EXPLORING HADRON STRUCTURE THROUGH MONTE-CARLO FITS AND MODEL CALCULATIONS

A Dissertation
Submitted to
the Temple University Graduate Board

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

---

by
Christopher Cocuzza
May 2023

Examining Committee Members:

Andreas Metz, Department of Physics, Advisory Chair
Martha Constantinou, Department of Physics
Bernd Surrow, Department of Physics
Leonard Gamberg, External Reader, Penn State Berks

ABSTRACT

Since the discovery in the 1960's that the proton is not a fundamental particle but instead composed of even smaller particles known as quarks and gluons, there has been a concerted effort to understand the proton's internal structure. There still remain many mysteries about the proton and the theory that describes the interactions within: Quantum Chromodynamics (QCD). The distributions of quarks and gluons are encoded in objects known as parton correlation functions. Physicists use high-energy scattering experiments to access these functions by means of QCD factorization. This process of extracting information is known as a global QCD analysis. Further insight can be gained through first-principles calculations in lattice QCD as well as models for the strong interaction.

In this thesis, we will use global QCD analyses to provide information on the one-dimensional (1D) structure of the proton using the latest experimental data available. Among the mysteries that remain within the proton, we provide insight on the non-perturbative nature of the proton's sea quarks, for both cases where the proton is unpolarized and longitudinally polarized. We also bring new information on the "proton spin puzzle," which concerns the delegation of the proton's spin into its constituent quarks and gluons. We shed light on the proton's transversely polarized structure, where current results from global QCD analyses and lattice QCD fail to paint a consistent picture. Our analyses also reveal a new feature of nuclear effects within light, highly asymmetric nuclei such as helium and tritium. Finally, we perform derivations in a spectator diquark model to glean information on the proton's 3D structure, and calculate moments that can be used in future lattice QCD studies.

ACKNOWLEDGEMENTS

First, I want to share my gratitude to my advisor Dr. Andreas Metz. I am lucky to have had such an excellent advisor, who gave me the freedom and trust to work, learn, and grow independently as a researcher. Our meetings would always provide valuable insight, as he was never afraid to raise difficult (and occasionally frustrating) questions and problems while still respecting my ability and perspective. His wisdom and advice ensured that I was always prepared for the challenges that I faced as a graduate student, and collaborating with him made the journey a smooth and enjoyable one. Thank you to my committee members Dr. Martha Constantinou, Dr. Bernd Surrow, and Dr. Leonard Gamberg for their time and thoughtful questions.

I would like to thank my collaborators Dr. Nobuo Sato, Dr. Wally Melnitchouk, Dr. Daniel Pitonyak, and Dr. Alexey Prokudin. Their time, insights, and knowledge were invaluable to my growth, and I am appreciative of their patience as I bombarded them with my ephemeral results and thoughts every week. Thank you to Nobuo in particular for the abundance of help while I learned how to program. My sincere gratitude also to Dr. Yiyu Zhou and Dr. Patrick Barry from the JAM collaboration. They were there to answer all of my silly questions when I first joined and was first learning Python, which made the most frustrating part of my Ph. D. career a lot easier. I want to thank my one and only group member during my time at Temple University: Dr. Shohini Bhattacharya. Beyond being an incredible collaborator and a supportive mentor, they are an even better friend whose levity and emotional support

I could not have done without during my first few years at Temple.

Thank you to my friends and classmates Dr. Nick Lukow, Kanishka Wijesekara, Santosh Neupane, Navagyan Ghimire, Babu Pokhrel, Biru Kc, and Anoj Aryal. I will always remember the fun times we had together playing volleyball and going to the gym. Thank you for the hospitality whenever I would invite myself over to your homes and for sharing your delicious vegan food. I would also like to appreciate my friends Drew Calhoun and Akida Soto for all of the deep and stimulating conversations, delicious dinners, and fun activities that we've shared. Thank you to my cats Buttons, Finnegan, Deuce, Pip, Penny, Raccoon, Elsie, Darling, Claudia, Charlie, George 2.0, John, Grayson, Nosy Nora, Pipsqueak, Derek, Squid, and Scallop; and my chicken Paulette. May Oreo, Poopydoo, Stuart, Stewie, Clover, Gizmo, Maggie, Mimi, Boomer, Harvey, George, Ms. Hare, Velveeta, Ariel, Cinderella, Sprinkles, Curlycue, Leah, Mr. Bates, and Mark rest in peace.

