma(z, b_\perp, P^z, a, L)$ is built with a gauge-invariant nonlocal quark bilinear operator as

$$\tilde{h}^0_\Gamma(z, b_\perp, P^z, a, L) = \langle P^z | \tilde{O}^0_{\Gamma,\sqsubset}(z, b_\perp, P^z; L) | P^z \rangle, \tag{4.84}$$

$$\tilde{O}^0_{\Gamma,\sqsubset}(z, b_\perp, L) \equiv \bar{\psi}(b_\perp \hat{n}_\perp) \Gamma U_{\sqsubset,L}(b_\perp \hat{n}_\perp, z\hat{n}_z) \psi(z\hat{n}_z). \tag{4.85}$$

In the above, $|P^z\rangle$ denotes the unpolarized nucleon state and L denotes the “infinity” that the gauge link can reach. The staple-shaped Wilson link is chosen as

$$U_{\sqsubset,L}\left(b_\perp\hat{n}_\perp, z\hat{n}_z\right) \equiv U_z^\dagger\left((z+L)\hat{n}_z + b_\perp\hat{n}_\perp, b_\perp\hat{n}_\perp\right) \tag{4.86}$$

$$\times U_\perp\left((z+L)\hat{n}_z + b_\perp\hat{n}_\perp, (z+L)\hat{n}_z\right) U_z\left((z+L)\hat{n}_z, z\hat{n}_z\right), \tag{4.87}$$

in which U_z is the path-ordered Euclidean gauge link along the z-direction, and $U_\perp$ is along the transverse direction at the "infinity" position $(z+L)$ on the finite lattice. Their explicit forms are:

$$U_z(\xi_1^z\hat{n}_z + \xi_\perp\hat{n}_\perp, \xi_2^z\hat{n}_z + \xi_\perp\hat{n}_\perp) = \mathcal{P}\exp\left[-ig\int_{\xi_1^z}^{\xi_2^z} d\lambda\hat{n}_z\cdot A\left(\lambda\hat{n}_z + \xi_\perp\hat{n}_\perp\right)\right], \tag{4.88}$$

$$U_\perp(\xi^z\hat{n}_z + \xi_{\perp 1}\hat{n}_\perp, \xi^z\hat{n}_z + \xi_{\perp 2}\hat{n}_\perp) = \mathcal{P}\exp\left[-ig\int_{\xi_{\perp 1}}^{\xi_{\perp 2}} d\lambda\hat{n}_\perp\cdot A\left(\xi^z\hat{n}_z + \lambda\hat{n}_\perp\right)\right]. \tag{4.89}$$

An illustration of the quasi TMDPDFs is given in Fig. 4.21, in which the staple-shaped Wilson link is depicted as double lines.

The bare matrix elements of quasi TMDPDFs contain linear divergence, pinch-pole singularity and logarithmic divergence. Both the linear divergence and pinch-pole singularity can be renormalized by the square root of Wilson loop $\sqrt{Z_E(2L+z, b_\perp, a)}$ [123, 153, 169, 351, 352]; and the logarithmic divergence can be renormalized by a factor $Z_O(1/a, \mu)$, which is extracted from matching between the lattice and perturbative calculation of zero-momentum matrix elements in the perturbative region [208, 257, 353, 354].

The Wilson loop $Z_E(r = 2L+z, b_\perp, a)$ is defined as the vacuum expectation of a rectangular shaped space-like gauge links with size $r\times b_\perp$, which can be written as

$$Z_E(r, b_\perp, a) = \langle U_z^\dagger\left((z+L)\hat{n}_z + b_\perp\hat{n}_\perp, (-z)\hat{n}_z + b_\perp\hat{n}_\perp\right) \tag{4.90}$$

$$\times U_\perp\left((z+L)\hat{n}_z + b\hat{n}_\perp, (z+L)\hat{n}_z\right) U_z\left((z+L)\hat{n}_z, (-z)\hat{n}_z\right) \tag{4.91}$$

$$\times\, U_\perp^\dagger\left((-z)\hat{n}_z + b_\perp\hat{n}_\perp, (-z)\hat{n}_z\right)\Big\rangle . \tag{4.92}$$

The Wilson loop is introduced to eliminate the linear divergence of the form $e^{-\delta\bar{m}r}$, which comes from the self-energy corrections to the gauge link [123, 354], as well as the pinch-pole singularity, which is related to the heavy quark effective potential term $e^{-V(b_\perp)L}$ describing the interactions between the two Wilson lines along the z direction in the staple link [281].

The logarithmic divergence factor Z_O can be extracted from the zero-momentum bare matrix elements $\tilde{h}^0_\Gamma(z, b_\perp, 0, a, L)$. In order to keep the renormalized matrix elements consistent with perturbation theory, Z_O should be determined from the condition:

$$Z_O\left(\frac{1}{a}, \mu, \Gamma\right) = \lim_{L\to\infty} \frac{\tilde{h}^0_\Gamma(z, b_\perp, 0, a, L)}{\sqrt{Z_E(2L+z, b_\perp, a)}\tilde{h}^{\overline{\text{MS}}}_\Gamma(z, b_\perp, \mu)} \tag{4.93}$$

in a specific window where $z \ll \Lambda^{-1}_{\text{QCD}}$ so that perturbation theory is valid.

4.5.2 Numerical Results

We use the valence tadpole improved clover fermion on the hypercubic (HYP) smeared [293] $2+1+1$ flavors MILC configurations with highly improved staggered quark (HISQ) sea and 1-loop Symanzik improved gauge action [285]. We analyze a single ensemble with lattice spacing $a = 0.12$ fm and volume $n_s^3 \times n_t = 48^3 \times 64$ using physical sea quark masses, and two choices of light valence quark mass corresponding to $m_\pi^{\text{val}} = \{220,\ 310\}$ MeV. The HYP smearing is also used for nonlocal correlation functions to improve the statistical signal. In order to explore the momentum dependence, we employ three different nucleon momenta as $P^z = 2\pi/(n_s a) \times \{8, 10, 12\} =$

$\{1.72, 2.15, 2.58\}$ GeV.

We adopt momentum-smearing point source [340] at several time slices, and average correlation functions for both the forward and backward directions in z and transverse space of the gauge link. In total, there are 1000 (configurations) $\times 16$ (source time slices) $\times 4$ (forward/backward directions of the z and transverse axes) measurements for the $m_\pi^{\text{val}} = 220$ MeV case and $1000 \times 4 \times 4$ measurements for the 310 MeV case.

The two-point correlation functions of the nucleon are defined as

$$C_2(t, \vec{P}) = \Big\langle \sum_{\vec{y}} e^{i\vec{P}\cdot\vec{y}} T_{\alpha\beta} N_\alpha(\vec{y}, t) \bar{N}_\beta(\vec{0}, 0) \Big\rangle, \tag{4.94}$$

where $T = (1 + \gamma^t)/2$ denotes the unpolarized projector, and the nucleon interpolating operator N is defined as Eq. (4.2). For the three momenta $P^z = 1.72, 2.15, 2.52$ GeV, the statistics of the two-point correlation functions are 1000 (configurations) $\times 16$ (source time slices) for $m_\pi = 220$ MeV and 1000×4 for $m_\pi = 310$ MeV; while for the cases with a momentum smaller than 1.72 GeV, there are 1000×4 measurements for $m_\pi = 220$ MeV and 1000×1 measurements for $m_\pi = 310$ MeV.

To extract the quasi TMDPDFs, one constructs the three-point functions

$$C_3^\Gamma(t, t_s) = \Big\langle \sum_{\vec{y}} e^{i\vec{P}\cdot\vec{y}} T N(\vec{y}, t_s) \sum_{\vec{x}} \tilde{O}^\Gamma_{\text{TMD}}(\vec{x}, t)\, \bar{N}(\vec{0}, 0) \Big\rangle. \tag{4.95}$$

We adopt the sequential source method [355] to reduce the number of propagators in three-point functions, t_s denotes the time position of the sequential source. The operator $\tilde{O}^\Gamma_{\text{TMD}}(\vec{x}, t)$ is short for the TMD nonlocal quark bilinear operator $\tilde{O}^0_\sqsubset(\vec{x} + z\hat{n}_z, \vec{x} + b_\perp \hat{n}_\perp, \Gamma, L)$ at discrete time slice $t \in [0, t_s]$. The three-momentum is chosen as $\vec{P} = (0, 0, P^z)$.

After inserting the single particle intermediate states, one can parametrize the ratio of three- and two-point functions as

$$\frac{C_3^{\Gamma}\left(t, t_s\right)}{C_2(t_s)}=\frac{\tilde{h}_{\Gamma}^{0}+c_2(e^{-\Delta E t}+e^{-\Delta E(t_s-t)})+c_3 e^{-\Delta E t_s}}{1+c_1 e^{-\Delta E t_s}}, \tag{4.96}$$

in which $\tilde{h}_{\Gamma}^{0} \equiv \tilde{h}^{0}\left(z, b_{\perp}, P^z, \Gamma\right)$, ΔE is the mass gap between the ground-state and excited state, and $c_{1,2,3}$ are parameters for the excited-state contaminations. Combining the parametrization form of 2pt functions: $C_2(t)=c_0 e^{-E_0 t}\left(1+c_1 e^{-\Delta E t}\right)$, one can extract the values of $\tilde{h}_{\Gamma}^{0}$ at fixed $(z, b_{\perp}, P^z)$ through a correlated joint fit.

In our calculation, we use bootstrap resampling to establish correlations among all datasets, and the correlations are maintained consistently throughout the entire analysis. For the three-point functions, we have excluded the contact points ($t=0$ and $t=t_{\text{seq}}$) for the five values of the source-sink separation time range $t_{\text{seq}} \in (0.48 \sim 0.84)$ fm. For small $b_{\perp}$ ($b_{\perp} \leq 3a$), we have further removed two points near the source and sink($t=1$ and $t=t_{\text{seq}}-1$).

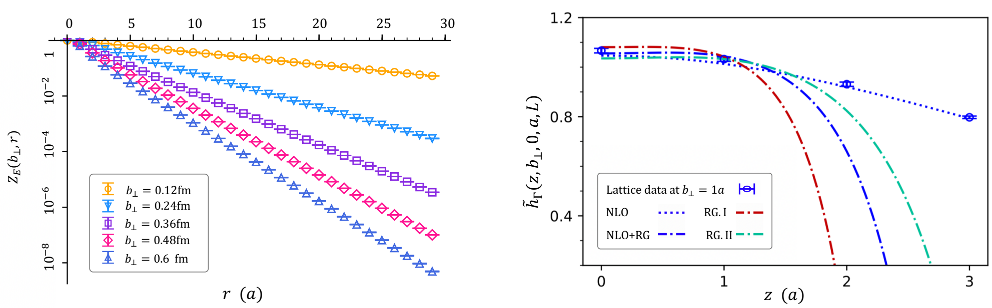

Figure 4.22: Upper panel: Wilson loop with $r = 2L + z$. Lower panel: renormalized zero-momentum matrix element $\tilde{h}_{\Gamma}\left(z, b_{\perp}, 0, a, L\right)$ at $b_{\perp} = 1a$, compare with 1-loop results before (NLO) and after (NLO+RG) RG evolution to $\mu = 2$ GeV in the $\overline{\text{MS}}$ scheme. Besides, we vary the starting point of evolution from $0.8\mu_0$ ("RG. I") to $1.2\mu_0$ ("RG. II") to estimate the higher-order correction effects.

As mentioned above, we use the square root of the Wilson loop $\sqrt{Z_E}$ and logarithmic divergence factor Z_O to renormalize the bare quasi-TMD matrix elements. In practice, the signal-to-noise ratio of $Z_E(r, b_\perp, a)$ defined in Eq. (4.92) decreases fast with r and $b_\perp$ such that it is hardly available for large r or $b_\perp$. To address this, we fit the effective energies of the Wilson loop, which gives the QCD static potentials, and then extrapolate them to large r and/or $b_\perp$, as in Ref. [208]. Numerical results for the Wilson loop are shown in the upper panel of Fig. 4.22.

The logarithmic divergence factor Z_O can be extracted from the matching between lattice and perturbative results in the region where both two theories work well. To preserve a good convergence of the perturbation theory before and after RG evolution, we choose the region where $b_\perp = 1a$, and $z = 0$ or $1a$, where both perturbation theory and lattice calculations work. The extracted Z_O values at these two points are $1.0734(93)$ and $1.0509(92)$, respectively. Averaging these data points yields an aggregated result of $Z_O = 1.0622(87)$.

With Eq. (4.93), it can be concluded that after dividing the bare matrix elements $\tilde{h}^0_\Gamma$ by $\sqrt{Z_E}$ and Z_O, the renormalized matrix elements should approximately be equal to the RG evolved perturbation results $\tilde{h}^{\overline{\mathrm{MS}}}_\Gamma$. The lower panel of Fig. 4.22 shows the consistency at points where we extract the Z_O factor, which serves as a check of the numerical result for Z_O. The points at $z = 0$ and $z = 1$ depicted in Fig. 4.22 demonstrate that, following renormalization with the extracted factor $Z_O = 1.0622(87)$, the renormalized matrix elements on the lattice show a consistent trend with the perturbative results obtained in $\overline{\mathrm{MS}}$ scheme at short distance. To verify the accuracy of scale evolution from the intrinsic physical scale in lattice calculation to $\overline{\mathrm{MS}}$ scale $\mu = 2$GeV, we vary the origin of the evolution from $0.8\mu_0$ (shown as "RG. I") to $1.2\mu_0$ (shown as "RG. II"), which is sensitive to higher-order corrections. One can see that the perturbative results from different μ_0 are consistent with the lattice data in the matching region ($b_\perp = 1a$, $z = 0$ or $1a$).

For a well-defined quasi TMDPDF, the length of the Wilson link L should be large enough to ensure that the final results are independent of L. In the definition of the staple-shaped Wilson link in Eq .(4.87), L should be extended to infinity. The L-dependence is included in the linear divergence of the Wilson link self-energy and pinch-pole singularity, which can be subtracted by the square root of the Wilson loop. Therefore, the subtracted quasi TMD matrix elements

$$\tilde{h}_{\Gamma}(z, b_{\perp}, a, L) \equiv \frac{\tilde{h}^{0}_{\Gamma}(z, b_{\perp}, P^{z}, a, L)}{\sqrt{Z_E(2L + z, b_{\perp}, a)} Z_O(1/a, \mu, \Gamma)} \tag{4.97}$$

will saturate to a plateau at large enough L to ensure that the link can extend outside the region of the parton, and exhibit independence of L.

Ref. [208] explored the L-dependence of renormalized TMD matrix elements and reached the conclusion that setting $L = 6a$ for MILC12 is sufficient, aligning well with the ensemble we have employed. To verify this, we further examine the L-dependence of the subtracted quasi TMD matrix elements, illustrated in Fig. 4.23.

In Fig. 4.23, the three panels on the left show the real parts of the ratios defined in Eq. (4.96) divided by the square root of the Wilson loop with various combinations of z and $b_{\perp}$ and $L = \{6, 8, 10\}a$. The right panel gives the fit results of subtracted quasi TMD matrix elements with different L. From the comparison, one can see that the fitted results for $L = \{6, 8, 10\}a$ are consistent with each other.

There is a phenomenological explanation for the earlier saturation of the Wilson line; however, it is essential to note that a definitive proof is currently unavailable. When calculating the TMDPDF of a proton, which reflects the quark correlation within the proton, a reasonable estimate for the correlation length could be the size of a proton, approximately 1 fm. Beyond this distance, quarks and gluons may escape the proton, potentially diminishing their impact. This suggests a typical saturation length of around 1 fm, with $L = 6a$ (or more precisely, $L = 8a$) being in proximity to this length scale.

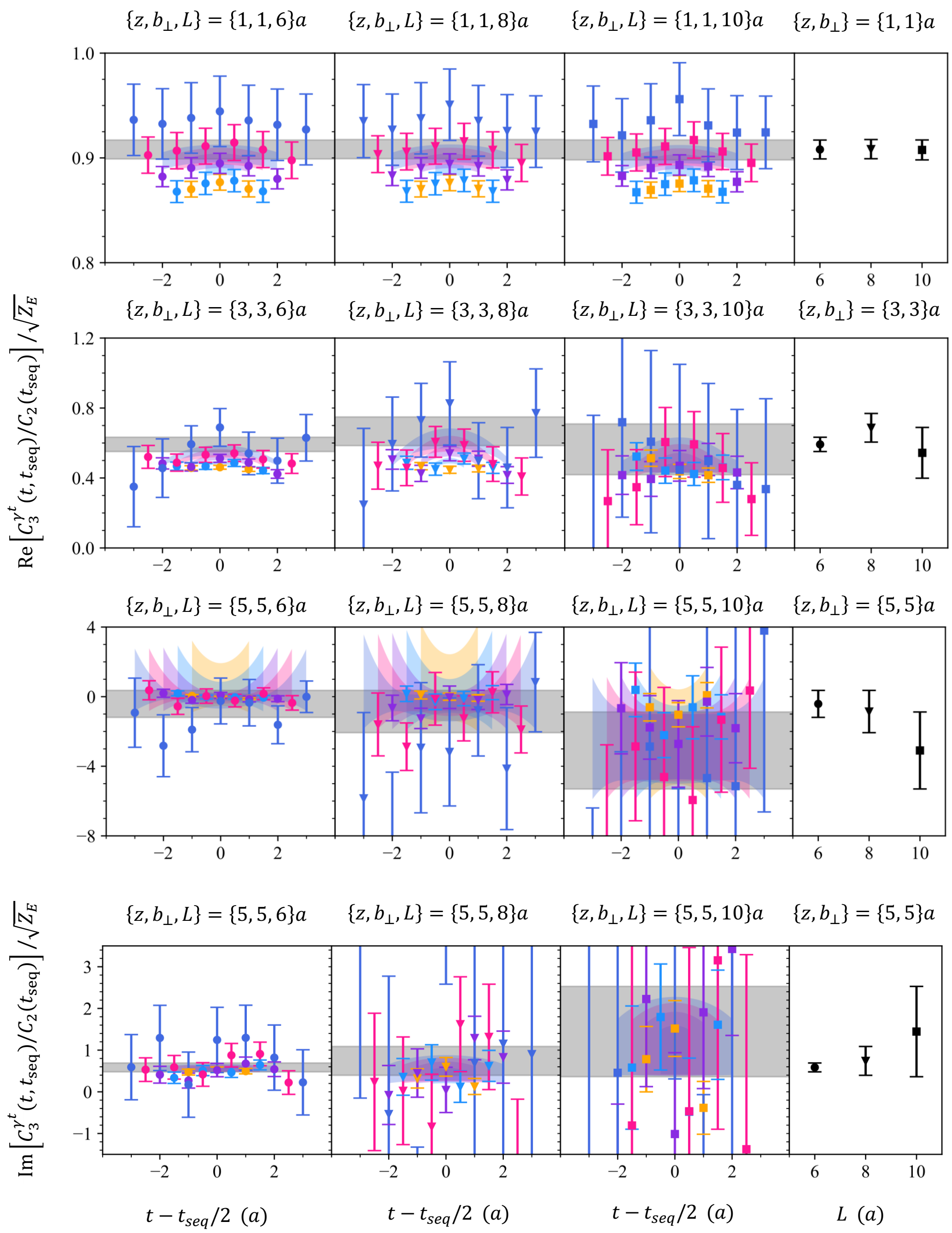

Figure 4.23: The ratios defined in Eq. (4.96) divided by the Wilson loop (three left panels) and the L dependence of the ground state results after fitting (right panels) at $\{z, b_\perp\} = \{1, 1\}, \{3, 3\}$ and $\{5, 5\}a$, with $P^z = 1.72$ GeV, $m_\pi = 220$ MeV and $\Gamma = \gamma^t$. Since the real parts for $\{z, b_\perp\} = \{5, 5\}a$ are close to zero, we also give the imaginary parts in this example.

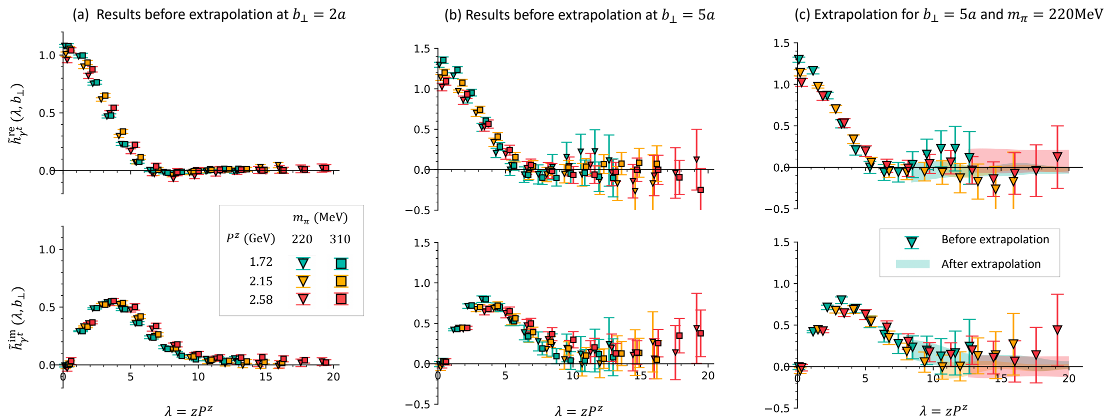

Figure 4.24: (a, b): Renormalized quasi TMDPDFs in coordinate space as function of $\lambda = zP^z$ with various m_π and P^z at $b_\perp = 2a$ and $5a$; (c): Comparison of original (data points) and extrapolated results (colored bands) at $b_\perp = 5a$.

Combining the bare quasi TMDPDFs matrix elements with the corresponding Wilson loop and renormalization factor, we obtain numerical results for renormalized matrix elements at different $\lambda = zP^z$. The two panels on the left in Fig. 4.24 exhibit the λ dependence of the renormalized matrix elements with $b_\perp = 2a$ and $5a$ with various P^z and m_π. Similar to the previous calculations, we apply an asymptotic extrapolation in λ to address the large λ behavior of the quasi TMDPDFs, using the form given in Eq. (4.4). It should be noticed that this asymptotic form was initially proposed for the one-dimensional parton distribution functions [169]. For TMDPDFs, whether this form holds and whether the involved parameters depend on the transverse separation remain to be found out.

After Fourier transforming the renormalized quasi TMDPDFs in coordinate space to momentum space, one can obtain the quasi TMDPDF $\tilde{f}_\Gamma(x, b_\perp, \zeta_z, \mu)$, shown as the dashdotted (magenta) line in Fig. 4.25. Combing the updated results of Collins-Soper kernel [356] and intrinsic soft function [222] calculated on the same ensembles, we obtain the TMDPDFs through the matching formula in Eq. (4.14). Fig. 4.25 shows an example of matched TMDPDF from the NLO [70,281] and NNLO [220, 283] kernel with RG running to scale $\mu = \sqrt{\zeta} = 2$ GeV. One can see that the results agree within

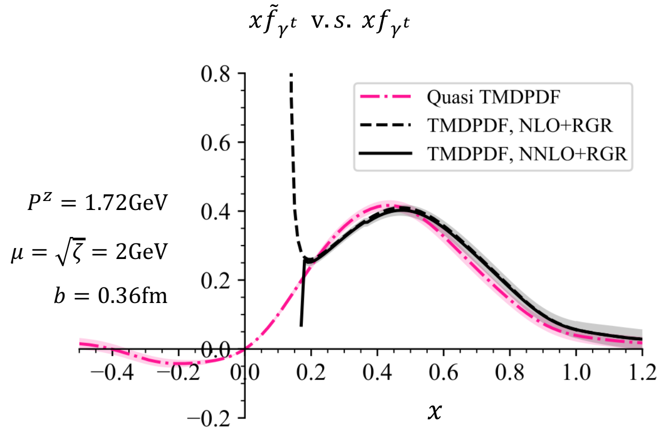

Figure 4.25: Quasi TMDPDF (dashed-red line) with nucleon boosted momentum $P^z = 1.72$ GeV and matched TMDPDF from NLO kernel (dashed-black line, no error) and NNLO kernel (solid-black line) with RG running to scale $\mu = \sqrt{\zeta} = 2$ GeV at $b_\perp = 3a$. Only statistical errors are included in the bands. Deviations between NLO and NNLO results at $x < 0.2$ indicate that the perturbative matching fails in the small-x region.

uncertainties, except for the endpoint region.

It should be noted that the TMDPDFs can be obtained after employing the factorization formula, while residual dependence on the lattice inputs (such as P^z and m_π) may still reside in the obtained results. To diminish the dependence, we extrapolate the above results to the physical m_π value (135 MeV) and infinite momentum using the following ansatz:

$$f(m_\pi, P^z) = f_{\text{phy}}\left[1 + d_0\left(m_\pi^2 - m_{\pi,\text{phy}}^2\right) + d_0' \ln(m_\pi^2/m_{\pi,\text{phy}}^2) + \frac{d_1}{(P^z)^2} + \frac{d_1'}{P^z}\right], \tag{4.98}$$

where the $d_0^{(\prime)}$ term characterizes the pion mass dependence, and $d_1^{(\prime)}$ accounts for the momentum-dependent discretization error. An interesting analysis has recently derived the chiral logarithms for quasi-PDFs [357]. In this approach, one first performs an operator product expansion and constructs the moments of quasi PDFs with derivative operators. Using chiral perturbation theory, one can find hadron-level operators which have the same symmetry. Calculating the one-loop perturbative contributions and summing these contributions for moment operators will lead to the chiral logarithms in quasi-PDFs. However, this analysis has not been generalized to TMDPDFs at present. In addition, since TMDPDFs are three-dimensional distributions, expanding the nonlocal operators will require the derivatives not only in the z direction, but also in the transverse direction. This will likely introduce additional complexities.

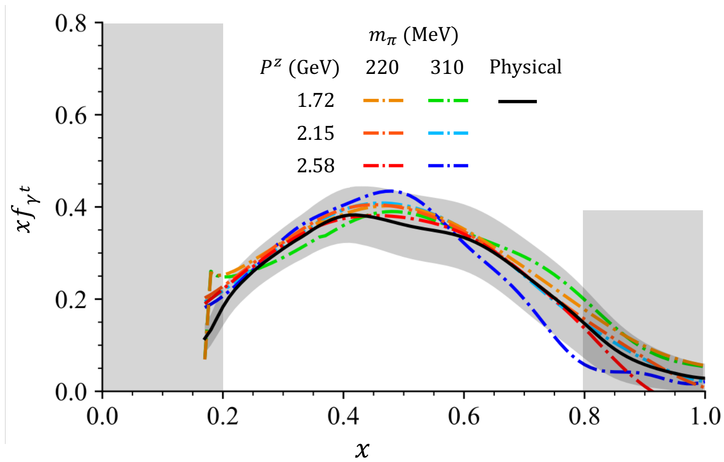

Figure 4.26: TMDPDFs obtained from different m_π and P^z (dashed-color lines) at $b_\perp = 3a$, and their physical results after extrapolation. The shaded grey band indicates the endpoint regions where LaMET predictions are not reliable. Only statistical errors of physical results are exhibited.

Taking $d_0' = d_1' = 0$ as our default case, we obtain the physical TMDPDF from a joint fit of $f(m_\pi, P^z)$ shown in Fig. 4.26. In order to explore the systematic bias in the extrapolation ansatz, we also estimate the m_π dependence by using the d_0' term (with $d_0 = 0$) and examine the $1/P^z$ contributions by adding the d_1' term, shown as the following two different strategies:

- Default form:

$$f(m_\pi, P^z) = f_{\text{phy}} \left[1 + d_0 \left(m_\pi^2 - m_{\pi,\text{phy}}^2 \right) + \frac{d_1}{(P^z)^2} \right], \tag{4.99}$$

- Strategy I:

$$f(m_\pi, P^z) = f_{\rm phy}\left[1 + d_0' \log\frac{m_\pi^2}{m_{\pi,\rm phy}^2} + \frac{d_1}{(P^z)^2}\right], \tag{4.100}$$

- Strategy II:

$$f(m_\pi, P^z) = f_{\rm phy}\left[1 + d_0\left(m_\pi^2 - m_{\pi,\rm phy}^2\right) + \frac{d_1}{(P^z)^2} + \frac{d_1'}{P^z}\right]. \tag{4.101}$$

Specifically, we introduce the chiral log term as well as the linear $1/P^z$ term into the fit formula, and consider the deviation between them and the default form as the systematic uncertainty. A comparison of them is shown in the left panel in Fig. 4.27 from which one can see that the results agree with each other.

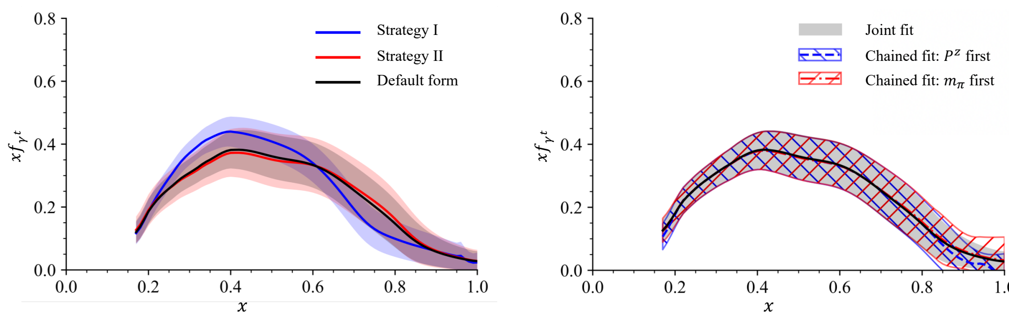

Figure 4.27: Left panel: Default extrapolation form in Eq.(4.99), considering the systematic uncertainties from chiral extrapolation in Eq.(4.100) (strategy I) and from the large-momentum extrapolation form in Eq.(4.101) (strategy II). Right panel: Comparison of the results from joint fit and chained fits with different order.

Furthermore, we have examined two different sequences of extrapolation: First, extrapolation to the physical m_π followed by the $P^z \to \infty$ extrapolation, and second, the $P^z \to \infty$ extrapolation followed by the extrapolation to the physical m_π. A comparison of the results is presented in the right panel of Fig. 4.27. It is evident that the results are in good agreement with each other.

In quasi TMDPDFs, both Lorentz structures $\Gamma = \gamma^t$ and γ^z can project onto γ^+ in the large momentum limit. An interesting observation in Ref. [150] is that the γ^t case has fewer operator

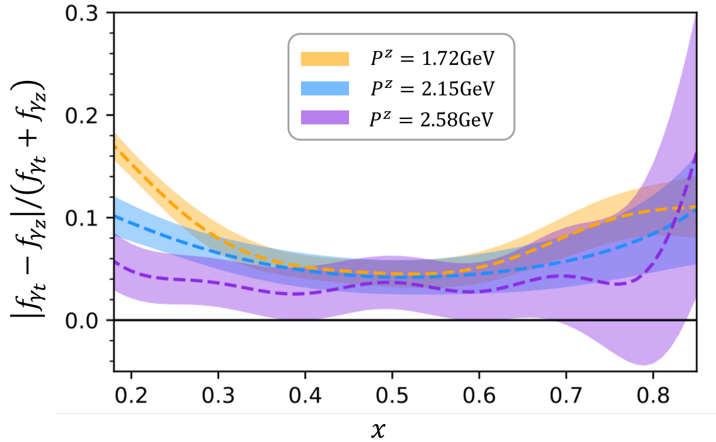

Figure 4.28: Ratio of $|f_{\gamma^t} - f_{\gamma^z}|/(f_{\gamma^t} + f_{\gamma^z})$ with $m_\pi = 220$ MeV and different P^z at $b_\perp = 3a$

mixing effects. Therefore, we use the $\Gamma = \gamma^t$ in Eq. (4.83) to extract the TMDPDFs. For the γ^z case, in addition to the sizable operator mixing effects, it is anticipated that its deviations from the γ^t results may come from power corrections from the operator product expansion of quasi correlators. These corrections are of order $\mathcal{O}(1/(P^z)^2)$ with opposite signs at large P^z. In order to analyze the impact caused by different structures, we plot the ratio $|f_{\gamma^t} - f_{\gamma^z}|/(f_{\gamma^t} + f_{\gamma^z})$ with different P^z in Fig. 4.28. One can see that with increasing nucleon momentum, the ratio becomes smaller, except for the endpoint region with large uncertainties. The differences will be incorporated as a systematic uncertainty.

To control the precision of the final results, in this work, we consider the following sources of systematic uncertainties:

- different fitting ranges in the λ-extrapolation;
- different strategies used for the chiral and P^z extrapolation;
- the difference between γ^t and γ^z results.

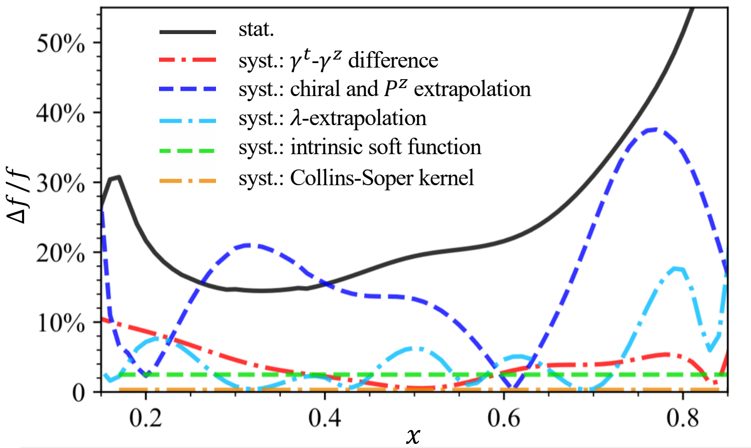

Figure 4.29: Ratios of various uncertainties and central value of final TMDPDF at $b_\perp = 3a$, include the statistical one and systematic one from: (1) λ extrapolation, (2) chiral and P^z extrapolation, (3) $\gamma^t - \gamma^z$ differences, (4) intrinsic soft function [222] and (5) Collins-Soper kernel [356].

In addition, we also consider the error propagating from the intrinsic soft function [222] and Collins-Soper kernel [356], which are calculated on the same configurations. The comparison of the statistic as well as all systematic uncertainties are shown in Fig. 4.29.

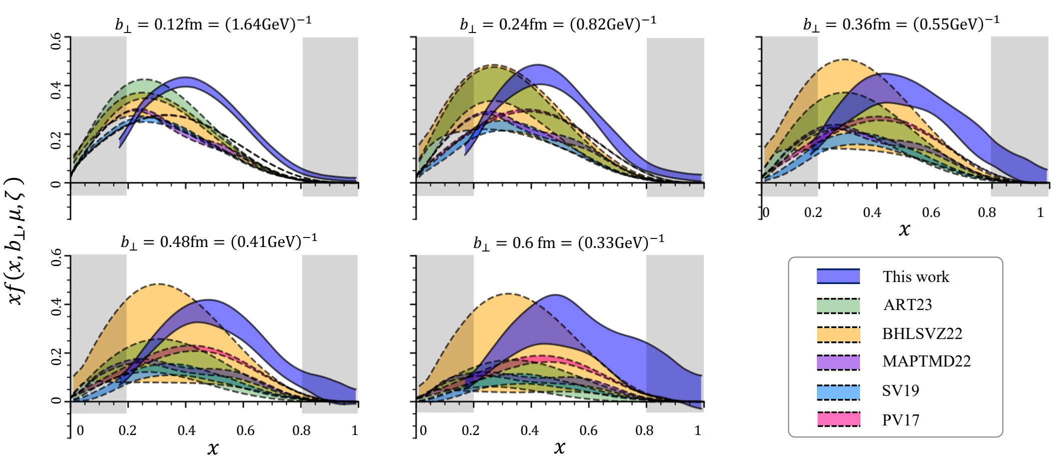

Figure 4.30: Our final results for unpolarized nucleon's isovector TMDPDFs $xf(x, b_\perp, \mu, \zeta)$ at renormalization and rapidity scales at $\mu = \sqrt{\zeta} = 2$ GeV, extrapolated to physical pion mass 135 MeV and infinite momentum limit $P^z \to \infty$, compared with ART23 [330], BHLSVZ22 [325], MAPTMD22 [326], SV19 [322]and PV17 [319] global fits. The colored bands denote our results with both statistical and systematic uncertainties, the shaded grey regions imply the endpoint regions where LaMET predictions are not reliable.

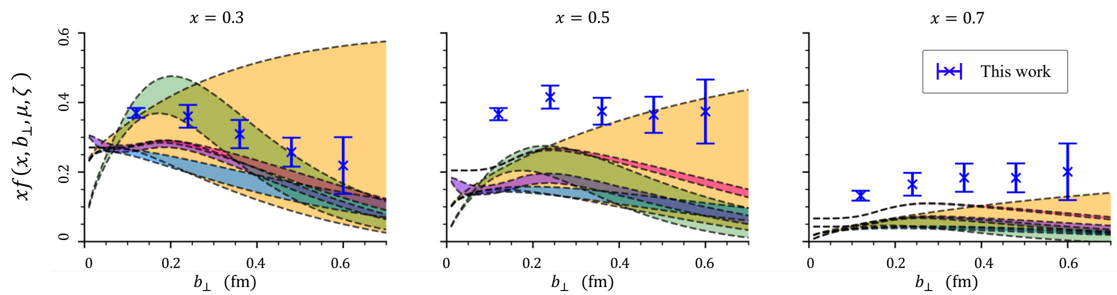

Figure 4.31: The TMDPDFs $xf(x, b_\perp, \mu = 2\ \text{GeV}, \zeta = 4\ \text{GeV}^2)$ with longitudinal momentum fraction $x = \{0.3, 0.5, 0.7\}$, together with the phenomenological results. The labels of the latter are the same as Fig. 4.30.

Combing all the known uncertainties indicated in Fig. 4.29, we obtain numerical results of unpolarized nucleon's isovector TMDPDFs from our lattice simulation. Fig. 4.30 shows $xf(x, b_\perp, \mu, \zeta)$ at renormalization and rapidity scales $\mu = \sqrt{\zeta} = 2$ GeV as a function of x, together with the phenomenological results from global analyses [319, 322, 325, 326, 330]. From the comparison one can see that, our results are in qualitative agreement with phenomenological results and share a similar behavior in $b_\perp$ space: the central values slowly decrease and uncertainties are gradually increasing with the increase of $b_\perp$.

In Fig. 4.30, a bump in the x distribution can be observed, representing the highest probability for the longitudinal momentum distributions. It is worth noting that the peak positions do not align precisely between the lattice results and the respective phenomenological results. Further investigations are required to elucidate this phenomenon.