Thank you to my Mommy and Dada for their never-ending love and support, especially while I was at home during COVID, and for driving me back and forth to Philadelphia, cooking me food, and playing Jeopardy! every night. Finally, I want to thank my partner Chetna Johri, who was there for all of the good and bad of this adventure. She was there as I showed off my pretty plots and regaled her with my favorite part of the dissertation, offering insightful questions, genuine understanding, and even predicting what I would say next. Her ability and willingness to understand my convoluted research is invaluable to me. She was also there to help me with the daunting task of job searching and preparing for interviews, and to leave surprise desserts and notes after my return from stressful conferences. On the day of my defense, knowing that I would be surprised with desserts afterwards helped me power through. Our partnership made this journey an enjoyable one.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Perturbative QCD and Factorization

Over the past five to six decades, there has been tremendous interest in the structure of the proton, discovered in 1911 by E. Rutherford [1], and the neutron, discovered in 1932 by J. Chadwick [2, 3], collectively known as "nucleons." Starting from the 1960's, it was proposed that the nucleons are composed of even smaller particles called quarks, antiquarks, and gluons [4–6], collectively called "partons." According to our present understanding, the partons belong to the fundamental building blocks of all matter.

The theory describing such sub-atomic particles and their interactions is known as Quantum Field Theory (QFT), which posits that the interactions are due to fields composed of particles known as "gauge bosons." In the case of Quantum Electrodynamics (QED), which governs the electromagnetic interaction and is the simplest such QFT, the boson is the photon, corresponding to the electromagnetic field. Any particle that carries electric charge can interact through the electromagnetic force, such as quarks, antiquarks and charged leptons (electrons, muons, tauons and their corresponding anti-particles). The gauge bosons of the weak interaction, on the other hand, are the $W^{\pm}$ and Z bosons. All particles that carry a weak charge are subject to the weak interaction. This includes the quarks, charged leptons and the electrically neutral leptons, that is, the neutrinos, along with all the respective anti-particles. The electromagnetic and weak interactions are part of the (unified) electroweak theory. In the case of Quantum Chromodynamics (QCD), which governs the strong interaction, the gauge boson is the gluon, and any particle with color charge can interact through

the strong interaction. Quarks, as well as gluons themselves, carry color charge and thus the strong force is the force through which partons interact and is that which binds nucleons (or more generally, hadrons) together.

All of the aforementioned particles (quarks, leptons, photons, $W^{\pm}/Z$ bosons, gluons), in addition to the Higgs boson [7–10], comprise what is known as the "Standard Model" of particle physics [11], which deals with all known fundamental particles and describes their electroweak and strong interactions in terms of a gauge theory. The Standard Model has proven extremely capable at describing most observed phenomena in particle physics, although there have been proposals for "physics beyond the Standard Model" [12]. The work of this thesis is mostly related to the strong interaction, and will thus remain within the bounds of the Standard Model.

The dynamics of the strong interaction is governed by the Lagrangian for QCD,

$$\mathcal{L}_{\mathrm{QCD}} = \sum_{q=1}^{N_q} \bar{\psi}_q \big(i \not{D} - m_q\big) \psi_q - \frac{1}{4} F_a^{\mu\nu} F^a_{\mu\nu}, \tag{1.1}$$

where ψ_q is the quark field for quark flavor q. Additionally, $D^\mu = \partial^\mu - igA_a^\mu T_a$ is the covariant derivative where T_a are the 3×3 Gell-Mann matrices, and $F_a^{\mu\nu} = \partial^\mu A_a^\nu - \partial^\nu A_a^\mu + g f_{abc} A_b^\mu A_c^\nu$, where A_a^μ is the gluon field and f_{abc} is the structure constant of the SU(3) color group. The strength of the strong interaction is determined by g, and we define the QCD coupling constant $\alpha_s \equiv g^2/4\pi$, which is given approximately by

$$\alpha_s(\mu_R) \approx \frac{12\pi}{(11C_A - 4N_f T_R) \ln\big(\mu_R^2/\Lambda_{\mathrm{QCD}}^2\big)}. \tag{1.2}$$

Here $C_A = 3$ and $T_R = \frac{1}{2}$ are SU(3) color factors [11], N_f is the number of active quark flavors, μ_R is the renormalization energy scale, and Λ_{QCD} is a QCD constant of integration corresponding to the scale where the perturbatively-defined coupling would diverge. Eq. (1.2), which is the leading-order (LO) term in an expansion of $1/\ln\big(\mu_R^2/\Lambda_{\mathrm{QCD}}^2\big)$, implies that the larger the scale μ_R the smaller α_s. In the limit $\mu_R \to \infty$, corresponding to vanishing distances between partons, the coupling constant goes to zero, which is called asymptotic freedom [13–15]. At the (experimentally accessible) scale of the mass of the Z boson ($\mu_R \approx 91$ GeV), one has $\alpha_s(\mu_R) \approx 0.118$ [11]. The large coupling constant at low values of μ_R of around 1 GeV has some relation to the so-called confinement property of QCD, which expresses the fact that individual

partons carrying a color charge cannot be isolated for distance scales of more than about $1\,\mathrm{fm} = 10^{-15}\,\mathrm{m}$. We emphasize that, in the first place, this color confinement is an empirical feature of the strong interaction which so far lacks a generally accepted theoretical explanation.

The QED counterpart of α_s is the fine structure constant $\alpha_{\mathrm{em}} = e^2/(4\pi)$, where e is the elementary charge. The QED coupling also depends on the renormalization scale μ_R. In contrast to the QCD case though, α_{em} increases as μ_R increases. But, since $\alpha_{\mathrm{em}} \approx 1/137$ at lower energies, and since its increase with increasing μ_R is (also) logarithmic and therefore mild, QED perturbation theory with the fine structure constant as the expansion parameter typically works very well for presently accessible energies.

On the other hand, the above discussion implies that a perturbative expansion in QCD is only valid at large energy scales. In fact, several high-energy scattering processes can be computed completely using perturbative QCD. One example is the total cross section for electron-positron annihilation. This is no longer true for processes with hadrons in the initial state and/or identified hadrons in the final state, regardless of how high the center-of-mass (COM) energy of the process is. In such cases one can often use what are known as QCD factorization theorems which allow access to the parton structure of hadrons. In order to intuitively understand the essence of factorization for a high-energy process, one can always consider the hadron in a frame where it travels at nearly the speed of light [16, 17]. Further, since high-energy scales correspond to short time scales, only the minimal number (typically one) of partons inside a hadron directly participate in the hard scattering. Put differently, the parton involved in a hard scattering is essentially a free particle since the interaction between this parton and the other partons inside the same hadron occurs on a much longer time scale. One can then factorize the cross section into a "hard part" and a "soft part" [18]. The hard part describes the interaction between the probe and the parton, which is calculable in perturbative QCD and dependent on the experimental process. The soft part describes the internal structure of the hadron that is being probed, and is governed by the non-perturbative dynamics of QCD [19]. An important ingredient in this context is "universality" which states that the soft part does not depend on the experimental process, and thus it is the object that provides general information about hadron substructure. QCD factorization, together with universality, also provides a framework with predictive power in the sense that the non-perturbative

contribution can be extracted from experimental data for one hard process and then used to make predictions for cross sections of other processes.

So it becomes the goal to determine the soft part, which is given by objects known as parton correlation functions. The simplest of these functions are the Parton Distribution Functions (PDFs), which describe the one-dimensional (1D) longitudinal momentum distribution of partons. They are dependent upon x, the parton momentum fraction, which is given by the ratio of the parton's longitudinal momentum to that of the hadron's longitudinal momentum. In addition, they, and all of the following distributions, depend upon the renormalization scale μ_R, which is largely determined by the hard scale of the experimental process (it is often chosen to be equal to the hard scale to reduce errors in the perturbative expansion in α_s, and it can be varied to estimate the size of the higher order terms). Generalizing to 3D, one has Transverse Momentum-Dependent Parton Distributions (TMDs) [20–23], which depend additionally on the transverse momentum of the parton, and Generalized Parton Distributions (GPDs) [24–28], which depend additionally on the longitudinal and transverse momentum transfer to the nucleon. The most general 5D functions depending on all of the aforementioned variables are known as Generalized Transverse Momentum-Dependent Parton Distributions (GTMDs) [29]. This thesis involves work on PDFs and GPDs. Furthermore, in semi-inclusive experimental processes, where one or more outgoing hadrons are detected, Fragmentation Functions (FFs) appear as the soft part. These describe the fragmentation of a parton into a hadron or hadrons, where in this thesis we consider single-hadron FFs as well as di-hadron FFs (DiFFs).

In general, all of these distributions depend on the parton, which may be an up, down, strange, charm, bottom, or top quark (and all of their corresponding anti-quarks), or a gluon. There are also different distributions depending on the hadron's spin relative to its momentum. If the spin has no particular direction and is averaged over, we speak of "spin-averaged" or "unpolarized" distributions. If the spin is instead parallel or transverse to its momentum, we speak of "helicity" or "transversity" distributions, respectively, often collectively called "polarized" distributions.

Currently, there are three methods to gain information on the parton correlation functions. The first, as mentioned above, is through high-energy scattering experiments. With the experimental cross section matched with the corresponding factorization theorem, the functions can be extracted through comparison of the two. This process of extracting the functions through a comparison of theory and experiment is

known as a "QCD analysis" and, if all relevant data are made use of, a "global QCD analysis." The second method is through model calculations, where one proposes a simplified model of the strong interaction that allows one to calculate the functions without experimental input. The third method is known as lattice QCD, which is the only computational method that allows one to calculate the functions from first principles. This thesis will primarily focus on using global QCD analyses to extract the functions with four chapters dedicated to the subject, along with another chapter on a model calculation.