The two shaded bands at the endpoint regions ($x < 0.2$ and $x > 0.8$) in each subplot of Fig. 4.30 indicate that LaMET predictions are not reliable there, which is estimated from the power correction terms $\Lambda^2_{\text{QCD}}/(xP^z)^2$ and threshold logarithms $\ln((1-x)P^z)$ [187, 255] . Besides, since only one lattice spacing is used, discretization uncertainties are not properly handled at this stage. Especially the $b_\perp \sim a$ case might suffer sizable discretization effects, which can be improved by a more detailed analysis in the future.

Fig. 4.31 shows the results for $xf(x, b_\perp, \mu, \zeta)$ with $x = 0.3, 0.5, 0.7$ as a function of spatial separation $b_\perp$. The spatial distributions reflect correlations between the partons at transverse interval $b_\perp$ inside a highly boosted nucleon, and will reveal aspects of nucleon structure. It could be conjectured that the distributions vanish when $b_\perp$ is larger than the nucleon radius, however, neither the present day lattice results nor the phenomenological results are precise enough to draw such a conclusion.

4.5.3 Conclusion

In summary, we have performed the first calculation of TMDPDFs within a nucleon using the LaMET framework. State-of-the-art techniques in renormalization and extrapolation on the lattice have been employed, including the consideration of the perturbative kernel up to NNLO with RG resummation. We investigate the dependence on pion mass and hadron momentum, incorporating both statistical and systematic errors to provide a robust characterization of the inner structure of nucleons through the parton distributions.

While the current findings are promising, further improvements are required. Firstly, a comprehensive analysis of multiple ensembles with varying lattice spacings, pion masses, and volumes is essential. This approach will not only reassess all uncertainties addressed in this study but also systematically explore additional significant factors. Secondly, our results exhibit notable theoretical uncertainties in the endpoint regions, which could stem from power corrections or threshold logarithms. Thirdly, we expect a decay of the TMDPDF with increasing $b_\perp$ which is not (yet) visible in our data. Going to larger $b_\perp$ is thus imperative. Moreover, the presence of chiral logarithms in the analysis of quasi PDFs (as demonstrated in Ref. [357]) highlights the potential for similar logarithms in our case, warranting dedicated theoretical investigation.

Chapter 5: Applications of Machine Learning in Lattice Gauge Theory

In recent years, machine learning has become a powerful paradigm in scientific research, enabling data driven discovery and significantly accelerating traditional computational and experimental workflows. In biology and chemistry, a prominent example is AlphaFold [358, 359], which demonstrated that deep learning models trained on large structural databases can predict protein three-dimensional structures with near experimental accuracy, transforming structural biology and molecular design. In materials science, the GNoME (Graph Networks for Materials Exploration) models represent a significant leap; by scaling deep learning to predict the stability of crystalline compounds, it has discovered over 2.2 million new structures below the current convex hull, many of which escaped previous human chemical intuition [360]. In astronomy, the shift toward foundation models is exemplified by AION-1 (Astronomical Omni-modal Network) [361], a multi-billion parameter model developed by PolymathicAI. Unlike previous CNN-based classifiers, AION-1 integrates heterogeneous data—including imaging, spectroscopy, and metadata—across multiple sky surveys, enabling a unified approach to tasks such as spectral super-resolution and galaxy property estimation without the need for task-specific retraining. Collectively, these advancements illustrate how AI is evolving from a specialized tool into a foundational methodological layer across different scientific domains.

Building on these representative successes, it is important to clarify the types of scientific problems for which machine learning is particularly well suited. In practice, such problems are typically outcome oriented rather than mechanism centered, and their solutions can be independently verified without introducing additional, uncontrolled systematic uncertainties from the learning model itself. In these settings, neural networks function primarily as high dimensional function approximators that map inputs to outputs with high accuracy, while the ultimate scientific validity of the result is established through external benchmarks such as experiments, higher fidelity simulations, or well-defined evaluation metrics. This separation between prediction and validation mitigates concerns about limited interpretability, since correctness does not rely on understanding the internal representations of the model.

AlphaFold [358, 359] provides a paradigmatic example. Protein structure prediction is intrinsically predictive: given an amino acid sequence, the goal is to determine its three-dimensional structure. The predicted structures can then be directly compared against independent experimental measurements, including X-ray crystallography and cryo electron microscopy, using clear quantitative metrics. Because validation is external and objective, the use of a deep neural network does not introduce uncontrolled systematic bias into the final scientific claim. The transformative impact of AlphaFold on structural biology and molecular design illustrates why such verifiable, result driven problems are especially well-matched to modern AI methods, and why this achievement was recognized with the 2024 Nobel Prize in Chemistry.

Building on this perspective, meaningful applications of machine learning in lattice gauge theory should satisfy the same criteria outlined above: they should be outcome oriented and admit independent validation through established consistency checks. Such caution is essential because lattice gauge theory provides a nonperturbative, first-principles formulation of quantum field theory on a discrete spacetime. Its predictive power rests on systematically controllable uncertainties, including finite lattice spacing effects, finite volume effects, and statistical errors, which can be quantified and removed through well-defined extrapolation procedures. Thus, a good application of machine learning should preserve this rigorous framework and avoid introducing additional, uncontrolled sources of systematic uncertainty.

Broadly speaking, current applications of machine learning in lattice gauge theory can be grouped into three categories. The first concerns gauge configuration generation, where learned proposals or generative models are used to accelerate sampling and mitigate long autocorrelation times. The second involves data analysis, where machine learning supports tasks such as operator optimization and parameter inference from correlators or ensembles. The third emerging direction encompasses AI agents and foundation models, which go beyond simple automation by learning to accelerate end to end workflows and propose improved strategies. In this chapter, we will discuss these three classes of applications in turn.

5.1 Applications in Gauge Configuration Generation

The generation of gauge configurations constitutes the computational foundation of lattice gauge theory simulations. In this framework, physical observables are defined through Euclidean path integrals over gauge fields, which are discretized on a spacetime lattice to provide a nonperturbative regularization. In practice, these high dimensional integrals are evaluated by importance sampling of gauge field configurations according to the Boltzmann weight of the lattice action. The central task of lattice gauge theory simulations is therefore to construct a Markov chain whose stationary distribution is proportional to $e^{-S(U)}$, so that expectation values of observables can be approximated by ensemble averages over sampled configurations.

Formally, the partition function is given by

$$Z = \int \mathcal{D}U \; e^{-S(U)} \; , \tag{5.1}$$

where U denotes the set of link variables and S(U) the lattice action. Expectation values are estimated as Monte Carlo averages over configurations generated via Markov Chain Monte Carlo (MCMC) methods. A variety of MCMC algorithms are employed in practice, all designed to ensure exactness through detailed balance and ergodicity. As the continuum limit is approached, however, these algorithms exhibit critical slowing down: autocorrelation times grow rapidly with decreasing lattice spacing and increasing physical volume. This effect is particularly severe for topological observables, whose dynamics may effectively freeze over long Monte Carlo histories. The resulting computational cost represents one of the central algorithmic bottlenecks in high-precision lattice QCD.

5.1.1 Overview

Motivated by the algorithmic limitations of conventional MCMC methods, a variety of machine learning approaches have recently been proposed to accelerate gauge configuration generation. These developments can be broadly categorized into two directions. The first aims at direct generative modeling of the target distribution, while the second focuses on integrating learned components into existing MCMC methods.

Among generative approaches, normalizing flow [91, 92] have received particular attention. A normalizing flow constructs an invertible and differentiable map between a simple base distribution and the target gauge-field distribution, with a tractable Jacobian determinant. In principle, such flows can serve as exact bridges between distributions, enabling either reweighting or independence Metropolis sampling. If a sufficiently accurate flow were available, one could perform approximate direct sampling of gauge configurations, thereby eliminating long autocorrelation times and potentially resolving topological freezing.

In practice, however, applying normalizing flow to four-dimensional non-abelian gauge theories presents substantial challenges. The dimensionality of configuration space scales with the lattice volume, gauge equivariance must be preserved, and locality constraints must be respected. Recent progress has improved model expressivity through gauge-equivariant architectures and learned loop constructions, and has demonstrated advantages in correlated ensemble generation and replica exchange frameworks. Nevertheless, fully replacing HMC at state-of-the-art lattice sizes remains an open challenge.

These considerations motivate hybrid approaches in which learned transformations are integrated into established MCMC frameworks. By retaining the exactness guarantees of traditional algorithms while reshaping the geometry of configuration space through machine-learned maps, one may achieve algorithmic acceleration without abandoning theoretical control.

5.1.2 Neural Field Transformations for Hybrid Monte Carlo

A particularly promising direction along this line is the construction of invertible neural field transformations to modify the geometry of configuration space prior to HMC. Conceptually, this idea is closely related to Lüscher's notion of trivializing maps [362], where a change of variables in the path integral is introduced to simplify the effective action and smooth the energy landscape explored by HMC. Although exact trivializing maps are not known for interacting non-abelian gauge theories, approximate transformations can still reshape the distribution in a way that reduces force fluctuations and mitigates critical slowing down.

Based on this insight, our recent work [363] presents a systematic of neural network transformed HMC (NTHMC) [364,365], focusing on how neural architectural designs such as residual connections, attention mechanisms, channel-dependent activation, and batch size affect performance and scaling. Testing on a two-dimensional U(1) gauge theory, we evaluate both small-lattice performance and the transferability to larger volumes and smaller lattice spacing. Our results suggest that wider receptive fields and channel-dependent activation improve efficiency while maintaining favorable scaling, thereby informing extensions to four-dimensional SU(3) gauge fields and large-scale lattice QCD.

The NTHMC framework uses an invertible, gauge-covariant field transformation $\mathcal{F} : \tilde{U} \mapsto U$ parameterized by a neural network to perform a change of variables in the integration over the Boltzmann weight $\exp(-S(U))$, leading to an effective action,

$$S_{\mathrm{FT}}(\tilde{U}) = S(\mathcal{F}(\tilde{U})) - \log|\det \mathcal{F}*(\tilde{U})| \,, \tag{5.2}$$

where $\mathcal{F}*(\tilde{U})$ is the Jacobian matrix of the transformation. HMC is then performed in the space of auxiliary variables $\tilde{U}$ with the Boltzmann weight $\exp(-S_{\mathrm{FT}}(\tilde{U}))$.

Intuitively, the field transformation acts like a learned preconditioner. We construct the field transformation from gauge-covariant, locally equivariant convolutional neural networks (CNNs). For each link variable $\tilde{U}_{x,\mu}$ (a gauge field element going from site x to $x+\hat{\mu}$), the update takes the form

$$\mathcal{F} : \tilde{U}_{x,\mu} \mapsto U_{x,\mu} = e^{\Pi_{x,\mu}} \tilde{U}_{x,\mu}, \quad \Pi_{x,\mu} = \sum_l \epsilon_{x,\mu,l} \, W_{x,\mu,l} \,. \tag{5.3}$$

The coefficients are given by $\epsilon_{x,\mu,l} = \phi[N_l(X, Y, \ldots)]$, with N_l a CNN mapping local gauge-invariant inputs $(X, Y, \ldots)$ to the coefficients for the loop functions $W_{x,\mu,l}$, and ϕ a bounded activation function (e.g. $\tanh$ or $\arctan$) that enforces the invertibility of the transformation. $W_{x,\mu,l}$ is the Lie algebra projection of a specific (labeled by l) Wilson loop (path ordered products of gauge fields along a closed loop) that starts and ends at site x and first moves in the μ direction (the first on the right of the product). For the $U(1)$ gauge theory, the projection just takes the imaginary part.

As suggested in Ref. [362], the lattice links are partitioned into eight disjoint subsets and updated

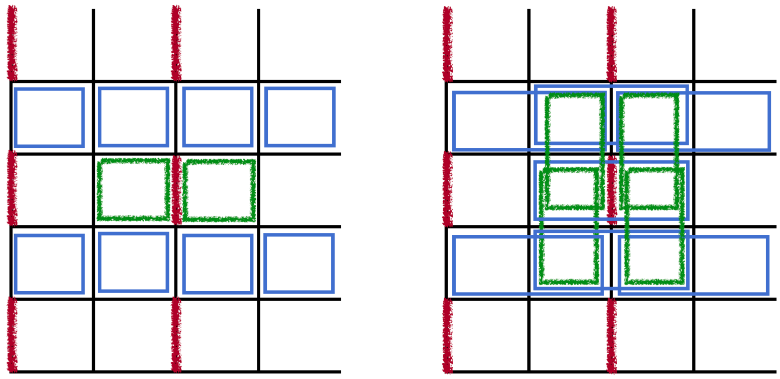

Figure 5.1: Cropped section of a lattice configuration illustrating the field transformation. Red lines mark one of the eight link subsets updated sequentially, green lines indicate Wilson loops $W_{x,\mu,l}$, and blue lines show the gauge-invariant features $(X, Y, \ldots)$ used as CNN inputs.

sequentially. This ensures that the Jacobian determinant of each update step is reduced to the product of diagonal elements. Moreover, by choosing CNN inputs $(X, Y, \ldots)$ to be independent of the links in the active subset, these diagonal elements can be written in a closed form, making the Jacobian determinant explicitly tractable. In this work, we take the l index to run over 1×1 plaquettes and 1×2 rectangles, as illustrated in Fig. 5.1, where one link subset (red) is updated, while the corresponding loops (green) and local gauge-invariant CNN inputs (blue) are highlighted.

For U(1) gauge theory, we use $U = e^{i\theta}$, and the Wilson plaquette action [58] as $S(U)$. The transformation reduces to exponentials of sine terms of plaquette and rectangle loops,

$$\mathcal{F} :\ \tilde{U}_{x,\mu} \ \mapsto\ U_{x,\mu} = e^{\sum_l \epsilon_{x,\mu,l} W_{x,\mu,l}} \,\tilde{U}_{x,\mu} = e^{\sum_p i\epsilon_{x,\mu,p} \sin(\theta_p) + \sum_r i\epsilon_{x,\mu,r} \sin(\theta_r)} \tilde{U}_{x,\mu}\ , \tag{5.4}$$

with p and r denoting individual plaquette and rectangle loops, respectively, and θ_p, θ_r the phase angle of the corresponding Wilson loop. The Jacobian has the matrix elements

$$\mathcal{F}_*(\tilde{U})_{x,\mu,x',\mu'} = \delta_{x,x'}\delta_{\mu,\mu'} \left[1 + \sum_p \epsilon_{x,\mu,p} \cos(\theta_p) + \sum_r \epsilon_{x,\mu,r} \cos(\theta_r) \right] . \tag{5.5}$$

In order to have an invertible transformation, the Jacobian determinant may not change sign. The invertibility condition is then

$$\sum_p |\epsilon_{x,\mu,p}| + \sum_r |\epsilon_{x,\mu,r}| < 1 \, . \tag{5.6}$$

The field transformation is trained on configurations generated from regular HMC by minimizing the transformed action force:

$$\mathcal{L}(\tilde{U}) = \sum_{p\in\{2,4,6,8\}} \frac{c_p}{V^{1/p}} \left\| \frac{\partial S_{\mathrm{FT}}(\tilde{U})}{\partial \tilde{U}} \right\|_p , \tag{5.7}$$

where V is the lattice volume and c_p are hyperparameters that weight different force norms, thereby balancing the emphasis between average and peak forces. In our experiments, we set $c_2 = c_4 = c_6 = c_8 = 1$. Different from Ref. [364], which matched the transformed force to that at a larger lattice spacing, we directly minimize the transformed force itself, effectively training the transformation to smooth the action landscape, reduce the stiff directions and improve the sampling efficiency.

Inside the HMC update we employ the 2nd order minimum norm symplectic integrator of Omelyan et al. [366] with 10 MD steps per trajectory, tuning the step size to achieve $\sim 80\%$ acceptance. All HMC simulations were verified by checking the plaquette values of the generated configurations, which were found to agree with the theoretical expectations within statistical uncertainties. We benchmark five architectures: a baseline (`Base`) model similar to the design in Ref. [364], and four variants (`Tanh`, `Resn`, `Attn`, `Comb`) introducing different modifications.

Base The baseline model is a two-layer CNN that processes six input channels (two plaquette, four rectangle). A first 3×3 convolution maps them to 12 hidden channels with GELU activation, followed by a second 3×3 convolution producing coefficients, which are scaled by $\arctan$ to $(-1/6, 1/6)$, satisfying the invertibility condition in Eq. 5.6. This design yields an effective 5×5 receptive field with ∼2,000 parameters.

Tanh Same as `Base`, but replaces the $\arctan$ scaling with a channel-dependent $\tanh$ activation, ensuring $|\epsilon_{x,\mu,p}| < 0.4$ and $|\epsilon_{x,\mu,r}| < 0.05$, with $\sim 2,000$ parameters.

Resn Adds two residual blocks (each has two 3×3 convolutions with skip connections) before the output, and replaces the second 3×3 convolution of `Base` with a 1×1 projection. This enlarges the receptive field to 11×11 and increases the parameters to $\sim 6,000$.

Attn Augments `Base` with a channel-attention module between the two convolutions, reweighting channels without changing the 5×5 receptive field. The parameter count is $\sim 2,100$.

Comb Combines residual blocks, channel attention, and channel-dependent $\tanh$ activation ($|\epsilon_{x,\mu,p}| < 0.4$, $|\epsilon_{x,\mu,r}| < 0.05$). The resulting architecture has an 11×11 receptive field and $\sim 6,100$ parameters.

All models are trained with the p-norm force loss (Eq. (5.7)) on 2^{12} independent configurations at $\beta = 3$, $V = 32^2$ using an 80/20 train–test split. The input gauge coupling, β, enters in the action $S(U)$, and in this theory, the lattice spacing is proportional to $\beta^{-1/2}$. The default batch size is 64, except that `Base` and `Comb` are also trained with 32 (denoted `Base32`, `Comb32`) to study batch-size effects. Each model is trained for 32 epochs in PyTorch [367], with one single run taking about 12–20 minutes on a single NVIDIA A100 GPU for batch size 64, and roughly twice as long for batch size 32.

To assess the sampling efficiency of regular HMC and NTHMC, we consider two complementary metrics. The first is $\gamma(\delta)$, which measures the relative independence of the configuration [364]:

$$\gamma(\delta) = \frac{1}{1 - \Gamma_t(\delta)}\,, \quad \Gamma_t(\delta) \equiv \frac{\langle Q_\tau Q_{\tau+\delta} \rangle}{\langle Q^2 \rangle} = 1 - \frac{\langle (Q_\tau - Q_{\tau+\delta})^2 \rangle}{2V \chi_t^\infty(\beta)}\,. \tag{5.8}$$

Here $\Gamma_t(\delta)$ is the autocorrelation of the topological charge Q at separation δ in HMC steps, and $\chi_t^\infty(\beta)$ is the topological susceptibility [364]. The larger $\gamma(\delta)$ indicates stronger autocorrelations and therefore less efficient sampling.

The second metric is the average topological change ΔQ, the mean absolute shift in topological charge per HMC step, which directly probes the tunneling between sectors; larger values of ΔQ indicate more frequent topological transitions and thus more efficient sampling. All training and evaluation experiments are repeated with 8 random seeds. Statistical uncertainties are estimated and consistently propagated through derived quantities using Jackknife resampling [368] across the seed variability.