In this thesis we aim to address some of the open questions that remain about hadron structure. The global QCD analyses to follow primarily concern the 1D PDFs, including the spin-averaged, helicity, and transversity distributions. Even with these simplest of functions, many questions remain. The first such question that we address is the nonperturbative origin of both the spin-averaged and helicity "sea asymmetry," which quantifies the difference between anti-up and anti-down quarks in the proton. We also comment on the "proton spin puzzle," a long-standing problem in describing how the proton's spin is delegated among its valence quarks, sea quarks, and gluons. The proton has spin 1/2, which, according to the Jaffe and Manohar spin sum rule [30], can be attributed to the partons through

$$\frac{1}{2} = \frac{1}{2}\Delta\Sigma(\mu_R) + \Delta G(\mu_R) + L_{Q+G}(\mu_R), \tag{1.3}$$

where $\Delta\Sigma$ is the quark and antiquark helicity, ΔG is the gluon helicity, and L_{Q+G} is the parton orbital angular momentum. Through a global analysis of helicity PDFs we are able to gain information on the first two terms. Finally, one has the transversity PDFs, which are the most difficult to access of the 1D PDFs and are still poorly known. Currently, results from global analyses of DiFFs [31, 32] and TMDs [33, 34], as well as those from lattice QCD [35–37], do not provide a consistent picture of the proton's transverse spin structure. Further work has also been conducted to estimate the impact of the future Electron Ion Collider (EIC), specifically its impact on the strange quark's helicity [38], the spin-averaged strange distribution, and the weak mixing angle [39]. These impact studies will not be covered in this thesis.

1.2 Global QCD Analysis

Here we cover the basics of extracting parton correlation functions from experimental data through the process of a global QCD analysis, with further details provided in Chapter 2. The first step of each global QCD analysis is to parameterize the functions at a chosen input energy scale. The choices for the parameterizations and the input scale are heavily dependent on the analysis, and will be discussed individually for each one. Constraints on these parameterizations can be found in sum rules and isospin relations. In the following discussion, we focus on the proton, which has 2 valence up quarks (u_v), 1 valence down quark (d_v), and no strange valence (s_v). Further, the total momentum of all partons, including the gluon (g) must be equal to that of the proton. Valence quarks are defined as $q_v = q - \bar{q}$, with $\bar{q}$ the antiquark of q and $q = u, d$ for the proton. Sea quarks are defined as all non-valence quarks which, for the proton and including only the lightest three quarks, are $q = s, \bar{u}, \bar{d}, \bar{s}$. These facts lead to the following valence and momentum sum rules on the spin-averaged PDFs:

$$\int_0^1 \mathrm{d}x\ u_v(x, \mu_R) = 2, \tag{1.4a}$$

$$\int_0^1 \mathrm{d}x\ d_v(x, \mu_R) = 1, \tag{1.4b}$$

$$\int_0^1 \mathrm{d}x\ s_v(x, \mu_R) = 0, \tag{1.4c}$$

$$\int_0^1 \mathrm{d}x\ x\Big(u^+(x, \mu_R) + d^+(x, \mu_R) + s^+(x, \mu_R) + g(x, \mu_R)\Big) = 1, \tag{1.4d}$$

where $q^+ \equiv q + \bar{q}$. Corresponding sum rules apply to other hadrons such as the pion and kaon. While many experimental measurements are performed directly on a proton target, there are also a number that are performed on nuclear targets which also contain neutrons. Fortunately, thanks to isospin symmetry, one does not need to consider entirely new functions for the neutron. Instead, the functions for the neutron and proton can be related through

$$u^n(x, \mu_R) = d^p(x, \mu_R), \quad d^n(x, \mu_R) = u^p(x, \mu_R), \tag{1.5}$$

where p (n) denotes the function within a proton (neutron), and all other PDFs are the same between the proton and neutron. The only approximation entering these

relations is the neglect of the mass difference between up and down quarks.

With the functions parameterized, one then uses the factorization theorem to relate the functions to the experimental observable. However, the experimental measurements often take place at energy scales that are different from the input scale, and so the functions must first be evolved to that scale. For all of the functions involved in the global analyses of this thesis, the equation governing this evolution is known as the DGLAP evolution equation [40–42]. It is given by

$$\frac{\partial f_i(x;\mu_R)}{\partial \ln \mu_R^2} = \frac{\alpha_s}{2\pi} \sum_j \left[P_{ij} \otimes f_i\right](x;\mu_R), \tag{1.6}$$

where f_i is the parton correlation function of flavor i and P_{ij} are the splitting functions, which can be computed in perturbative QCD order by order in α_s. The symbol $\otimes$ represents the convolution

$$[C \otimes f](x) = \int_x^1 \frac{\mathrm{d}w}{w} C(w) f(\frac{x}{w}). \tag{1.7}$$

We note that for PDFs the so-called space-like splitting functions are to be used, while for FFs the time-like splitting functions are to be used. At LO in α_s, they are equal upon switching $i \leftrightarrow j$. The LO space-like and time-like functions were computed in Refs. [40–43] and Refs. [44–46] respectively, and calculations to Next-to-Leading Order (NLO) [47–52] and even Next-to-Next-to-Leading Order (NNLO) [53–55] are available.

With the functions evolved to the experimental energy scale, the factorization theorem allows a direct comparison between theory and data. Through this comparison, one can determine the parameters that enter the parameterization of the functions through the minimization of some metric, chosen in the following analyses to be the χ^2. With the parameters determined, one has thus "extracted" the functions through the use of experimental data and factorization theorems.

1.3 Dissertation Outline

The following research can be divided into two categories: global QCD analyses, where the results are presented in Chapters 2 – 5, and model calculations, where the results are presented in Chapter 6.

In Chapter 2 we present a simultaneous global QCD analysis of spin-averaged PDFs and FFs [56]. In particular, we focus on the question of the sea asymmetry, defined as the difference between anti-down and anti-up quarks: $\bar{d} - \bar{u}$. Previous experiments have demonstrated that this asymmetry is nonzero, despite the fact that a significant asymmetry cannot be generated perturbatively from gluons splitting into quark-antiquark pairs due to the (approximately) equal masses of up and down quarks. As such, nonperturbative approaches are required to explain this phenomenon. In this chapter we will present the results of the PDFs from an analysis including the latest data relevant for extracting the sea asymmetry. Our results will be compared to other global QCD analyses and to nonperturbative models.

In Chapter 3 we present a simultaneous global QCD analysis of nuclear effects in deuterium, helium, and tritium as well as spin-averaged PDFs [57]. We include the latest data from the MARATHON collaboration, which measured the Deep Inelastic Scattering (DIS) process on helium and tritium targets. We use this new data and the mirror symmetry of helium and tritium to extract "isovector" nuclear effects, which are nuclear effects that differ between the two $A = 3$ nuclei. We present the first extraction of such effects, as well as results for the d/u quark ratio and nuclear effects in deuterium.

In Chapter 4 we present a simultaneous global QCD analysis of helicity PDFs, spin-averaged PDFs, and FFs [58]. Analogous to Chapter 2, we focus again on the sea asymmetry except this time for the helicity PDFs, defined as $\Delta\bar{u} - \Delta\bar{d}$, and include the latest data relevant for its extraction. We compare our result to other global QCD analyses and model predictions, and calculate the contribution to the proton's spin from the light quarks and antiquarks. We also present results for the gluon's helicity, relevant for the proton spin sum rule in Eq. (1.3). For all of the results we also look at the impact of imposing positivity constraints on the helicity PDFs. Finally, we present the results for the pion, kaon, and unidentified hadron FFs and compare to other analyses.

In Chapter 5 we present a simultaneous global QCD analysis of DiFFs and transversity PDFs [59,60]. We start by proposing a new definition of unintegrated DiFFs that is compatible with the probability interpretation of collinear DiFFs, and we discuss the corresponding evolution equations. We then present results for the DiFFs using the latest data to constrain them, as well as data generated from the Monte Carlo

event generator PYTHIA. With the DiFFs determined from data, we show the results for the extracted transversity PDFs as well as the up, down, and isovector tensor charges. These results are compared to previous extractions of transversity PDFs using the di-hadron channel, as well as to extractions using single-hadron production in a TMD framework. Finally, we also compare our tensor charges to those found from lattice QCD.

In Chapter 6, we present results for quasi-distributions of GPDs (quasi-GPDs) within a model approximation known as the scalar diquark spectator model [61, 62]. We perform a perturbative calculation of quasi-GPDs and observe consistent features of the model. From the quasi-GPDs, we are also able to easily derive the quasi-PDFs. All of the quasi-distributions are compared to their corresponding light-cone distribution, to which they should reduce to in the limit of large forward momentum. We also look into the axial-vector diquark spectator model and compare its features to the scalar case. Finally, we discuss some model-independent results of quasi-PDFs and quasi-GPDs which can be used to shed light on systematic uncertainties of lattice-QCD calculations.

In Chapter 7 we provide a general summary of the thesis and an outlook on the topics discussed in this thesis.

CHAPTER 2

SPIN-AVERAGED PARTON DISTRIBUTION FUNCTIONS

In this chapter we present the details of a global QCD analysis of spin-averaged PDFs and FFs. We focus in particular on the sea asymmetry $\bar{d} - \bar{u}$, where high-energy scattering experiments and global QCD analyses of the data have conclusively demonstrated a sea asymmetry in the proton [63]. The result cannot be explained perturbatively through the splitting of gluons into quark-antiquark pairs [64], and requires nonperturbative mechanisms, such as dynamical chiral symmetry breaking and the pion cloud of the nucleon [65–75], or dynamics related to the Fermi-Dirac statistics of quarks and the Pauli exclusion principle [76–81]. More recently, exploratory studies have been made in extracting information on the isovector sea quark distributions directly from lattice QCD calculations [37, 82–85].