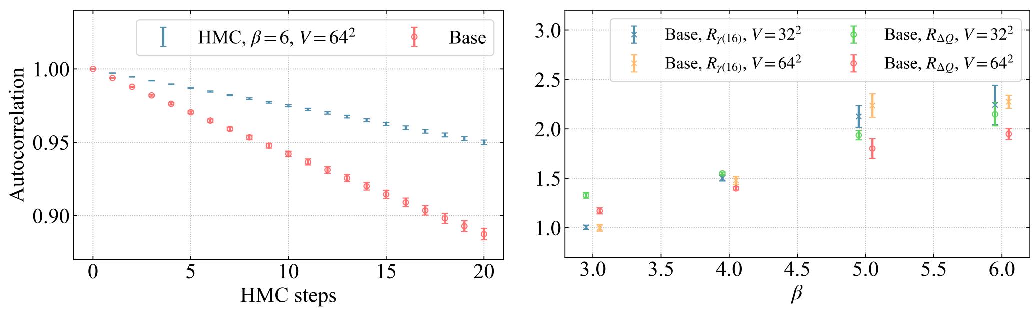

Figure 5.2: Comparison of regular HMC and NTHMC with the `Base` model. Left: autocorrelation decay at $\beta = 6.0$, $V = 64^2$. Right: scaling of two efficiency ratios across β and volume, with $R_{\gamma(16)} \equiv \gamma(16)_{\mathrm{HMC}}/\gamma(16)_{\mathrm{Base}}$ and $R_{\Delta Q} \equiv \Delta Q_{\mathrm{Base}}/\Delta Q_{\mathrm{HMC}}$.

As shown in the left panel of Fig. 5.2, NTHMC with the `Base` model achieves a markedly faster decay of autocorrelations than regular HMC. The right panel reports the efficiency ratios $R_{\gamma(16)} = \gamma(16)_{\mathrm{HMC}}/\gamma(16)_{\mathrm{Base}}$ and $R_{\Delta Q} = \Delta Q_{\mathrm{Base}}/\Delta Q_{\mathrm{HMC}}$. Despite being trained only at $\beta = 3$, $V = 32^2$, NTHMC yields even larger relative gains at higher β (smaller lattice spacing) and larger volumes, demonstrating favorable scaling and generalization. At $\beta = 3$, where regular HMC is already efficient, both ratios remain close to unity.

Fig. 5.3 summarizes the performance of different CNN architectures within NTHMC across several lattice couplings and volumes. The comparison includes results at $\beta = 5$, 6, and 7 with lattice

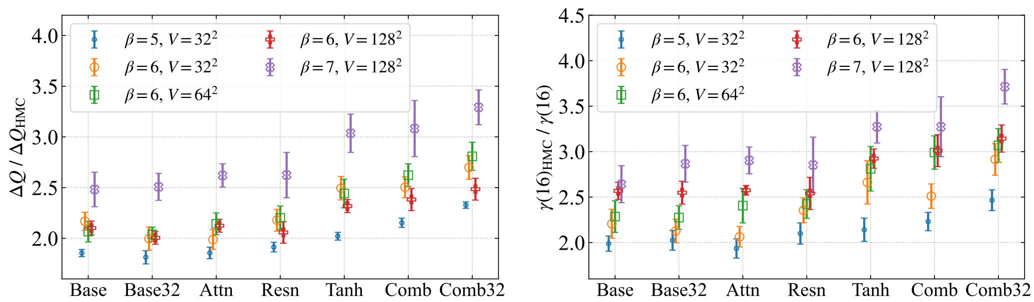

Figure 5.3: Comparison of NTHMC with various network architectures across different lattice couplings and volumes. The plots show ΔQ (left panel) and $\gamma(16)$ (right panel), both normalized to standard HMC results.

sizes $V = 32^2$, 64^2, and 128^2, allowing us to examine both coupling and volume dependence relative to standard HMC. Residual blocks and channel-dependent activation lead to improvements over the `Base` model, and the combined design (`Comb`) yields the strongest effect, achieving nearly a 40% reduction in $\gamma(16)$ with a corresponding increase in the average topological change ΔQ. Moreover, as the coupling β increases, all NTHMC models exhibit a higher sampling efficiency compared to standard HMC, and the efficiency does not show statistically significant changes with increasing lattice volume, indicating that our method can be effectively extended to larger volumes and finer lattices.

Interestingly, the `Tanh` model achieves competitive gains despite having the same parameter count as `Base`, which we attribute to the faster saturation of $\tanh$ compared to $\arctan$ and, more importantly, to the larger magnitude of plaquette coefficients $\epsilon_{x,\mu,p}$ enabled by channel-dependent activation. Likewise, residual connections expand the receptive field, which might be expected to hinder transferability, yet we find that an appropriate enlargement does not compromise scaling or generalization and may even be beneficial. A possible explanation is that, as the continuum limit is approached in LGT, infrared (long distance) structures are preserved while ultraviolet (short distance) details become increasingly refined, so the transferable features are precisely the long-range correlations that benefit from a relatively wider receptive field.

Finally, batch size interacts mildly with architecture. Although the baseline shows little sensitivity, the richer `Comb` model benefits from smaller batches, consistent with gradient noise acting as an implicit regularizer that helps larger capacity models generalize more effectively.

Our preliminary results demonstrate that neural field transformations can substantially reduce autocorrelations, yet several open directions remain. Future work includes exploring integrator choices such as varying the number of MD steps at fixed trajectory length, refining training objectives beyond minimizing the force norms, and systematically analyzing training regimes (e.g. number of epoches, learning rate, training set size and data from different β and volumes). Finally, extending from U(1) to SU(3) will be essential to connect this approach with large-scale lattice QCD applications.

5.2 Applications in Data Analysis

Beyond configuration generation, machine learning has found increasingly important applications in the data analysis stage of lattice gauge theory. In contrast to sampling algorithms, which operate at the level of the path integral measure, data-driven methods act on correlation functions or ensembles that have already been generated. These approaches can be broadly divided into two categories: operator optimization and parameter inference from correlators or gauge ensembles.

5.2.1 Operator Optimization

The extraction of physical observables from lattice simulations relies critically on the choice of interpolating operators. An ideal operator has strong overlap with the ground state and suppressed overlap with excited states, thereby reducing statistical noise and systematic contamination. Traditionally, operator construction is guided by symmetry considerations and physical intuition, supplemented by variational techniques such as the GEVP.

Machine learning provides a complementary strategy: treat operator construction as an optimization problem. Neural networks can be trained to maximize ground-state overlap or minimize excited-state contamination, effectively learning optimal linear combinations of operator bases. Related ideas have been explored in the context of neural-network-assisted spectroscopy and variational optimization, where data-driven methods improve the efficiency of spectral extraction [369]. More recently, automated operator discovery and structure learning techniques have been proposed to systematically search large operator spaces [370]. These developments suggest that operator optimization is a natural interface between traditional lattice methodologies and modern representation learning.

5.2.2 Parameter Inference from Correlators and Ensembles

A second major application concerns parameter inference. In lattice QCD, physical observables such as hadron masses, form factors, and matrix elements are obtained from Euclidean correlation functions through multi-exponential fits, Bayesian model averaging, or spectral reconstruction techniques. This inverse problem is inherently ill-conditioned: correlators decay exponentially in Euclidean time, and statistical noise grows at large separations.

Machine learning approaches have been proposed to assist in this inference task. Neural networks can be trained to extract spectral parameters directly from correlator data, effectively learning non-linear regression mappings between correlation functions and physical parameters. In addition, Bayesian neural networks and Gaussian-process-based methods provide flexible frameworks for incorporating prior information and quantifying uncertainties. Ensemble-level inference tasks, such as identifying phase transitions or estimating renormalization constants, have similarly benefited from

ML-based regression and classification techniques.

Importantly, these data-analysis applications do not alter the underlying field theory or sampling algorithm. Instead, they enhance the extraction of physics from already-generated data, often reducing human intervention and improving robustness against model selection ambiguities. As lattice calculations continue to increase in precision and complexity, systematic and automated data analysis tools will become increasingly valuable.

5.3 AI Agent and Foundation Model

Recent advances in artificial intelligence suggest a broader paradigm beyond task-specific machine learning models: the development of AI agents and foundation models capable of operating across complex scientific workflows. An AI agent may be understood as a system that interacts with an environment, makes sequential decisions under specified objectives, and adapts its strategy based on feedback. In computational science, such an agent would not merely perform a single regression or classification task, but rather orchestrate multi-step pipelines, select among competing algorithms, tune hyperparameters, and iteratively refine intermediate results. A foundation model, by contrast, refers to a large-scale model trained on broad and diverse data within a domain, capable of learning general-purpose representations that transfer across tasks. Rather than solving one narrowly defined problem, a foundation model encodes structural regularities of the domain and serves as a reusable backbone for downstream applications.

Lattice gauge theory provides a particularly well-structured environment in which to explore these ideas. Euclidean correlation functions possess strict symmetry constraints, dimensional consistency, and well-defined spectral decompositions. The transformation from correlation functions to physical observables such as spectra, matrix elements, and form factors follows a systematic and theoretically controlled pipeline. Unlike many empirical domains, the correctness of intermediate and final results can be verified through known identities, scaling relations, and continuum-limit consistency. These characteristics make lattice QCD an ideal testbed for evaluating the reasoning capabilities, robustness, and limitations of AI-based scientific systems.

Initial exploratory work has begun to examine the role of large models and AI systems in structured physics problems [371]. However, a comprehensive framework tailored to lattice gauge theory remains largely undeveloped. A natural first step toward such a framework is the construction of quantitative benchmarks. The lattice community possesses substantial collections of correlation function data, including ensembles that were generated but not fully analyzed or not publicly disseminated. For many of these datasets, the extracted ground-state spectra and matrix elements are already known from conventional analyses. This setting enables the creation of controlled benchmark problems in which an AI system is provided with raw correlators and evaluated on its ability to reproduce established results, quantify uncertainties, and choose appropriate fitting strategies. Because the spectral decomposition of two- and three-point functions is well understood, such benchmarks can be scaled in difficulty by varying noise levels, operator bases, or excited-state contamination.

Beyond benchmarking, one may envision the construction of an AI agent capable of automating the full analysis workflow from ensemble to physical observables. Such a workflow would encompass the design and optimization of interpolating operators, the choice of smearing schemes and lattice actions, the implementation of Wick contractions, the extraction of spectra using multi-state fits or generalized eigenvalue methods, the application of Feynman-Hellmann techniques for matrix elements, and the extrapolation to physical limits through continuum and chiral fits. Many of these steps involve model selection, numerical stability considerations, and trade-offs between statistical precision and systematic control. An effective agent would not only reproduce established procedures but also evaluate alternatives, compare competing methods such as GEVP-based approaches, model averaging, Lanczos-type analyses, or current-inspired constructions, and identify strategies that improve stability or efficiency. In this sense, the goal is not merely automation, but the possibility of creative methodological exploration guided by physical constraints.

Foundation models offer a complementary long-term perspective. By training on large collections of correlators and associated metadata across different lattice spacings, volumes, hadron channels, and actions, a foundation model could learn universal representations of Euclidean observables. Such representations might enable transfer across related tasks, facilitate anomaly detection in new datasets, or accelerate parameter inference by providing strong domain-aware priors. Crucially, the theoretical rigidity of lattice gauge theory provides natural validation criteria: symmetries, Ward identities, scaling behavior, and renormalization properties serve as internal consistency checks. This allows the assessment of foundation models not only in terms of predictive accuracy, but also in terms of physical consistency and generalization across regimes.

In summary, lattice gauge theory occupies a unique position at the intersection of high-performance computing, structured statistical inference, and first-principles theoretical control. The development of AI agents and foundation models in this domain is both technically feasible and scientifically meaningful. By establishing rigorous benchmarks and progressively integrating learned systems into established workflows, one may systematically evaluate the capabilities and limitations of artificial intelligence in a controlled, theory-driven environment.

Chapter 6: Summary and Outlook

This dissertation has been motivated by a question: how to determine partonic structures from QCD in a framework that remains first-principles while being quantitatively relevant for precision phenomenology. Across the previous chapters, the work has been organized around a consistent strategy that combines lattice gauge theory, LaMET framework, control of systematic uncertainties, and machine-learning-based methods. The objective is to present a coherent computational pathway from Euclidean matrix elements to light-cone partonic information in a model-independent manner with controlled precision.

This thesis presents a set of successful lattice calculations of collinear and transverse partonic observables, demonstrating that the LaMET framework has evolved into a practically implementable and systematically improvable approach. For meson structure, the lattice extraction of pion DAs demonstrates that the LaMET framework can now be carried through with mature renormalization and extrapolation strategies, yielding distributions that exhibit physically interpretable deviations from asymptotic forms and broad consistency with complementary nonperturbative approaches in the reliable kinematic regions. For baryon structure, the studies of nucleon PDFs and TMDs have extended this logic to substantially more demanding systems, where dense excitation spectra, noisier correlators, and richer operator structures intensify the analysis burden. The results show that state-of-the-art treatments of renormalization, finite-separation behavior, and momentum dependence can produce stable nucleon-structure signals that align qualitatively with existing phenomenological results while preserving visibility of unresolved tensions, especially in TMD cases and in sectors where higher-order or power-suppressed contributions are still numerically important. Taken together, these outcomes indicate that lattice partonic physics has moved beyond proof-of-principle demonstrations toward a regime of structured precision building, where the key questions are less about feasibility and more about how systematically one can suppress the remaining uncertainty budget across observables.

In parallel with the results of lattice calculations, this dissertation emphasizes that computational methodology has become an increasingly important component of research in lattice gauge theory. The machine-learning chapter presents a pragmatic framework for identifying where learning-based tools can contribute in first-principles calculations: acceleration of gauge-configuration generation, enhancement of data-analysis workflows, and the longer-term development of AI systems that can assist in multi-stage scientific inference. In the configuration-generation domain, neural field transformations integrated with hybrid Monte Carlo illustrate a promising route to reducing autocorrelation and improving sampling efficiency while retaining physical distributions through the underlying Markov-chain framework. In data analysis, learning-based methods can support operator optimization and parameter inference when they are embedded in externally verifiable pipelines and evaluated against established physical consistency checks. At a broader horizon, the discussion of AI agents and foundation-model concepts has framed lattice gauge theory as a suitable environment for benchmark-driven development, because symmetry constraints, controlled limits, and cross-checkable intermediate quantities provide strict criteria for model assessment. The central conclusion is that machine learning is most useful in this domain when it supports controlled calculations rather than replacing theoretical structure or systematic uncertainty analysis.

At the same time, the thesis makes clear that several limitations remain decisive for the next stage of progress, and these limitations should be viewed as concrete research targets in future investigations. Excited-state contamination continues to be a dominant challenge in nucleon observables and can produce nontrivial biases if fit-model choices, source-sink separations, and analysis windows are not jointly constrained. Continuum, infinite-volume, and physical-mass control remains uneven across different quantities, as finite computational resources make precision studies near these limiting regimes technically demanding. In the LaMET context, achievable hadron boost factors, and residual higher-power corrections continue to limit the quantitative precision and systematic control of the extracted distributions, and these effects can couple to practical choices in non-perturbative renormalization, and perturbative matching. None of these issues invalidates the present results; instead, they define the concrete boundary between current precision and future precision, and they emphasize that progress depends on coordinated improvements in simulation design, analysis methodology, and theoretical control.

Looking ahead, further progress should proceed in clearly differentiated stages. In the near term, the priority is the systematic deployment of several recently developed innovations, including kinematically enhanced interpolating operators, the current-inspired GEVP method [309, 310], the Lanczos algorithm [306–308], and asymptotic analyses at large distances [279, 280]. These developments aim to better control excited-state contamination and Fourier reconstruction systematics, while improving the SNR, with the overarching objective of enhancing the stability of the analysis and enabling more reliable estimates of systematic uncertainties. In the mid term, emphasis should shift toward structural optimization at both the lattice and perturbative interfaces, including optimized source profiles and improved momentum injection strategies to enhance overlap with highly boosted states, higher-order precision in perturbative matching to reduce truncation uncertainties, and more robust nonperturbative renormalization procedures—particularly within CG method—to achieve a reduction of systematic uncertainties, while multi-observable programs are developed within unified uncertainty frameworks to guarantee internal consistency across collinear and transverse sectors. In parallel, algorithmic acceleration and AI-based tools should continue to mature beyond controlled prototypes into full SU(3) production environments, with strict validation criteria tied to physical consistency, reproducibility, and systematic uncertainty propagation, so that computational efficiency gains translate into scientifically reliable outputs. In the long term, once lattice extractions and phenomenological analyses achieve mutually compatible precision, the objective is a genuinely joint program in which first-principles calculations and global data constraints are incorporated into coherent combined analyses, enabling iterative refinement on both sides and ultimately delivering quantitatively sharper and more internally consistent determinations of partonic structure.

Bibliography

[1] P. W. Anderson. More Is Different. *Science*, 177(4047):393–396, 1972.

[2] David J. Gross and Frank Wilczek. Ultraviolet Behavior of Nonabelian Gauge Theories. *Phys. Rev. Lett.*, 30:1343–1346, 1973.

[3] H. David Politzer. Reliable Perturbative Results for Strong Interactions? *Phys. Rev. Lett.*, 30:1346–1349, 1973.

[4] John Collins. *Foundations of Perturbative QCD*, volume 32. Cambridge University Press, 2011.

[5] J. D. Bjorken. Asymptotic Sum Rules at Infinite Momentum. *Phys. Rev.*, 179:1547–1553, 1969.

[6] Richard P. Feynman. Very high-energy collisions of hadrons. *Phys. Rev. Lett.*, 23:1415–1417, 1969.

[7] A. Accardi et al. Electron Ion Collider: The Next QCD Frontier: Understanding the glue that binds us all. *Eur. Phys. J. A*, 52(9):268, 2016.

[8] R. Abdul Khalek et al. Science Requirements and Detector Concepts for the Electron-Ion Collider: EIC Yellow Report. *Nucl. Phys. A*, 1026:122447, 2022.

[9] Franz Gross et al. 50 Years of Quantum Chromodynamics. *Eur. Phys. J. C*, 83:1125, 2023.

[10] R. Angeles-Martinez et al. Transverse Momentum Dependent (TMD) parton distribution functions: status and prospects. *Acta Phys. Polon. B*, 46(12):2501–2534, 2015.

[11] Dave Gaskell. Deep inelastic scattering, part 1, 2015. CTEQ 2015 Summer School, July 8–9, 2015, SMU. Slides available at `http://www.physics.smu.edu/olness/ftp/misc2/cteq/2015/gaskell_dis_part1.pdf`.

[12] J. D. Bjorken and Emmanuel A. Paschos. Inelastic Electron Proton and gamma Proton Scattering, and the Structure of the Nucleon. *Phys. Rev.*, 185:1975–1982, 1969.

[13] V. N. Gribov and L. N. Lipatov. Deep inelastic e p scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15:438–450, 1972.

[14] Yuri L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov. Phys. JETP*, 46:641–653, 1977.

[15] Guido Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys. B*, 126:298–318, 1977.

[16] Kenneth G. Wilson. Nonlagrangian models of current algebra. *Phys. Rev.*, 179:1499–1512, 1969.

[17] S. D. Drell and Tung-Mow Yan. Partons and their Applications at High-Energies. *Annals Phys.*, 66:578, 1971.

[18] G. Peter Lepage and Stanley J. Brodsky. Exclusive Processes in Quantum Chromodynamics: Evolution Equations for Hadronic Wave Functions and the Form-Factors of Mesons. *Phys. Lett. B*, 87:359–365, 1979.

[19] G. Peter Lepage and Stanley J. Brodsky. Exclusive Processes in Perturbative Quantum Chromodynamics. *Phys. Rev. D*, 22:2157, 1980.

[20] A. V. Efremov and A. V. Radyushkin. Asymptotical Behavior of Pion Electromagnetic Form-Factor in QCD. *Theor. Math. Phys.*, 42:97–110, 1980.

[21] A. V. Efremov and A. V. Radyushkin. Factorization and Asymptotical Behavior of Pion Form-Factor in QCD. *Phys. Lett. B*, 94:245–250, 1980.

[22] Matthias Burkardt. Light front quantization. *Adv. Nucl. Phys.*, 23:1–74, 1996.

[23] Stanley J. Brodsky, Hans-Christian Pauli, and Stephen S. Pinsky. Quantum chromodynamics and other field theories on the light cone. *Phys. Rept.*, 301:299–486, 1998.

[24] Gerald A. Miller. Light front quantization: A Technique for relativistic and realistic nuclear physics. *Prog. Part. Nucl. Phys.*, 45:83–155, 2000.

[25] Xiangdong Ji and Yizhuang Liu. Computing light-front wave functions without light-front quantization: A large-momentum effective theory approach. *Phys. Rev. D*, 105(7):076014, 2022.

[26] John C. Collins, Davison E. Soper, and George F. Sterman. Factorization of Hard Processes in QCD. *Adv. Ser. Direct. High Energy Phys.*, 5:1–91, 1989.

[27] F. Herzog, B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt. The five-loop beta function of Yang-Mills theory with fermions. *JHEP*, 02:090, 2017.

[28] Georges Aad et al. Measurement of the double-differential high-mass Drell-Yan cross section in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *JHEP*, 08:009, 2016.

[29] Georges Aad et al. Determination of the parton distribution functions of the proton from ATLAS measurements of differential W and Z boson production in association with jets. *JHEP*, 07:223, 2021.

[30] Vardan Khachatryan et al. Measurement and QCD analysis of double-differential inclusive jet cross sections in pp collisions at $\sqrt{s} = 8$ TeV and cross section ratios to 2.76 and 7 TeV. *JHEP*, 03:156, 2017.

[31] Tie-Jiun Hou et al. New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC. *Phys. Rev. D*, 103(1):014013, 2021.

[32] S. Bailey, T. Cridge, L. A. Harland-Lang, A. D. Martin, and R. S. Thorne. Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs. *Eur. Phys. J. C*, 81(4):341, 2021.

[33] Richard D. Ball et al. The path to proton structure at 1% accuracy. *Eur. Phys. J. C*, 82(5):428, 2022.

[34] Richard D. Ball et al. The PDF4LHC21 combination of global PDF fits for the LHC Run III. *J. Phys. G*, 49(8):080501, 2022.

[35] 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. *Phys. Rev. D*, 93(3):033006, 2016.

[36] M. Beneke, G. Buchalla, M. Neubert, and Christopher T. Sachrajda. QCD factorization in B —> pi K, pi pi decays and extraction of Wolfenstein parameters. *Nucl. Phys. B*, 606:245–321, 2001.

[37] M. Beneke, G. Buchalla, M. Neubert, and Christopher T. Sachrajda. QCD factorization for B —> pi pi decays: Strong phases and CP violation in the heavy quark limit. *Phys. Rev. Lett.*, 83:1914–1917, 1999.

[38] Jing Gao, Tobias Huber, Yao Ji, and Yu-Ming Wang. Next-to-Next-to-Leading-Order QCD Prediction for the Photon-Pion Form Factor. *Phys. Rev. Lett.*, 128(6):062003, 2022.

[39] B. Melic, B. Nizic, and K. Passek. Complete next-to-leading order perturbative QCD prediction for the pion form-factor. *Phys. Rev. D*, 60:074004, 1999.

[40] J. Gronberg et al. Measurements of the meson - photon transition form-factors of light pseudoscalar mesons at large momentum transfer. *Phys. Rev. D*, 57:33–54, 1998.

[41] Bernard Aubert et al. Measurement of the gamma gamma* —> pi0 transition form factor. *Phys. Rev. D*, 80:052002, 2009.

[42] S. Uehara et al. Measurement of $\gamma\gamma^* \to \pi^0$ transition form factor at Belle. *Phys. Rev. D*, 86:092007, 2012.

[43] A. E. Dorokhov. Photon-Pion Transition Form Factor: BABAR Puzzle is Cracked. 3 2010.

[44] V. L. Chernyak and A. R. Zhitnitsky. Exclusive Decays of Heavy Mesons. *Nucl. Phys. B*, 201:492, 1982. [Erratum: Nucl.Phys.B 214, 547 (1983)].

[45] V. L. Chernyak and A. R. Zhitnitsky. Asymptotic Behavior of Exclusive Processes in QCD. *Phys. Rept.*, 112:173, 1984.

[46] Vladimir M. Braun and I. E. Filyanov. QCD Sum Rules in Exclusive Kinematics and Pion Wave Function. *Z. Phys. C*, 44:157, 1989.

[47] Lei Chang, I. C. Cloet, 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.

[48] Artur Avkhadiev. *Hadronic Structure with Classical and Quantum Computing*. PhD thesis, MIT, 2025.

[49] Renaud Boussarie et al. TMD Handbook. 4 2023.

[50] John C. Collins and Davison E. Soper. Parton Distribution and Decay Functions. *Nucl. Phys. B*, 194:445–492, 1982.

[51] John C. Collins, Davison E. Soper, and George F. Sterman. Transverse Momentum Distribution in Drell-Yan Pair and W and Z Boson Production. *Nucl. Phys. B*, 250:199–224, 1985.

[52] Aneesh V. Manohar and Iain W. Stewart. The Zero-Bin and Mode Factorization in Quantum Field Theory. *Phys. Rev. D*, 76:074002, 2007.

[53] John C. Collins and Davison E. Soper. Back-To-Back Jets: Fourier Transform from B to K-Transverse. *Nucl. Phys. B*, 197:446–476, 1982.

[54] Xiang-Dong Ji. Gauge-Invariant Decomposition of Nucleon Spin. *Phys. Rev. Lett.*, 78:610–613, 1997.

[55] Xiang-Dong Ji. Deeply virtual Compton scattering. *Phys. Rev. D*, 55:7114–7125, 1997.

[56] A. V. Radyushkin. Asymmetric gluon distributions and hard diffractive electroproduction. *Phys. Lett. B*, 385:333–342, 1996.

[57] John C. Collins, Leonid Frankfurt, and Mark Strikman. Factorization for hard exclusive electroproduction of mesons in QCD. *Phys. Rev. D*, 56:2982–3006, 1997.

[58] Kenneth G. Wilson. Confinement of Quarks. *Phys. Rev. D*, 10:2445–2459, 1974.

[59] K. Symanzik. Continuum Limit and Improved Action in Lattice Theories. 1. Principles and φ^4 Theory. *Nucl. Phys. B*, 226:187–204, 1983.

[60] M. Luscher and P. Weisz. Computation of the Action for On-Shell Improved Lattice Gauge Theories at Weak Coupling. *Phys. Lett. B*, 158:250–254, 1985.

[61] T. Reisz. Lattice Gauge Theory: Renormalization to All Orders in the Loop Expansion. *Nucl. Phys. B*, 318:417–463, 1989.

[62] M. Creutz. Monte Carlo Study of Quantized SU(2) Gauge Theory. *Phys. Rev. D*, 21:2308–2315, 1980.

[63] Don Weingarten. Masses and Decay Constants in Lattice QCD. *Nucl. Phys. B*, 215:1–22, 1983.

[64] H. Hamber, E. Marinari, G. Parisi, and C. Rebbi. Spectroscopy in a Lattice Gauge Theory. *Phys. Lett. B*, 108:314, 1982.

[65] H. Hamber and G. Parisi. Numerical Estimates of Hadronic Masses in a Pure SU(3) Gauge Theory. *Phys. Rev. Lett.*, 47:1792, 1981.

[66] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth. Hybrid Monte Carlo. *Phys. Lett. B*, 195:216–222, 1987.

[67] S. Aoki et al. FLAG Review 2019: Flavour Lattice Averaging Group (FLAG). *Eur. Phys. J. C*, 80(2):113, 2020.

[68] Xiangdong Ji. Parton Physics on a Euclidean Lattice. *Phys. Rev. Lett.*, 110:262002, 2013.

[69] Xiangdong Ji. Parton Physics from Large-Momentum Effective Field Theory. *Sci. China Phys. Mech. Astron.*, 57:1407–1412, 2014.

[70] Xiangdong Ji, Yu-Sheng Liu, Yizhuang Liu, Jian-Hui Zhang, and Yong Zhao. Large-momentum effective theory. *Rev. Mod. Phys.*, 93(3):035005, 2021.

[71] Xiangdong Ji. Euclidean effective theory for partons in the spirit of Steven Weinberg. *Nucl. Phys. B*, 1007:116670, 2024.

[72] Yong Zhao. Improving the Precision of First-Principles Calculation of Parton Physics from Lattice QCD. 8 2025.

[73] K. Symanzik. Continuum Limit and Improved Action in Lattice Theories. 2. O(N) Nonlinear Sigma Model in Perturbation Theory. *Nucl. Phys. B*, 226:205–227, 1983.

[74] Luuk H. Karsten and Jan Smit. Lattice Fermions: Species Doubling, Chiral Invariance, and the Triangle Anomaly. *Nucl. Phys. B*, 183:103, 1981.

[75] Holger Bech Nielsen and M. Ninomiya. No Go Theorem for Regularizing Chiral Fermions. *Phys. Lett. B*, 105:219–223, 1981.

[76] Kenneth G. Wilson. Quarks and Strings on a Lattice. In *13th International School of Subnuclear Physics: New Phenomena in Subnuclear Physics*, 11 1975.

[77] B. Sheikholeslami and R. Wohlert. Improved Continuum Limit Lattice Action for QCD with Wilson Fermions. *Nucl. Phys. B*, 259:572, 1985.

[78] John B. Kogut and Leonard Susskind. Hamiltonian Formulation of Wilson's Lattice Gauge Theories. *Phys. Rev. D*, 11:395–408, 1975.

[79] Leonard Susskind. Lattice Fermions. *Phys. Rev. D*, 16:3031–3039, 1977.

[80] Paul H. Ginsparg and Kenneth G. Wilson. A Remnant of Chiral Symmetry on the Lattice. *Phys. Rev. D*, 25:2649, 1982.

[81] Herbert Neuberger. Exactly massless quarks on the lattice. *Phys. Lett. B*, 417:141–144, 1998.

[82] David B. Kaplan. A Method for simulating chiral fermions on the lattice. *Phys. Lett. B*, 288:342–347, 1992.

[83] Yigal Shamir. Chiral fermions from lattice boundaries. *Nucl. Phys. B*, 406:90–106, 1993.

[84] Stephen L. Adler. Overrelaxation Algorithms for Lattice Field Theories. *Phys. Rev. D*, 37:458, 1988.

[85] Stephen L. Adler. An Overrelaxation Method for the Monte Carlo Evaluation of the Partition Function for Multiquadratic Actions. *Phys. Rev. D*, 23:2901, 1981.

[86] Frank R. Brown and Thomas J. Woch. Overrelaxed Heat Bath and Metropolis Algorithms for Accelerating Pure Gauge Monte Carlo Calculations. *Phys. Rev. Lett.*, 58:2394, 1987.

[87] G. Parisi and Yong-shi Wu. Perturbation Theory Without Gauge Fixing. *Sci. Sin.*, 24:483, 1981.

[88] Konstantin S Turitsyn, Michael Chertkov, and Marija Vucelja. Irreversible monte carlo algorithms for efficient sampling. *Physica D: Nonlinear Phenomena*, 240(4-5):410–414, 2011.

[89] Martin Luscher. Topology, the Wilson flow and the HMC algorithm. *PoS*, LATTICE2010:015, 2010.

[90] Stefan Schaefer, Rainer Sommer, and Francesco Virotta. Critical slowing down and error analysis in lattice QCD simulations. *Nucl. Phys. B*, 845:93–119, 2011.

[91] M. S. Albergo, G. Kanwar, and P. E. Shanahan. Flow-based generative models for Markov chain Monte Carlo in lattice field theory. *Phys. Rev. D*, 100(3):034515, 2019.

[92] Michael S. Albergo, Denis Boyda, Daniel C. Hackett, Gurtej Kanwar, Kyle Cranmer, Sébastien Racanière, Danilo Jimenez Rezende, and Phiala E. Shanahan. Introduction to Normalizing Flows for Lattice Field Theory. 1 2021.

[93] M. P. Mendenhall et al. Precision measurement of the neutron β-decay asymmetry. *Phys. Rev. C*, 87(3):032501, 2013.

[94] M. A. P. Brown et al. New result for the neutron β-asymmetry parameter A_0 from UCNA. *Phys. Rev. C*, 97(3):035505, 2018.

[95] D. Mund, B. Maerkisch, M. Deissenroth, J. Krempel, M. Schumann, H. Abele, A. Petoukhov, and T. Soldner. Determination of the Weak Axial Vector Coupling from a Measurement of the Beta-Asymmetry Parameter A in Neutron Beta Decay. *Phys. Rev. Lett.*, 110:172502, 2013.

[96] M. Ademollo and Raoul Gatto. Nonrenormalization Theorem for the Strangeness Violating Vector Currents. *Phys. Rev. Lett.*, 13:264–265, 1964.

[97] John F. Donoghue and D. Wyler. Isospin Breaking and the Precise Determination of $V(ud)$. *Phys. Lett. B*, 241:243–248, 1990.

[98] Y. Aoki et al. FLAG review 2024. *Phys. Rev. D*, 113(1):014508, 2026.

[99] R. L. Workman et al. Review of Particle Physics. *PTEP*, 2022:083C01, 2022.

[100] Jinchen He et al. Detailed analysis of excited-state systematics in a lattice QCD calculation of gA. *Phys. Rev. C*, 105(6):065203, 2022.

[101] Nolan Miller et al. Scale setting the Möbius domain wall fermion on gradient-flowed HISQ action using the omega baryon mass and the gradient-flow scales t_0 and w_0. *Phys. Rev. D*, 103(5):054511, 2021.

[102] Huey-Wen Lin et al. Parton distributions and lattice QCD calculations: a community white paper. *Prog. Part. Nucl. Phys.*, 100:107–160, 2018.

[103] Xiang Gao, Jinchen He, Joshua Lin, Swagato Mukherjee, Peter Petreczky, Rui Zhang, and Yong Zhao. Nucleon Parton Distribution Functions from Boosted Correlations in the Coulomb gauge. 2 2026.

[104] Xiaonu Xiong, Xiangdong Ji, Jian-Hui Zhang, and Yong Zhao. One-loop matching for parton distributions: Nonsinglet case. *Phys. Rev. D*, 90(1):014051, 2014.

[105] Huey-Wen Lin, Jiunn-Wei Chen, Saul D. Cohen, and Xiangdong Ji. Flavor Structure of the Nucleon Sea from Lattice QCD. *Phys. Rev. D*, 91:054510, 2015.

[106] Xiangdong Ji, Peng Sun, Xiaonu Xiong, and Feng Yuan. Soft factor subtraction and transverse momentum dependent parton distributions on the lattice. *Phys. Rev. D*, 91:074009, 2015.

[107] Xiangdong Ji, Jian-Hui Zhang, and Yong Zhao. Justifying the Naive Partonic Sum Rule for Proton Spin. *Phys. Lett. B*, 743:180–183, 2015.

[108] Raza Sabbir Sufian, Michael J. Glatzmaier, Yi-Bo Yang, Keh-Fei Liu, and Mingyang Sun. Glue Spin S_G in The Longitudinally Polarized Nucleon. *PoS*, LATTICE2014:166, 2015.

[109] Xiangdong Ji, Andreas Schäfer, Xiaonu Xiong, and Jian-Hui Zhang. One-Loop Matching for Generalized Parton Distributions. *Phys. Rev. D*, 92:014039, 2015.

[110] Xiaonu Xiong and Jian-Hui Zhang. One-loop matching for transversity generalized parton distribution. *Phys. Rev. D*, 92(5):054037, 2015.

[111] Constantia Alexandrou, Krzysztof Cichy, Martha Constantinou, Kyriakos Hadjiyiannakou, Karl Jansen, Fernanda Steffens, and Christian Wiese. Parton Distributions from Lattice QCD with Momentum Smearing. *PoS*, LATTICE2016:151, 2016.

[112] Jiunn-Wei Chen, Saul D. Cohen, Xiangdong Ji, Huey-Wen Lin, and Jian-Hui Zhang. Nucleon Helicity and Transversity Parton Distributions from Lattice QCD. *Nucl. Phys. B*, 911:246–273, 2016.

[113] Huey-Wen Lin. From C to Parton Sea: Bjorken-x Dependence of the PDFs. *PoS*, LATTICE2016:005, 2016.

[114] Yi-Bo Yang, Raza Sabbir Sufian, Andrei Alexandru, Terrence Draper, Michael J. Glatzmaier, and Keh-Fei Liu. Glue Spin of the Proton. *PoS*, LATTICE2015:129, 2016.

[115] Yi-Bo Yang, Raza Sabbir Sufian, Andrei Alexandru, Terrence Draper, Michael J. Glatzmaier, Keh-Fei Liu, and Yong Zhao. Glue Spin and Helicity in the Proton from Lattice QCD. *Phys. Rev. Lett.*, 118(10):102001, 2017.

[116] Constantia Alexandrou, Simone Bacchio, Krzysztof Cichy, Martha Constantinou, Kyriakos Hadjiyiannakou, Karl Jansen, Giannis Koutsou, Aurora Scapellato, and Fernanda Steffens. Computation of parton distributions from the quasi-PDF approach at the physical point. *EPJ Web Conf.*, 175:14008, 2018.

[117] Constantia Alexandrou, Krzysztof Cichy, Martha Constantinou, Kyriakos Hadjiyiannakou, Karl Jansen, Haralambos Panagopoulos, and Fernanda Steffens. A complete non-perturbative renormalization prescription for quasi-PDFs. *Nucl. Phys. B*, 923:394–415, 2017.

[118] Constantia Alexandrou, Krzysztof Cichy, Martha Constantinou, Kyriakos Hadjiyiannakou, Karl Jansen, Haralambos Panagopoulos, Aurora Scapellato, and Fernanda Steffens. Progress in computing parton distribution functions from the quasi-PDF approach. *EPJ Web Conf.*, 175:06021, 2018.

[119] Jiunn-Wei Chen, Tomomi Ishikawa, Luchang Jin, Huey-Wen Lin, Jian-Hui Zhang, and Yong Zhao. Symmetry properties of nonlocal quark bilinear operators on a Lattice. *Chin. Phys. C*, 43(10):103101, 2019.

[120] Jiunn-Wei Chen, Tomomi Ishikawa, Luchang Jin, Huey-Wen Lin, Yi-Bo Yang, Jian-Hui Zhang, and Yong Zhao. Parton distribution function with nonperturbative renormalization from lattice QCD. *Phys. Rev. D*, 97(1):014505, 2018.