The first experimental indications of a light-quark sea asymmetry came from the CFS group at Fermilab in 1981 [86]. Measuring the Drell-Yan process, where a quark and antiquark from colliding hadrons annihilate into a virtual photon that subsequently decays into a lepton-antilepton pair, they found that the $\bar{d}$ distribution, integrated over parton momentum fraction x, was larger than the integrated $\bar{u}$ distribution. The first high-precision experimental evidence for a sea asymmetry came a decade later from the New Muon Collaboration (NMC) at CERN [87, 88], which used measurements of inclusive deep-inelastic scattering (DIS) on hydrogen and deuterium to test the Gottfried sum rule [89] and determine that the integral $\int_0^1 \mathrm{d}x \big[\bar{d}(x) - \bar{u}(x)\big]$ must be positive. Further evidence was provided by the NA51 Collaboration at CERN [90] using the Drell-Yan process, and indications for a nonzero

asymmetry were also found by the HERMES Collaboration [91] in semi-inclusive DIS of charged pions.

The most conclusive evidence for an excess of $\bar{d}$ over $\bar{u}$ was provided in 1998 by the Fermilab E866 (NuSea) experiment [92–94], which measured Drell-Yan lepton-pair production cross sections in proton-proton (pp) and proton-deuteron (pD) scattering. While the NuSea experiment probed the asymmetry up to momentum fractions $x \approx 0.3$, the subsequent Fermilab E906 (SeaQuest) experiment [95] extended the range up to $x \approx 0.4$, finding some tensions with the NuSea result in the high-x region. The NuSea data had suggested a significant fall in the $\bar{d}/\bar{u}$ ratio for $x \gtrsim 0.3$, albeit with large uncertainties, which is difficult to accommodate in many of the nonperturbative models [65–80]. The SeaQuest experiment was partially motivated to verify this behavior at large x.

An alternative method for extracting the $\bar{d} - \bar{u}$ asymmetry involves W-lepton production in hadronic collisions, whereby a quark and antiquark annihilate into a W boson that decays into a detected lepton and a neutrino. This has been measured in $p\bar{p}$ scattering at the Tevatron and in pp collisions at the Large Hadon Collider (LHC) and the Relativistic Heavy Ion Collider (RHIC). Recently, high-precision data from the STAR Collaboration at RHIC on the ratios of W-lepton production cross sections have been made available [96]. Since RHIC has a lower COM energy of $\sqrt{s} = 0.51$ TeV compared to the Tevatron ($\sqrt{s} = 1.96$ TeV) and the LHC ($\sqrt{s} \geq 7$ TeV), these data are sensitive to parton distributions at higher values of x, which can potentially provide information on the distributions of light sea quarks in this difficult to measure region.

Ref. [56] was the first global QCD analysis to include the SeaQuest and STAR W-lepton data, and these new results will be discussed in this chapter. In addition, results will be shown for the many other datasets in this analysis, which include data on inclusive and semi-inclusive DIS, Drell-Yan, W and Z boson production, and jet production. With this data combined, the resulting sea asymmetry, as well as the rest of the PDFs, will be shown and discussed.

2.1 Bayesian Global QCD Analysis

In this section, we cover the fundamentals of global QCD analyses, relevant for Chapters 2 – 5, starting with the Bayesian aspects of the analysis. In general, a Bayesian

analysis uses Bayes' theorem to determine the probability of a hypothesis given certain information. In the specific case of a global QCD analysis, we start by parametrizing the non-perturbative objects which, in the chapters to follow, includes PDFs, single hadron FFs, and DiFFs. The collection of parameters used to parametrize all of the objects is denoted by $\boldsymbol{a}$ and acts as the hypothesis of our Bayesian analysis, while the experimental measurements act as the evidence. We then sample the posterior distribution given by

$$\mathcal{P}(\boldsymbol{a}|\text{data}) \propto \mathcal{L}(\boldsymbol{a}, \text{data})\, \pi(\boldsymbol{a}), \tag{2.1}$$

with a likelihood function of Gaussian form,

$$\mathcal{L}(\boldsymbol{a}, \text{data}) = \exp\Big(-\frac{1}{2}\chi^2(\boldsymbol{a}, \text{data})\Big), \tag{2.2}$$

and a flat prior function $\pi(\boldsymbol{a})$ that vanishes in regions where the parameters $\boldsymbol{a}$ give unphysical PDFs, FFs, or DiFFs. The χ^2 function in (2.2) is defined as [97]

$$\chi^2(\boldsymbol{a}) = \sum_{i,e} \left(\frac{d_{i,e} - \sum_k r_e^k \beta_{i,e}^k - T_{i,e}(\boldsymbol{a})/N_e}{\alpha_{i,e}} \right)^2 + \sum_k \left(r_e^k\right)^2 + \sum_e \left(\frac{1 - N_e}{\delta N_e} \right)^2, \tag{2.3}$$

where $d_{i,e}$ is the experimental data point i from dataset e, and $T_{i,e}$ is the corresponding theoretical value. All uncorrelated uncertainties are added in quadrature and labeled by $\alpha_{i,e}$, while $\beta_{i,e}^k$ represents the k-th source of point-to-point correlated systematic uncertainties for the i-th data point weighted by r_e^k. The latter are optimized per values of the parameters $\boldsymbol{a}$ via $\partial\chi^2/\partial r_e^k = 0$, which introduces the second penalty term in Eq. (2.3). We include normalization parameters N_e for each dataset e as part of the posterior distribution per data set, with a Gaussian penalty controlled by the experimentally quoted normalization uncertainties δN_e.