[121] Martha Constantinou and Haralambos Panagopoulos. Perturbative renormalization of quasi-parton distribution functions. *Phys. Rev. D*, 96(5):054506, 2017.

[122] Tomomi Ishikawa, LuChang Jin, Huey-Wen Lin, Andreas Schäfer, Yi-Bo Yang, Jian-Hui Zhang, and Yong Zhao. Gaussian-weighted parton quasi-distribution (Lattice Parton Physics Project (LP^3)). *Sci. China Phys. Mech. Astron.*, 62(9):991021, 2019.

[123] Xiangdong Ji, Jian-Hui Zhang, and Yong Zhao. Renormalization in Large Momentum Effective Theory of Parton Physics. *Phys. Rev. Lett.*, 120(11):112001, 2018.

[124] Xiangdong Ji, Jian-Hui Zhang, and Yong Zhao. More On Large-Momentum Effective Theory Approach to Parton Physics. *Nucl. Phys. B*, 924:366–376, 2017.

[125] Huey-Wen Lin, Jiunn-Wei Chen, Tomomi Ishikawa, and Jian-Hui Zhang. Improved parton distribution functions at the physical pion mass. *Phys. Rev. D*, 98(5):054504, 2018.

[126] Iain W. Stewart and Yong Zhao. Matching the quasiparton distribution in a momentum subtraction scheme. *Phys. Rev. D*, 97(5):054512, 2018.

[127] Xiaonu Xiong, Thomas Luu, and Ulf-G. Meißner. Quasi-Parton Distribution Function in Lattice Perturbation Theory. 4 2017.

[128] Jian-Hui Zhang, Jiunn-Wei Chen, Xiangdong Ji, Luchang Jin, and Huey-Wen Lin. Pion Distribution Amplitude from Lattice QCD. *Phys. Rev. D*, 95(9):094514, 2017.

[129] Jian-Hui Zhang, Luchang Jin, Huey-Wen Lin, Andreas Schäfer, Peng Sun, Yi-Bo Yang, Rui Zhang, Yong Zhao, and Jiunn-Wei Chen. Kaon Distribution Amplitude from Lattice QCD and the Flavor SU(3) Symmetry. *Nucl. Phys. B*, 939:429–446, 2019.

[130] Constantia Alexandrou, Krzysztof Cichy, Martha Constantinou, Karl Jansen, Aurora Scapellato, and Fernanda Steffens. Transversity parton distribution functions from lattice QCD. *Phys. Rev. D*, 98(9):091503, 2018.

[131] Constantia Alexandrou, Krzysztof Cichy, Martha Constantinou, Karl Jansen, Aurora Scapellato, and Fernanda Steffens. Light-Cone Parton Distribution Functions from Lattice QCD. *Phys. Rev. Lett.*, 121(11):112001, 2018.

[132] Constantia Alexandrou, Krzysztof Cichy, Martha Constantinou, Kyriakos Hadjiyiannakou, Karl Jansen, Aurora Scapellato, and Fernanda Steffens. Light-cone PDFs from Lattice QCD. *PoS*, LATTICE2018:094, 2018.

[133] Jiunn-Wei Chen, Luchang Jin, Huey-Wen Lin, Yu-Sheng Liu, Yi-Bo Yang, Jian-Hui Zhang, and Yong Zhao. Lattice Calculation of Parton Distribution Function from LaMET at Physical Pion Mass with Large Nucleon Momentum. 3 2018.

[134] Markus A. Ebert, Iain W. Stewart, and Yong Zhao. Determining the Nonperturbative Collins-Soper Kernel From Lattice QCD. *Phys. Rev. D*, 99(3):034505, 2019.

[135] Zhou-You Fan, Yi-Bo Yang, Adam Anthony, Huey-Wen Lin, and Keh-Fei Liu. Gluon Quasi-Parton-Distribution Functions from Lattice QCD. *Phys. Rev. Lett.*, 121(24):242001, 2018.

[136] Taku Izubuchi, Xiangdong Ji, Luchang Jin, Iain W. Stewart, and Yong Zhao. Factorization Theorem Relating Euclidean and Light-Cone Parton Distributions. *Phys. Rev. D*, 98(5):056004, 2018.

[137] Xiangdong Ji, Lu-Chang Jin, Feng Yuan, Jian-Hui Zhang, and Yong Zhao. Transverse momentum dependent parton quasidistributions. *Phys. Rev. D*, 99(11):114006, 2019.

[138] Yu-Sheng Liu et al. Unpolarized isovector quark distribution function from lattice QCD: A systematic analysis of renormalization and matching. *Phys. Rev. D*, 101(3):034020, 2020.

[139] Huey-Wen Lin, Jiunn-Wei Chen, Xiangdong Ji, Luchang Jin, Ruizi Li, Yu-Sheng Liu, Yi-Bo Yang, Jian-Hui Zhang, and Yong Zhao. Proton Isovector Helicity Distribution on the Lattice at Physical Pion Mass. *Phys. Rev. Lett.*, 121(24):242003, 2018.

[140] Yu-Sheng Liu, Jiunn-Wei Chen, Luchang Jin, Ruizi Li, Huey-Wen Lin, Yi-Bo Yang, Jian-Hui Zhang, and Yong Zhao. Nucleon Transversity Distribution at the Physical Pion Mass from Lattice QCD. 10 2018.

[141] Yu-Sheng Liu, Wei Wang, Ji Xu, Qi-An Zhang, Shuai Zhao, and Yong Zhao. Matching the meson quasidistribution amplitude in the RI/MOM scheme. *Phys. Rev. D*, 99(9):094036, 2019.

[142] Jian-Hui Zhang, Xiangdong Ji, Andreas Schäfer, Wei Wang, and Shuai Zhao. Accessing Gluon Parton Distributions in Large Momentum Effective Theory. *Phys. Rev. Lett.*, 122(14):142001, 2019.

[143] 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. D*, 100(3):034505, 2019.

[144] Jianhui Zhang, Jiunn-Wei Chen, Luchang Jin, Huey-Wen Lin, Yu-Sheng Liu, Andreas Schäfer, Yi-Bo Yang, and Yong Zhao. Direct lattice-QCD calculation of pion valence quark distribution. *PoS*, LATTICE2018:108, 2018.

[145] Yong Zhao. Unraveling high-energy hadron structures with lattice QCD. *Int. J. Mod. Phys. A*, 33(36):1830033, 2019.

[146] Constantia Alexandrou, Krzysztof Cichy, Martha Constantinou, Kyriakos Hadjiyiannakou, Karl Jansen, Aurora Scapellato, and Fernanda Steffens. Quasi-PDFs with Twisted Mass Fermions. *PoS*, LATTICE2019:036, 2019.

[147] Constantia Alexandrou, Krzysztof Cichy, Martha Constantinou, Kyriakos Hadjiyiannakou, Karl Jansen, Aurora Scapellato, and Fernanda Steffens. Systematic uncertainties in parton distribution functions from lattice QCD simulations at the physical point. *Phys. Rev. D*, 99(11):114504, 2019.

[148] Yahui Chai et al. Parton distribution functions of Δ^+ on the lattice. *PoS*, LATTICE2019:270, 2019.

[149] Jiunn-Wei Chen, Huey-Wen Lin, and Jian-Hui Zhang. Pion generalized parton distribution from lattice QCD. *Nucl. Phys. B*, 952:114940, 2020.

[150] Martha Constantinou, Haralambos Panagopoulos, and Gregoris Spanoudes. One-loop renormalization of staple-shaped operators in continuum and lattice regularizations. *Phys. Rev. D*, 99(7):074508, 2019.

[151] Markus A. Ebert, Iain W. Stewart, and Yong Zhao. Renormalization and Matching for the Collins-Soper Kernel from Lattice QCD. *JHEP*, 03:099, 2020.

[152] Yu-Sheng Liu, Wei Wang, Ji Xu, Qi-An Zhang, Jian-Hui Zhang, Shuai Zhao, and Yong Zhao. Matching generalized parton quasidistributions in the RI/MOM scheme. *Phys. Rev. D*, 100(3):034006, 2019.

[153] Phiala Shanahan, Michael L. Wagman, and Yong Zhao. Nonperturbative renormalization of staple-shaped Wilson line operators in lattice QCD. *Phys. Rev. D*, 101(7):074505, 2020.

[154] Wei Wang, Yu-Ming Wang, Ji Xu, and Shuai Zhao. B-meson light-cone distribution amplitude from Euclidean quantities. *Phys. Rev. D*, 102(1):011502, 2020.

[155] Rui Zhang, Zhouyou Fan, Ruizi Li, Huey-Wen Lin, and Boram Yoon. Machine-learning prediction for quasiparton distribution function matrix elements. *Phys. Rev. D*, 101(3):034516, 2020.

[156] Constantia Alexandrou, Krzysztof Cichy, Martha Constantinou, Jeremy R. Green, Kyriakos Hadjiyiannakou, Karl Jansen, Floriano Manigrasso, Aurora Scapellato, and Fernanda Steffens. Lattice continuum-limit study of nucleon quasi-PDFs. *Phys. Rev. D*, 103:094512, 2021.

[157] Constantia Alexandrou, Martha Constantinou, Kyriakos Hadjiyiannakou, Karl Jansen, and Floriano Manigrasso. Flavor decomposition for the proton helicity parton distribution functions. *Phys. Rev. Lett.*, 126(10):102003, 2021.

[158] Constantia Alexandrou, Krzysztof Cichy, Martha Constantinou, Kyriakos Hadjiyiannakou, Karl Jansen, Aurora Scapellato, and Fernanda Steffens. Unpolarized and helicity generalized parton distributions of the proton within lattice QCD. *Phys. Rev. Lett.*, 125(26):262001, 2020.

[159] Shohini Bhattacharya, Krzysztof Cichy, Martha Constantinou, Andreas Metz, Aurora Scapellato, and Fernanda Steffens. Insights on proton structure from lattice QCD: The twist-3 parton distribution function $g_T(x)$. *Phys. Rev. D*, 102(11):111501, 2020.

[160] Shohini Bhattacharya, Krzysztof Cichy, Martha Constantinou, Andreas Metz, Aurora Scapellato, and Fernanda Steffens. The role of zero-mode contributions in the matching for the twist-3 PDFs $e(x)$ and $h_L(x)$. *Phys. Rev. D*, 102:114025, 2020.

[161] Yahui Chai et al. Parton distribution functions of Δ^+ on the lattice. *Phys. Rev. D*, 102(1):014508, 2020.

[162] Long-Bin Chen, Wei Wang, and Ruilin Zhu. Quasi parton distribution functions at NNLO: flavor non-diagonal quark contributions. *Phys. Rev. D*, 102(1):011503, 2020.

[163] Long-Bin Chen, Wei Wang, and Ruilin Zhu. Master integrals for two-loop QCD corrections to quark quasi PDFs. *JHEP*, 10:079, 2020.

[164] Long-Bin Chen, Wei Wang, and Ruilin Zhu. Next-to-Next-to-Leading Order Calculation of Quasiparton Distribution Functions. *Phys. Rev. Lett.*, 126(7):072002, 2021.

[165] Markus A. Ebert, Stella T. Schindler, Iain W. Stewart, and Yong Zhao. One-loop Matching for Spin-Dependent Quasi-TMDs. *JHEP*, 09:099, 2020.

[166] Zhouyou Fan, Xiang Gao, Ruizi Li, Huey-Wen Lin, Nikhil Karthik, Swagato Mukherjee, Peter Petreczky, Sergey Syritsyn, Yi-Bo Yang, and Rui Zhang. Isovector parton distribution functions of the proton on a superfine lattice. *Phys. Rev. D*, 102(7):074504, 2020.

[167] 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. *Phys. Rev. D*, 102(9):094513, 2020.

[168] Jun Hua, Min-Huan Chu, Peng Sun, Wei Wang, Ji Xu, Yi-Bo Yang, Jian-Hui Zhang, and Qi-An Zhang. Distribution Amplitudes of K* and ϕ at the Physical Pion Mass from Lattice QCD. *Phys. Rev. Lett.*, 127(6):062002, 2021.

[169] Xiangdong Ji, Yizhuang Liu, Andreas Schäfer, Wei Wang, Yi-Bo Yang, Jian-Hui Zhang, and Yong Zhao. A Hybrid Renormalization Scheme for Quasi Light-Front Correlations in Large-Momentum Effective Theory. *Nucl. Phys. B*, 964:115311, 2021.

[170] Xiangdong Ji, Yizhuang Liu, Andreas Schäfer, and Feng Yuan. Single Transverse-Spin Asymmetry and Sivers Function in Large Momentum Effective Theory. *Phys. Rev. D*, 103(7):074005, 2021.

[171] Qi-An Zhang et al. Lattice-QCD Calculations of TMD Soft Function Through Large-Momentum Effective Theory. *Phys. Rev. Lett.*, 125(19):192001, 2020.

[172] Huey-Wen Lin, Jiunn-Wei Chen, and Rui Zhang. Lattice Nucleon Isovector Unpolarized Parton Distribution in the Physical-Continuum Limit. 11 2020.

[173] Huey-Wen Lin. Nucleon Tomography and Generalized Parton Distribution at Physical Pion Mass from Lattice QCD. *Phys. Rev. Lett.*, 127(18):182001, 2021.

[174] 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. *Phys. Rev. D*, 103(1):014516, 2021.

[175] Phiala Shanahan, Michael Wagman, and Yong Zhao. Collins-Soper kernel for TMD evolution from lattice QCD. *Phys. Rev. D*, 102(1):014511, 2020.

[176] Charles Shugert, Xiang Gao, Taku Izubichi, Luchang Jin, Christos Kallidonis, Nikhil Karthik, Swagato Mukherjee, Peter Petreczky, Sergey Syritsyn, and Yong Zhao. Pion valence quark PDF from lattice QCD. In *37th International Symposium on Lattice Field Theory*, 1 2020.

[177] Alexey A. Vladimirov and Andreas Schäfer. Transverse momentum dependent factorization for lattice observables. *Phys. Rev. D*, 101(7):074517, 2020.

[178] Rui Zhang, Huey-Wen Lin, and Boram Yoon. Probing nucleon strange and charm distributions with lattice QCD. *Phys. Rev. D*, 104(9):094511, 2021.

[179] Rui Zhang, Carson Honkala, Huey-Wen Lin, and Jiunn-Wei Chen. Pion and kaon distribution amplitudes in the continuum limit. *Phys. Rev. D*, 102(9):094519, 2020.

[180] Kuan Zhang, Yuan-Yuan Li, Yi-Kai Huo, Andreas Schäfer, Peng Sun, and Yi-Bo Yang. RI/MOM renormalization of the parton quasidistribution functions in lattice regularization. *Phys. Rev. D*, 104(7):074501, 2021.

[181] Constantia Alexandrou, Krzysztof Cichy, Martha Constantinou, Kyriakos Hadjiyiannakou, Karl Jansen, Aurora Scapellato, and Fernanda Steffens. Transversity GPDs of the proton from lattice QCD. *Phys. Rev. D*, 105(3):034501, 2022.

[182] Constantia Alexandrou, Martha Constantinou, Kyriakos Hadjiyiannakou, Karl Jansen, and Floriano Manigrasso. Flavor decomposition of the nucleon unpolarized, helicity, and transversity parton distribution functions from lattice QCD simulations. *Phys. Rev. D*, 104(5):054503, 2021.

[183] Shohini Bhattacharya, Krzysztof Cichy, Martha Constantinou, Andreas Metz, Aurora Scapellato, and Fernanda Steffens. Parton distribution functions beyond leading twist from lattice QCD: The hL(x) case. *Phys. Rev. D*, 104(11):114510, 2021.

[184] Shohini Bhattacharya, Krzysztof Cichy, Martha Constantinou, Andreas Metz, Aurora Scapellato, and Fernanda Steffens. Twist-3 partonic distributions from lattice QCD. *SciPost Phys. Proc.*, 8:057, 2022.

[185] Martha Constantinou, Shohini Bhattacharya, Krzysztof Cichy, Andreas Metz, Aurora Scapellato, and Fernanda Steffens. First study of twist-3 PDFs for the proton from lattice QCD. *PoS*, LATTICE2021:391, 2022.

[186] Jack Dodson, Shohini Bhattacharya, Krzysztof Cichy, Martha Constantinou, Andreas Metz, Aurora Scapellato, and Fernanda Steffens. First Lattice QCD Study of Proton Twist-3 GPDs. *PoS*, LATTICE2021:054, 2022.

[187] Xiang Gao, Kyle Lee, Swagato Mukherjee, Charles Shugert, and Yong Zhao. Origin and resummation of threshold logarithms in the lattice QCD calculations of PDFs. *Phys. Rev. D*, 103(9):094504, 2021.

[188] Xiang Gao, Andrew D. Hanlon, Swagato Mukherjee, Peter Petreczky, Philipp Scior, Sergey Syritsyn, and Yong Zhao. Lattice QCD Determination of the Bjorken-x Dependence of Parton Distribution Functions at Next-to-Next-to-Leading Order. *Phys. Rev. Lett.*, 128(14):142003, 2022.

[189] Yi-Kai Huo et al. Self-renormalization of quasi-light-front correlators on the lattice. *Nucl. Phys. B*, 969:115443, 2021.

[190] Yuan Li et al. Lattice QCD Study of Transverse-Momentum Dependent Soft Function. *Phys. Rev. Lett.*, 128(6):062002, 2022.

[191] Huey-Wen Lin. Nucleon helicity generalized parton distribution at physical pion mass from lattice QCD. *Phys. Lett. B*, 824:136821, 2022.

[192] Huey-Wen Lin. Recent Progress in Lattice Parton-Distribution Calculations. *PoS*, LATTICE2021:276, 2022.

[193] Aurora Scapellato, Constantia Alexandrou, Krzysztof Cichy, Martha Constantinou, Kyriakos Hadjiyiannakou, Karl Jansen, and Fernanda Steffens. Generalized parton distributions of the proton from lattice QCD. *PoS*, LATTICE2021:129, 2022.

[194] Maximilian Schlemmer, Alexey Vladimirov, Christian Zimmermann, Michael Engelhardt, and Andreas Schäfer. Determination of the Collins-Soper Kernel from Lattice QCD. *JHEP*, 08:004, 2021.

[195] Phiala Shanahan, Michael Wagman, and Yong Zhao. Lattice QCD calculation of the Collins-Soper kernel from quasi-TMDPDFs. *Phys. Rev. D*, 104(11):114502, 2021.

[196] Shohini Bhattacharya, Krzysztof Cichy, Martha Constantinou, Jack Dodson, Xiang Gao, Andreas Metz, Swagato Mukherjee, Aurora Scapellato, Fernanda Steffens, and Yong Zhao. Generalized parton distributions from lattice QCD with asymmetric momentum transfer: Unpolarized quarks. *Phys. Rev. D*, 106(11):114512, 2022.

[197] Martha Constantinou, Shohini Bhattacharya, Krzysztof Cichy, Jack Dodson, Xiang Gao, Andreas Metz, Swagato Mukherjee, Aurora Scapellato, Fernanda Steffens, and Yong Zhao. Accessing proton GPDs in asymmetric frames: Numerical implementation. *PoS*, LATTICE2022:096, 2023.

[198] Zhi-Fu Deng, Wei Wang, and Jun Zeng. Transverse-momentum-dependent wave functions and soft functions at one-loop in large momentum effective theory. *JHEP*, 09:046, 2022.