The posterior distribution is sampled via data re-sampling, whereby multiple maximum likelihood optimizations are carried out by adding Gaussian noise with width $\alpha_{i,e}$ to each data point across all data sets. The resulting ensemble of n parameter samples $\{\boldsymbol{a}_k; k = 1, \ldots, n\}$, referred to as a "replica," is then used to obtain statistical estimators for a given non-perturbative function or observable, generically defined as

$\mathcal{O}(\boldsymbol{a})$, such as the mean and variance,

$$\mathrm{E}[\mathcal{O}] = \frac{1}{n}\sum_k \mathcal{O}(\boldsymbol{a}_k), \tag{2.4a}$$

$$\mathrm{V}[\mathcal{O}] = \frac{1}{n}\sum_k \big[\mathcal{O}(\boldsymbol{a}_k) - \mathrm{E}[\mathcal{O}]\big]^2. \tag{2.4b}$$

In the chapters to follow, the agreement between data and theory will be assessed by defining the "reduced" χ^2 for each dataset e as

$$\chi^2_{\mathrm{red},e} \equiv \frac{1}{N^e_{\mathrm{dat}}}\sum_i \left(\frac{d_{i,e} - \mathrm{E}\big[\sum_k r^k_e \beta^k_{i,e} + T_{i,e}/N_e\big]}{\alpha_{i,e}}\right)^2 + \sum_k \big(r^k_e\big)^2, \tag{2.5}$$

with N^e_{dat} the total number of data points for each experiment, and E[...] represents the mean theory as defined in Eq. (2.4a). This definition of the χ^2 is more convenient compared to Eq. (2.3) as it condenses the χ^2 distribution of the replicas (of which there are generally $\mathcal{O}(1000)$) into a single number.

Another aspect of the following global analyses is data resampling, which is used to propagate the errors on the data to the extracted distributions, as an alternative to the Hessian [98] or Lagrange multiplier [99] methods. Instead of using the data directly, we generate "pseudo-data." Dropping the i, e subscripts, each pseudo-data point $\tilde{d}$ is generated through [100]

$$\tilde{d} = (1 + R^{\mathrm{nor}}\delta N)\left[d + R^{\mathrm{uncorr}}\alpha + \sum_k R^{\mathrm{corr}}\beta^k\right], \tag{2.6}$$

where R^{nor}, R^{uncorr}, and R^{corr} are random numbers from the normal distribution. This recipe takes into account all three types of errors defined in Eq. (2.3). This pseudo-data generation process is repeated for each replica, so that each replica has a corresponding unique set of pseudo-data. In this way the errors on the experimental data are propagated to the resulting distributions.

The final aspect that we discuss here is the "multi-step strategy" of extracting distributions, with a simplified example shown in Table 2.1. The essence of this strategy is to start with a minimal amount of data and parameters and to add more data and more parameters in subsequent steps. This eases the computational burden of the analysis, and also makes it easier for the minimization algorithm to find the global minimum so that we are less likely to end up with "failed" replicas with very

step	data	fitted distributions	fixed distributions
01	DIS	PDFs	—
02	+Drell-Yan	PDFs	—
03	$+W/Z$ production	PDFs	—
04	+jet production	PDFs	—
05	SIA (π)	π FFs	PDFs
06	+SIDIS (π)	π FFs	PDFs
07	SIA (K)	K FFs	PDFs, π FFs
08	+SIDIS (K)	K FFs	PDFs, π FFs
09	Inclusive + π	PDFs, π FFs	K FFs
10	Inclusive + π + K	PDFs, π FFs, K FFs	—

Table 2.1: Multi-step process. A simplified example of the multi-step process. We start by fitting only DIS data with PDF parameters free to vary. In subsequent steps, we add more data and fit more distributions. In some steps, some of the distributions are fixed.

large χ^2. In the simplified example, we start only with DIS data and with only fitting the spin-averaged PDFs. Since DIS is a computationally simple observable (relative to the others) and there are not too many parameters, the fit will converge quickly and most of the replicas will succeed in finding the global minimum of the χ^2. While the PDFs started with random parameters, after this first step they will now reside in a "reasonable" region of the parameter space that is capable of describing the DIS data. This helps in the following steps where we add data that is more computationally expensive, as the PDFs are no longer starting off with a completely random guess. In later steps we begin fitting FFs with the PDFs fixed in order to get them also into a reasonable position. With all of the distributions in a reasonable position, in the final steps we fit all of the data with all of the parameters free to vary. By choosing these steps intelligently, we are able to break a computationally intractable problem into multiple steps and thus make it possible.