[199] Markus A. Ebert, Stella T. Schindler, Iain W. Stewart, and Yong Zhao. Factorization connecting continuum & lattice TMDs. *JHEP*, 04:178, 2022.

[200] Xiang Gao, Andrew D. Hanlon, Nikhil Karthik, Swagato Mukherjee, Peter Petreczky, Philipp Scior, Shuzhe Shi, Sergey Syritsyn, Yong Zhao, and Kai Zhou. Continuum-extrapolated NNLO valence PDF of the pion at the physical point. *Phys. Rev. D*, 106(11):114510, 2022.

[201] Xiang Gao, Andrew D. Hanlon, Jack Holligan, Nikhil Karthik, Swagato Mukherjee, Peter Petreczky, Sergey Syritsyn, and Yong Zhao. Unpolarized proton PDF at NNLO from lattice QCD with physical quark masses. *Phys. Rev. D*, 107(7):074509, 2023.

[202] Fei Yao et al. Nucleon Transversity Distribution in the Continuum and Physical Mass Limit from Lattice QCD. *Phys. Rev. Lett.*, 131(26):261901, 2023.

[203] Jun Hua et al. Pion and Kaon Distribution Amplitudes from Lattice QCD. *Phys. Rev. Lett.*, 129(13):132001, 2022.

[204] Jin-Chen He, Min-Huan Chu, Jun Hua, Xiangdong Ji, Andreas Schäfer, Yushan Su, Wei Wang, Yi-Bo Yang, Jian-Hui Zhang, and Qi-An Zhang. Unpolarized transverse momentum dependent parton distributions of the nucleon from lattice QCD. *Phys. Rev. D*, 109(11):114513, 2024.

[205] Min-Huan Chu et al. Nonperturbative determination of the Collins-Soper kernel from quasitransverse-momentum-dependent wave functions. *Phys. Rev. D*, 106(3):034509, 2022.

[206] Aurora Scapellato, Constantia Alexandrou, Krzysztof Cichy, Martha Constantinou, Kyriakos Hadjiyiannakou, Karl Jansen, and Fernanda Steffens. Proton generalized parton distributions from lattice QCD. *Rev. Mex. Fis. Suppl.*, 3(3):0308104, 2022.

[207] Stella T. Schindler, Iain W. Stewart, and Yong Zhao. One-loop matching for gluon lattice TMDs. *JHEP*, 08:084, 2022.

[208] Kuan Zhang, Xiangdong Ji, Yi-Bo Yang, Fei Yao, and Jian-Hui Zhang. Renormalization of Transverse-Momentum-Dependent Parton Distribution on the Lattice. *Phys. Rev. Lett.*, 129(8):082002, 2022.

[209] Constantia Alexandrou et al. Nonperturbative renormalization of asymmetric staple-shaped operators in twisted mass lattice QCD. *Phys. Rev. D*, 108(11):114503, 2023.

[210] Artur Avkhadiev, Phiala E. Shanahan, Michael L. Wagman, and Yong Zhao. Collins-Soper kernel from lattice QCD at the physical pion mass. *Phys. Rev. D*, 108(11):114505, 2023.

[211] Shohini Bhattacharya et al. Generalized parton distributions from lattice QCD with asymmetric momentum transfer: Axial-vector case. *Phys. Rev. D*, 109(3):034508, 2024.

[212] Shohini Bhattacharya, Krzysztof Cichy, Martha Constantinou, Jack Dodson, Andreas Metz, Aurora Scapellato, and Fernanda Steffens. Chiral-even axial twist-3 GPDs of the proton from lattice QCD. *Phys. Rev. D*, 108(5):054501, 2023.

[213] Shohini Bhattacharya, Krzysztof Cichy, Martha Constantinou, Jack Dodson, Xiang Gao, Andreas Metz, Swagato Mukherjee, Aurora Scapellato, Fernanda Steffens, and Yong Zhao. GPDs in asymmetric frames. *PoS*, LATTICE2022:095, 2023.

[214] Krzysztof Cichy et al. Generalized Parton Distributions from Lattice QCD. *Acta Phys. Polon. Supp.*, 16(7):7–A6, 2023.

[215] Zhi-Fu Deng, Chao Han, Wei Wang, Jun Zeng, and Jia-Lu Zhang. Light-cone distribution amplitudes of a light baryon in large-momentum effective theory. *JHEP*, 07:191, 2023.

[216] Xiang Gao, Andrew D. Hanlon, Swagato Mukherjee, Peter Petreczky, Qi Shi, Sergey Syritsyn, and Yong Zhao. Transversity PDFs of the proton from lattice QCD with physical quark masses. *Phys. Rev. D*, 109(5):054506, 2024.

[217] Xiang Gao, Wei-Yang Liu, and Yong Zhao. Parton distributions from boosted fields in the Coulomb gauge. *Phys. Rev. D*, 109(9):094506, 2024.

[218] Jack Holligan and Huey-Wen Lin. Systematic improvement of x-dependent unpolarized nucleon generalized parton distributions in lattice-QCD calculation. *Phys. Rev. D*, 110(3):034503, 2024.

[219] Jack Holligan, Xiangdong Ji, Huey-Wen Lin, Yushan Su, and Rui Zhang. Precision control in lattice calculation of x-dependent pion distribution amplitude. *Nucl. Phys. B*, 993:116282, 2023.

[220] Xiangdong Ji, Yizhuang Liu, and Yushan Su. Threshold resummation for computing large-x parton distribution through large-momentum effective theory. *JHEP*, 08:037, 2023.

[221] Min-Huan Chu et al. Transverse-momentum-dependent wave functions of the pion from lattice QCD. *Phys. Rev. D*, 109(9):L091503, 2024.

[222] Min-Huan Chu et al. Lattice calculation of the intrinsic soft function and the Collins-Soper kernel. *JHEP*, 08:172, 2023.

[223] Huey-Wen Lin. Pion valence-quark generalized parton distribution at physical pion mass. *Phys. Lett. B*, 846:138181, 2023.

[224] Yong Zhao. Transverse Momentum Distributions from Lattice QCD without Wilson Lines. *Phys. Rev. Lett.*, 133(24):241904, 2024.

[225] Yizhuang Liu and Yushan Su. Renormalon cancellation and linear power correction to threshold-like asymptotics of space-like parton correlators. *JHEP*, 2024:204, 2024.

[226] Artur Avkhadiev, Phiala E. Shanahan, Michael L. Wagman, and Yong Zhao. Determination of the Collins-Soper Kernel from Lattice QCD. *Phys. Rev. Lett.*, 132(23):231901, 2024.

[227] Ethan Baker, Dennis Bollweg, Peter Boyle, Ian Cloët, Xiang Gao, Swagato Mukherjee, Peter Petreczky, Rui Zhang, and Yong Zhao. Lattice QCD calculation of the pion distribution amplitude with domain wall fermions at physical pion mass. *JHEP*, 07:211, 2024.

[228] Dennis Bollweg, Xiang Gao, Swagato Mukherjee, and Yong Zhao. Nonperturbative Collins-Soper kernel from chiral quarks with physical masses. *Phys. Lett. B*, 852:138617, 2024.

[229] Chen Chen, Yiqi Geng, Liuming Liu, Peng Sun, Yi-Bo Yang, Fei Yao, Jian-Hui Zhang, and Kuan Zhang. Parton distribution function of a deuteronlike dibaryon system from lattice QCD. *Phys. Rev. D*, 111(7):074506, 2025.

[230] Ian Cloet, Xiang Gao, Swagato Mukherjee, Sergey Syritsyn, Nikhil Karthik, Peter Petreczky, Rui Zhang, and Yong Zhao. Lattice QCD calculation of x-dependent meson distribution amplitudes at physical pion mass with threshold logarithm resummation. *Phys. Rev. D*, 110(11):114502, 2024.

[231] Heng-Tong Ding, Xiang Gao, Swagato Mukherjee, Peter Petreczky, Qi Shi, Sergey Syritsyn, and Yong Zhao. Three-dimensional imaging of pion using lattice QCD: generalized parton distributions. *JHEP*, 02:056, 2025.

[232] Xiang Gao, Jinchen He, Rui Zhang, and Yong Zhao. Systematic Uncertainties from Gribov Copies in Lattice Calculation of Parton Distributions in the Coulomb Gauge. *Chin. Phys. Lett.*, 41(12):121201, 2024.

[233] William Good, Kinza Hasan, and Huey-Wen Lin. Toward the first gluon parton distribution from the LaMET. *J. Phys. G*, 52(3):035105, 2025.

[234] Xue-Ying Han, Jun Hua, Xiangdong Ji, Cai-Dian Lü, Wei Wang, Ji Xu, Qi-An Zhang, and Shuai Zhao. Realistic method to access heavy meson light-cone distribution amplitudes from first-principle. *Phys. Rev. D*, 111(11):L111503, 2025.

[235] Jack Holligan and Huey-Wen Lin. Pion valence quark distribution at physical pion mass of $N_f = 2 + 1 + 1$ lattice QCD. *J. Phys. G*, 51(6):065101, 2024.

[236] Jack Holligan and Huey-Wen Lin. Nucleon helicity parton distribution function in the continuum limit with self-renormalization. *Phys. Lett. B*, 854:138731, 2024.

[237] Xiangdong Ji, Yizhuang Liu, Yushan Su, and Rui Zhang. Effects of threshold resummation for large-x PDF in large momentum effective theory. *JHEP*, 03:045, 2025.

[238] Lisa Walter et al. Quark transverse spin-momentum correlation of the pion from lattice QCD: The Boer-Mulders function. *Phys. Rev. D*, 111(9):094507, 2025.

[239] Min-Huan Chu et al. Light cone distribution amplitude for the Λ baryon from lattice QCD. *Phys. Rev. D*, 111(3):034510, 2025.

[240] Xue-Ying Han et al. Calculation of heavy meson light-cone distribution amplitudes from lattice QCD. *Phys. Rev. D*, 111(3):034503, 2025.

[241] Joshua Miller, Shohini Bhattacharya, Krzysztof Cichy, Martha Constantinou, Xiang Gao, Andreas Metz, Swagato Mukherjee, Peter Petreczky, Fernanda Steffens, and Yong Zhao. Proton Helicity GPDs from Lattice QCD. *PoS*, LATTICE2023:310, 2024.

[242] Swagato Mukherjee, Dennis Bollweg, Xiang Gao, and Yong Zhao. Non-perturbative Collins-Soper kernel: Chiral quarks and Coulomb-gauge-fixed quasi-TMD. *PoS*, DIS2024:238, 2025.

[243] Gregoris Spanoudes, Martha Constantinou, and Haralambos Panagopoulos. Renormalization of asymmetric staple-shaped Wilson-line operators in lattice and continuum perturbation theory. *Phys. Rev. D*, 109(11):114501, 2024.

[244] Kuan Zhang, Yi-Kai Huo, Xiangdong Ji, Andreas Schaefer, Chun-Jiang Shi, Peng Sun, Wei Wang, Yi-Bo Yang, and Jian-Hui Zhang. Impact of gauge fixing precision on the continuum limit of nonlocal quark-bilinear lattice operators. *Phys. Rev. D*, 110(7):074505, 2024.

[245] Dennis Bollweg, Xiang Gao, Swagato Mukherjee, and Yong Zhao. Lattice QCD Benchmark of Proton Helicity and Flavor-Dependent Unpolarized Transverse Momentum-Dependent Parton Distribution Functions at Physical Quark Masses. *Phys. Rev. Lett.*, 135(20):201901, 2025.

[246] Dennis Bollweg, Xiang Gao, Jinchen He, Swagato Mukherjee, and Yong Zhao. Transverse-momentum-dependent pion structures from lattice QCD: Collins-Soper kernel, soft factor, TMDWF, and TMDPDF. *Phys. Rev. D*, 112(3):034501, 2025.

[247] Wei Wang et al. Heavy meson lightcone distribution amplitudes from Lattice QCD. *PoS*, QCHSC24:111, 2025.

[248] Jack Holligan, Huey-Wen Lin, Rui Zhang, and Yong Zhao. Resummation for lattice QCD calculation of generalized parton distributions at nonzero skewness. *JHEP*, 207:241, 2025.

[249] Lingquan Ma et al. Quark transverse spin-momentum correlation of the nucleon from lattice QCD: the Boer-Mulders function. *JHEP*, 08:086, 2025.

[250] Qi-An Zhang, Jin-Chen He, Min-Huan Chu, Jun Hua, Xiangdong Ji, Andreas Schäfer, Yushan Su, Wei Wang, Yi-Bo Yang, and Jian-Hui Zhang. Lattice QCD determination of unpolarized TMDPDFs of the nucleon. *PoS*, QNP2024:052, 2025.

[251] Yao Ji, Fei Yao, and Jian-Hui Zhang. Extracting Meson Distribution Amplitudes from Nonlocal Euclidean Correlations at Next-to-Next-to-Leading Order. 4 2025.

[252] Xiangdong Ji and Jian-Hui Zhang. Renormalization of quasiparton distribution. *Phys. Rev. D*, 92:034006, 2015.

[253] Jiunn-Wei Chen, Xiangdong Ji, and Jian-Hui Zhang. Improved quasi parton distribution through Wilson line renormalization. *Nucl. Phys. B*, 915:1–9, 2017.

[254] Jian-Hui Zhang, Jiunn-Wei Chen, and Christopher Monahan. Parton distribution functions from reduced Ioffe-time distributions. *Phys. Rev. D*, 97(7):074508, 2018.

[255] Vladimir M. Braun, Alexey Vladimirov, and Jian-Hui Zhang. Power corrections and renormalons in parton quasidistributions. *Phys. Rev. D*, 99(1):014013, 2019.

[256] Wei Wang, Jian-Hui Zhang, Shuai Zhao, and Ruilin Zhu. Complete matching for quasidistribution functions in large momentum effective theory. *Phys. Rev. D*, 100(7):074509, 2019.

[257] Yao Ji, Jian-Hui Zhang, Shuai Zhao, and Ruilin Zhu. Renormalization and mixing of staple-shaped Wilson line operators on the lattice revisited. *Phys. Rev. D*, 104(9):094510, 2021.

[258] Yushan Su, Jack Holligan, Xiangdong Ji, Fei Yao, Jian-Hui Zhang, and Rui Zhang. Resumming quark's longitudinal momentum logarithms in LaMET expansion of lattice PDFs. *Nucl. Phys. B*, 991:116201, 2023.

[259] Ruilin Zhu, Yao Ji, Jian-Hui Zhang, and Shuai Zhao. Gluon transverse-momentum-dependent distributions from large-momentum effective theory. *JHEP*, 02:114, 2023.

[260] Fei Yao, Yao Ji, and Jian-Hui Zhang. Connecting Euclidean to light-cone correlations: from flavor nonsinglet in forward kinematics to flavor singlet in non-forward kinematics. *JHEP*, 11:021, 2023.

[261] Zhuoyi Pang, Fei Yao, and Jian-Hui Zhang. Total gluon helicity from lattice without effective theory matching. *JHEP*, 07:222, 2024.

[262] Chao Han, Wei Wang, Jia-Lu Zhang, and Jian-Hui Zhang. Power corrections to quasidistribution amplitudes of a heavy meson. *Phys. Rev. D*, 110(9):094038, 2024.

[263] Kostas Orginos, Anatoly Radyushkin, Joseph Karpie, and Savvas Zafeiropoulos. Lattice QCD exploration of parton pseudo-distribution functions. *Phys. Rev. D*, 96(9):094503, 2017.

[264] Zheng-Yang Li, Yan-Qing Ma, and Jian-Wei Qiu. Extraction of Next-to-Next-to-Leading-Order Parton Distribution Functions from Lattice QCD Calculations. *Phys. Rev. Lett.*, 126(7):072001, 2021.

[265] Rui Zhang, Jack Holligan, Xiangdong Ji, and Yushan Su. Leading power accuracy in lattice calculations of parton distributions. *Phys. Lett. B*, 844:138081, 2023.

[266] Chen Cheng, Li-Hong Huang, Xiang Li, Zheng-Yang Li, and Yan-Qing Ma. Lattice-QCD Computable Quark Correlation Functions at Three-Loop Order and Extraction of Splitting Functions. *Phys. Rev. Lett.*, 134(25):251902, 2025.

[267] Yoshitaka Hatta, Xiangdong Ji, and Yong Zhao. Gluon helicity ΔG from a universality class of operators on a lattice. *Phys. Rev. D*, 89(8):085030, 2014.

[268] Christian W. Bauer, Sean Fleming, and Michael E. Luke. Summing Sudakov logarithms in $B \to X_s\gamma$in effective field theory. *Phys. Rev. D*, 63:014006, 2000.

[269] Christian W. Bauer, Sean Fleming, Dan Pirjol, and Iain W. Stewart. An Effective field theory for collinear and soft gluons: Heavy to light decays. *Phys. Rev. D*, 63:114020, 2001.

[270] Christian W. Bauer and Iain W. Stewart. Invariant operators in collinear effective theory. *Phys. Lett. B*, 516:134–142, 2001.

[271] Christian W. Bauer, Dan Pirjol, and Iain W. Stewart. Soft collinear factorization in effective field theory. *Phys. Rev. D*, 65:054022, 2002.

[272] V. N. Gribov. Quantization of Nonabelian Gauge Theories. *Nucl. Phys. B*, 139:1, 1978.

[273] I. M. Singer. Some Remarks on the Gribov Ambiguity. *Commun. Math. Phys.*, 60:7–12, 1978.

[274] N. Vandersickel and Daniel Zwanziger. The Gribov problem and QCD dynamics. *Phys. Rept.*, 520:175–251, 2012.

[275] Leonardo Giusti, M. L. Paciello, C. Parrinello, S. Petrarca, and B. Taglienti. Problems on lattice gauge fixing. *Int. J. Mod. Phys. A*, 16:3487–3534, 2001.

[276] G. Martinelli, S. Petrarca, Christopher T. Sachrajda, and A. Vladikas. Nonperturbative renormalization of two quark operators with an improved lattice fermion action. *Phys. Lett. B*, 311:241–248, 1993. [Erratum: Phys.Lett.B 317, 660 (1993)].

[277] Maria Luigia Paciello, Silvano Petrarca, Bruno Taglienti, and Anastassios Vladikas. Gribov noise of the lattice axial current renormalization constant. *Phys. Lett. B*, 341:187–194, 1994.

[278] Rui Zhang, Anthony V. Grebe, Daniel C. Hackett, Michael L. Wagman, and Yong Zhao. Kinematically enhanced interpolating operators for boosted hadrons. *Phys. Rev. D*, 112(5):L051502, 2025.

[279] Xiangdong Ji, Yizhuang Liu, and Yushan Su. Asymptotic Long-Distance Expansion of Euclidean Correlators in Lattice Parton Applications. 1 2026.

[280] Jiunn-Wei Chen et al. Large-momentum effective theory's asymptotic extrapolation vs the inverse problem. *Phys. Rev. D*, 113(1):014509, 2026.

[281] Xiangdong Ji, Yizhuang Liu, and Yu-Sheng Liu. Transverse-momentum-dependent parton distribution functions from large-momentum effective theory. *Phys. Lett. B*, 811:135946, 2020.

[282] Markus A. Ebert, Iain W. Stewart, and Yong Zhao. Towards Quasi-Transverse Momentum Dependent PDFs Computable on the Lattice. *JHEP*, 09:037, 2019.

[283] Óscar del Río and Alexey Vladimirov. Quasitransverse momentum dependent distributions at next-to-next-to-leading order. *Phys. Rev. D*, 108(11):114009, 2023.

[284] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. P. Lepage, J. Shigemitsu, H. Trottier, and K. Wong. Highly improved staggered quarks on the lattice, with applications to charm physics. *Phys. Rev. D*, 75:054502, 2007.

[285] A. Bazavov et al. Lattice QCD Ensembles with Four Flavors of Highly Improved Staggered Quarks. *Phys. Rev. D*, 87(5):054505, 2013.

[286] Patricia Ball, V. M. Braun, and A. Lenz. Twist-4 distribution amplitudes of the K* and phi mesons in QCD. *JHEP*, 08:090, 2007.

[287] Craig D. Roberts, David G. Richards, Tanja Horn, and Lei Chang. Insights into the emergence of mass from studies of pion and kaon structure. *Prog. Part. Nucl. Phys.*, 120:103883, 2021.

[288] 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. [Addendum: JHEP 11, 037 (2020)].

[289] N. G. Stefanis. What binds quarks together at different momentum scales? A conceptual scenario. *Phys. Lett. B*, 738:483–487, 2014.

[290] S. Alekhin, J. Blümlein, S. Moch, and R. Placakyte. Parton distribution functions, α_s, and heavy-quark masses for LHC Run II. *Phys. Rev. D*, 96(1):014011, 2017.

[291] Juan Cruz-Martinez, Toon Hasenack, Felix Hekhorn, Giacomo Magni, Emanuele R. Nocera, Tanjona R. Rabemananjara, Juan Rojo, Tanishq Sharma, and Gijs van Seeventer. NNPDFpol2.0: a global determination of polarised PDFs and their uncertainties at next-to-next-to-leading order. *JHEP*, 07:168, 2025.

[292] A. Bazavov et al. Equation of state in (2+1)-flavor QCD. *Phys. Rev. D*, 90:094503, 2014.

[293] Anna Hasenfratz and Francesco Knechtli. Flavor symmetry and the static potential with hypercubic blocking. *Phys. Rev. D*, 64:034504, 2001.

[294] G. Martinelli, C. Pittori, Christopher T. Sachrajda, M. Testa, and A. Vladikas. A general method for non-perturbative renormalization of lattice operators. *Nucl. Phys. B*, 445:81–108, 1995.

[295] G. Burgio, M. Quandt, and H. Reinhardt. Coulomb gauge gluon propagator and the Gribov formula. *Phys. Rev. Lett.*, 102:032002, 2009.

[296] G. Burgio, M. Schrock, H. Reinhardt, and M. Quandt. Running mass, effective energy and confinement: the lattice quark propagator in Coulomb gauge. *Phys. Rev. D*, 86:014506, 2012.

[297] G. Curci, W. Furmanski, and R. Petronzio. Evolution of Parton Densities Beyond Leading Order: The Nonsinglet Case. *Nucl. Phys. B*, 175:27–92, 1980.

[298] Jack Holligan, Huey-Wen Lin, Rui Zhang, and Yong Zhao. Resummation for lattice QCD calculation of generalized parton distributions at nonzero skewness. *JHEP*, 207:241, 2025.

[299] 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. D*, 100(3):034516, 2019.

[300] Chris Bouchard, Chia Cheng Chang, Thorsten Kurth, Kostas Orginos, and Andre Walker-Loud. On the Feynman-Hellmann Theorem in Quantum Field Theory and the Calculation of Matrix Elements. *Phys. Rev. D*, 96(1):014504, 2017.

[301] Emanuele R. Nocera, Richard D. Ball, Stefano Forte, Giovanni Ridolfi, and Juan Rojo. A first unbiased global determination of polarized PDFs and their uncertainties. *Nucl. Phys. B*, 887:276–308, 2014.

[302] Ignacio Borsa, Marco Stratmann, Werner Vogelsang, Daniel de Florian, and Rodolfo Sassot. Next-to-Next-to-Leading Order Global Analysis of Polarized Parton Distribution Functions. *Phys. Rev. Lett.*, 133(15):151901, 2024.

[303] Valerio Bertone, Amedeo Chiefa, and Emanuele R. Nocera. Helicity-dependent parton distribution functions at next-to-next-to-leading order accuracy from inclusive and semi-inclusive deep-inelastic scattering data. *Phys. Lett. B*, 865:139497, 2025.

[304] C. Cocuzza, N. T. Hunt-Smith, W. Melnitchouk, N. Sato, and A. W. Thomas. Global QCD analysis of spin PDFs in the proton with high-x and lattice constraints. *Phys. Rev. D*, 112(11):114017, 2025.

[305] Leonard Gamberg, Michel Malda, Joshua A. Miller, Daniel Pitonyak, Alexei Prokudin, and Nobuo Sato. Updated QCD global analysis of single transverse-spin asymmetries: Extracting H˜, and the role of the Soffer bound and lattice QCD. *Phys. Rev. D*, 106(3):034014, 2022.

[306] Michael L. Wagman. Lanczos Algorithm, the Transfer Matrix, and the Signal-to-Noise Problem. *Phys. Rev. Lett.*, 134(24):241901, 2025.

[307] Daniel C. Hackett and Michael L. Wagman. Lanczos algorithm for lattice QCD matrix elements. *Phys. Rev. D*, 112(5):054506, 2025.

[308] Daniel C. Hackett and Michael L. Wagman. Block Lanczos algorithm for lattice QCD spectroscopy and matrix elements. *Phys. Rev. D*, 112(1):014514, 2025.

[309] Lorenzo Barca, Gunnar Bali, and Sara Collins. Nucleon sigma terms with a variational analysis from Lattice QCD. *Phys. Rev. D*, 111(3):L031505, 2025.

[310] Lorenzo Barca. Current-enhanced excited states in lattice QCD three-point functions. *Phys. Rev. D*, 112(9):L091503, 2025.

[311] Jozef Dudek et al. Physics Opportunities with the 12 GeV Upgrade at Jefferson Lab. *Eur. Phys. J. A*, 48:187, 2012.

[312] S. Amoroso et al. Snowmass 2021 Whitepaper: Proton Structure at the Precision Frontier. *Acta Phys. Polon. B*, 53(12):12–A1, 2022.

[313] C. T. H. Davies, B. R. Webber, and W. James Stirling. Drell-Yan Cross-Sections at Small Transverse Momentum. 1:I.95, 1984.

[314] G. A. Ladinsky and C. P. Yuan. The Nonperturbative regime in QCD resummation for gauge boson production at hadron colliders. *Phys. Rev. D*, 50:R4239, 1994.

[315] F. Landry, R. Brock, Pavel M. Nadolsky, and C. P. Yuan. Tevatron Run-1 Z boson data and Collins-Soper-Sterman resummation formalism. *Phys. Rev. D*, 67:073016, 2003.

[316] Anton V. Konychev and Pavel M. Nadolsky. Universality of the Collins-Soper-Sterman nonperturbative function in gauge boson production. *Phys. Lett. B*, 633:710–714, 2006.

[317] Peng Sun, Joshua Isaacson, C. P. Yuan, and Feng Yuan. Nonperturbative functions for SIDIS and Drell–Yan processes. *Int. J. Mod. Phys. A*, 33(11):1841006, 2018.

[318] Umberto D'Alesio, Miguel G. Echevarria, Stefano Melis, and Ignazio Scimemi. Nonperturbative QCD effects in q_T spectra of Drell-Yan and Z-boson production. *JHEP*, 11:098, 2014.

[319] Alessandro Bacchetta, Filippo Delcarro, Cristian Pisano, Marco Radici, and Andrea Signori. Extraction of partonic transverse momentum distributions from semi-inclusive deep-inelastic scattering, Drell-Yan and Z-boson production. *JHEP*, 06:081, 2017. [Erratum: JHEP 06, 051 (2019)].

[320] Ignazio Scimemi and Alexey Vladimirov. Analysis of vector boson production within TMD factorization. *Eur. Phys. J. C*, 78(2):89, 2018.

[321] Valerio Bertone, Ignazio Scimemi, and Alexey Vladimirov. Extraction of unpolarized quark transverse momentum dependent parton distributions from Drell-Yan/Z-boson production. *JHEP*, 06:028, 2019.

[322] Ignazio Scimemi and Alexey Vladimirov. Non-perturbative structure of semi-inclusive deep-inelastic and Drell-Yan scattering at small transverse momentum. *JHEP*, 06:137, 2020.

[323] Alessandro Bacchetta, Valerio Bertone, Chiara Bissolotti, Giuseppe Bozzi, Filippo Delcarro, Fulvio Piacenza, and Marco Radici. Transverse-momentum-dependent parton distributions up to N^{3}LL from Drell-Yan data. *JHEP*, 07:117, 2020.

[324] Francesco Hautmann, Ignazio Scimemi, and Alexey Vladimirov. Non-perturbative contributions to vector-boson transverse momentum spectra in hadronic collisions. *Phys. Lett. B*, 806:135478, 2020.

[325] Marcin Bury, Francesco Hautmann, Sergio Leal-Gomez, Ignazio Scimemi, Alexey Vladimirov, and Pia Zurita. PDF bias and flavor dependence in TMD distributions. *JHEP*, 10:118, 2022.

[326] Alessandro Bacchetta, Valerio Bertone, Chiara Bissolotti, Giuseppe Bozzi, Matteo Cerutti, Fulvio Piacenza, Marco Radici, and Andrea Signori. Unpolarized transverse momentum distributions from a global fit of Drell-Yan and semi-inclusive deep-inelastic scattering data. *JHEP*, 10:127, 2022.

[327] Joshua Isaacson, Yao Fu, and C. P. Yuan. Improving resbos for the precision needs of the LHC. *Phys. Rev. D*, 110(7):073002, 2024.

[328] F. Aslan, M. Boglione, J. O. Gonzalez-Hernandez, T. Rainaldi, T. C. Rogers, and A. Simonelli. Phenomenology of TMD parton distributions in Drell-Yan and Z0 boson production in a hadron structure oriented approach. *Phys. Rev. D*, 110(7):074016, 2024.

[329] Alessandro Bacchetta, Alessia Bongallino, Matteo Cerutti, Marco Radici, and Lorenzo Rossi. Exploring the Three-Dimensional Momentum Distribution of Longitudinally Polarized Quarks in the Proton. *Phys. Rev. Lett.*, 134(12):121901, 2025.

[330] Valentin Moos, Ignazio Scimemi, Alexey Vladimirov, and Pia Zurita. Extraction of unpolarized transverse momentum distributions from the fit of Drell-Yan data at N^{4}LL. *JHEP*, 05:036, 2024.

[331] Ke Yang, Tianbo Liu, Peng Sun, Yuxiang Zhao, and Bo-Qiang Ma. First Extraction of Transverse-Momentum Dependent Helicity Distributions. *Phys. Rev. Lett.*, 134(12):121902, 2025.

[332] Alessandro Bacchetta, Valerio Bertone, Chiara Bissolotti, Giuseppe Bozzi, Matteo Cerutti, Filippo Delcarro, Marco Radici, Lorenzo Rossi, and Andrea Signori. Flavor dependence of unpolarized quark transverse momentum distributions from a global fit. *JHEP*, 08:232, 2024.

[333] Valentin Moos, Ignazio Scimemi, Alexey Vladimirov, and Pia Zurita. Determination of unpolarized TMD distributions from the fit of Drell-Yan and SIDIS data at N^{4}LL. *JHEP*, 11:134, 2025.

[334] Alexey Vladimirov. Pion-induced Drell-Yan processes within TMD factorization. *JHEP*, 10:090, 2019.

[335] Matteo Cerutti, Lorenzo Rossi, Simone Venturini, Alessandro Bacchetta, Valerio Bertone, Chiara Bissolotti, and Marco Radici. Extraction of pion transverse momentum distributions from Drell-Yan data. *Phys. Rev. D*, 107(1):014014, 2023.

[336] P. C. Barry, L. Gamberg, W. Melnitchouk, E. Moffat, D. Pitonyak, A. Prokudin, and N. Sato. Tomography of pions and protons via transverse momentum dependent distributions. *Phys. Rev. D*, 108(9):L091504, 2023.

[337] Iain W. Stewart, Frank J. Tackmann, and Wouter J. Waalewijn. Factorization at the LHC: From PDFs to Initial State Jets. *Phys. Rev. D*, 81:094035, 2010.

[338] Xiangdong Ji, Yizhuang Liu, and Yu-Sheng Liu. TMD soft function from large-momentum effective theory. *Nucl. Phys. B*, 955:115054, 2020.

[339] Aneesh V. Manohar. Deep inelastic scattering as x —> 1 using soft collinear effective theory. *Phys. Rev. D*, 68:114019, 2003.

[340] Gunnar S. Bali, Bernhard Lang, Bernhard U. Musch, and Andreas Schäfer. Novel quark smearing for hadrons with high momenta in lattice QCD. *Phys. Rev. D*, 93(9):094515, 2016.

[341] Eigo Shintani, Rudy Arthur, Thomas Blum, Taku Izubuchi, Chulwoo Jung, and Christoph Lehner. Covariant approximation averaging. *Phys. Rev. D*, 91(11):114511, 2015.

[342] Ye Li and Hua Xing Zhu. Bootstrapping Rapidity Anomalous Dimensions for Transverse-Momentum Resummation. *Phys. Rev. Lett.*, 118(2):022004, 2017.

[343] Alexey A. Vladimirov. Correspondence between Soft and Rapidity Anomalous Dimensions. *Phys. Rev. Lett.*, 118(6):062001, 2017.

[344] Zhong-Bo Kang, Jani Penttala, and Congyue Zhang. Determination of the strong coupling constant and the Collins-Soper kernel from the energy-energy correlator in e^+e^- collisions. 10 2024.

[345] A. Bermudez Martinez, F. Hautmann, L. Keersmaekers, A. Lelek, M. Mendizabal Morentin, S. Taheri Monfared, and A. M. van Kampen. Soft-gluon coupling and the TMD parton branching Sudakov form factor. *Phys. Lett. B*, 868:139762, 2025.

[346] Alessandro Bacchetta, Valerio Bertone, Chiara Bissolotti, Matteo Cerutti, Marco Radici, Simone Rodini, and Lorenzo Rossi. Neural-Network Extraction of Unpolarized Transverse-Momentum-Dependent Distributions. *Phys. Rev. Lett.*, 135(2):021904, 2025.

[347] P. C. Barry, Chueng-Ryong Ji, N. Sato, and W. Melnitchouk. Global QCD Analysis of Pion Parton Distributions with Threshold Resummation. *Phys. Rev. Lett.*, 127(23):232001, 2021.

[348] John C. Collins and Davison E. Soper. Back-To-Back Jets in QCD. *Nucl. Phys. B*, 193:381, 1981. [Erratum: Nucl.Phys.B 213, 545 (1983)].

[349] Ian Moult, Hua Xing Zhu, and Yu Jiao Zhu. The four loop QCD rapidity anomalous dimension. *JHEP*, 08:280, 2022.

[350] Claude Duhr, Bernhard Mistlberger, and Gherardo Vita. Four-Loop Rapidity Anomalous Dimension and Event Shapes to Fourth Logarithmic Order. *Phys. Rev. Lett.*, 129(16):162001, 2022.

[351] Tomomi Ishikawa, Yan-Qing Ma, Jian-Wei Qiu, and Shinsuke Yoshida. Renormalizability of quasiparton distribution functions. *Phys. Rev. D*, 96(9):094019, 2017.

[352] Jeremy Green, Karl Jansen, and Fernanda Steffens. Nonperturbative Renormalization of Nonlocal Quark Bilinears for Parton Quasidistribution Functions on the Lattice Using an Auxiliary Field. *Phys. Rev. Lett.*, 121(2):022004, 2018.

[353] Xiang-Dong Ji and M. J. Musolf. Subleading logarithmic mass dependence in heavy meson form-factors. *Phys. Lett. B*, 257:409–413, 1991.

[354] Yi-Kai Huo et al. Self-renormalization of quasi-light-front correlators on the lattice. *Nucl. Phys. B*, 969:115443, 2021.

[355] G. W. Kilcup, Stephen R. Sharpe, Rajan Gupta, G. Guralnik, A. Patel, and T. Warnock. ϵ Beyond the Naive Mass Spectrum. *Phys. Lett. B*, 164:347–355, 1985.

[356] Min-Huan Chu et al. Nonperturbative determination of the Collins-Soper kernel from quasitransverse-momentum-dependent wave functions. *Phys. Rev. D*, 106(3):034509, 2022.

[357] Wei-Yang Liu and Jiunn-Wei Chen. Chiral perturbation for large momentum effective field theory. *Phys. Rev. D*, 104(5):054508, 2021.

[358] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *nature*, 596(7873):583–589, 2021.

[359] Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, Lindsay Willmore, Andrew J Ballard, Joshua Bambrick, et al. Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*, 630(8016):493–500, 2024.

[360] Amil Merchant, Simon Batzner, Samuel S Schoenholz, Muratahan Aykol, Gowoon Cheon, and Ekin Dogus Cubuk. Scaling deep learning for materials discovery. *Nature*, 624(7990):80–85, 2023.

[361] Liam Parker, Francois Lanusse, Jeff Shen, Ollie Liu, Tom Hehir, Leopoldo Sarra, Lucas Meyer, Micah Bowles, Sebastian Wagner-Carena, Helen Qu, et al. Aion-1: Omnimodal foundation model for astronomical sciences. *arXiv preprint arXiv:2510.17960*, 2025.

[362] Martin Luscher. Trivializing maps, the Wilson flow and the HMC algorithm. *Commun. Math. Phys.*, 293:899–919, 2010.

[363] Jinchen He, Xiao-Yong Jin, James C. Osborn, and Yong Zhao. Neural Field Transformations for Hybrid Monte Carlo: Architectural Design and Scaling. In *39th Annual Conference on Neural Information Processing Systems: Includes Machine Learning and the Physical Sciences (ML4PS)*, 11 2025.

[364] Xiao-Yong Jin. Neural Network Field Transformation and Its Application in HMC. *PoS*, LATTICE2021:600, 2022.

[365] Xiao-Yong Jin. Neural Network Gauge Field Transformation for 4D SU(3) gauge fields. *PoS*, LATTICE2023:040, 2023.

[366] IP Omelyan, IM Mryglod, and R Folk. Symplectic analytically integrable decomposition algorithms: classification, derivation, and application to molecular dynamics, quantum and celestial mechanics simulations. *Computer Physics Communications*, 151(3):272–314, 2003.

[367] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. 2017.

[368] Rupert G Miller. The jackknife-a review. *Biometrika*, 61(1):1–15, 1974.

[369] Shuzhe Shi, Lingxiao Wang, and Kai Zhou. Rethinking the ill-posedness of the spectral function reconstruction — Why is it fundamentally hard and how Artificial Neural Networks can help. *Comput. Phys. Commun.*, 282:108547, 2023.

[370] Verena Bellscheidt, Nora Brambilla, Andreas S. Kronfeld, and Julian Mayer-Steudte. Wilson loops with neural networks. 2 2026.

[371] Tingjia Miao, Jiawen Dai, Jingkun Liu, Jinxin Tan, Muhua Zhang, Wenkai Jin, Yuwen Du, Tian Jin, Xianghe Pang, Zexi Liu, et al. Physmaster: Building an autonomous ai physicist for theoretical and computational physics research. *arXiv preprint arXiv:2512.19799*, 2025.