2.2 Processes

In this analysis of spin-averaged PDFs we include inclusive data on DIS, Drell-Yan, W-lepton production, and jet production. Further, to constrain the FFs, we include

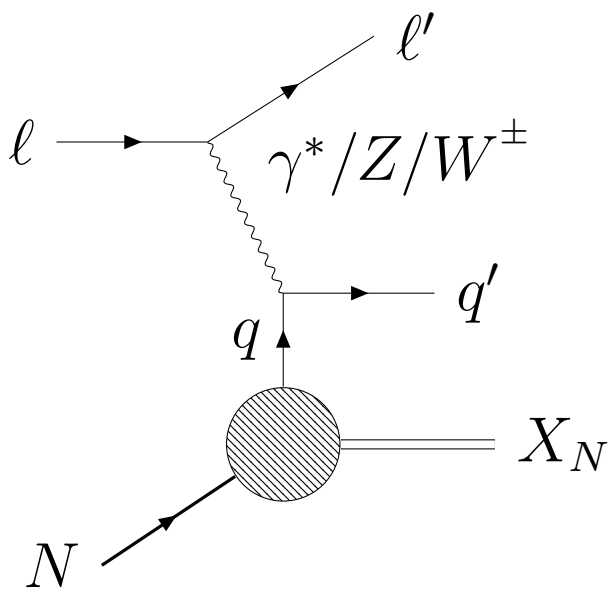

Figure 2.1: Deep inelastic scattering process. The process $\ell N \to \ell' X$ in the approximation of the exchange of a single electroweak gauge boson at LO in perturbative QCD.

semi-inclusive DIS and e^+e^- annihilation data. In the subsections below we summarize each of these processes and the relevant equations.

2.2.1 Deep Inelastic Scattering

In deep inelastic scattering, a high-energy lepton ℓ with four-momentum k scatters off of a nucleon or nucleus target and is detected (or reconstructed if the outgoing lepton is a neutrino) with four-momentum k'. The target N has four-momentum P, while the exchanged particle, which may be a photon (γ), Z boson, or $W^\pm$ boson, has four-momentum $q = k - k'$. The recoiling system X (which, in this and following figures, consists of all of the X shown in the figure as well as the remnants of any outgoing quark) has invariant mass W. This process, $\ell N \to \ell' X$, is shown in Fig. 2.1. In the unpolarized case, both the spin of the lepton and target are averaged over. In this subsection it will be assumed that the target is a proton of mass M.

Neglecting the mass of the lepton, one can define the invariant quantities:

$$s \equiv (P+k)^2, \quad Q^2 \equiv -q^2, \quad x_{\rm bj} \equiv \frac{Q^2}{2P\cdot q} = \frac{Q^2}{2M\nu}, \quad y \equiv \frac{P\cdot q}{P\cdot k} = \frac{Q^2}{x_{\rm bj}(s-M^2)}, \tag{2.7}$$

where Q^2 is the four momentum transfer squared, $x_{\rm bj}$ is the Bjorken scaling variable, y is the inelasticity and s is the squared COM energy of the reaction. The quantity ν is given by $\nu \equiv E - E'$, where E (E') is the energy of the initial (final) lepton in the target rest frame. We note that beyond s there are only two independent variables in the case of averaged or longitudinal polarizations for which one could choose $x_{\rm bj}$ and Q^2. The cross section for this process can then be written as [11]:

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}x_{\rm bj}\mathrm{d}y} = \frac{2\pi y\alpha^2}{Q^4}\sum_j \eta_j L^{\mu\nu}_{U,j} W^U_{\mu\nu}, \tag{2.8}$$

where α is the electromagnetic fine structure constant, $L^{\mu\nu}_{U,j}$ is the unpolarized leptonic tensor, and $W^U_{\mu\nu}$ is the unpolarized hadronic tensor. The summation is over $j = \gamma, \gamma Z, Z$ for the neutral current (NC) process, and $j = W$ for the charged current (CC) process. The factors η_j are given by:

$$\begin{aligned}
\eta_\gamma &= 1, \\
\eta_{\gamma Z} &= \left(\frac{G_F M_Z^2}{2\sqrt{2}\pi\alpha}\right)\left(\frac{Q^2}{Q^2 + M_Z^2}\right), \\
\eta_Z &= \eta_{\gamma Z}^2, \\
\eta_W &= \frac{1}{2}\left(\frac{G_F M_W^2}{4\pi\alpha}\frac{Q^2}{Q^2 + M_W^2}\right)^2,
\end{aligned} \tag{2.9}$$

where G_F is the Fermi constant, M_Z is the mass of the Z boson, and M_W is the mass of the W boson.

For spin-averaged electrons and NC, the unpolarized leptonic tensors are given by

$$\begin{aligned}
L^{U,\gamma}_{\mu\nu} &= 2(k_\mu k'_\nu + k'_\mu k_\nu - g_{\mu\nu} k\cdot k'), \\
L^{U,\gamma Z}_{\mu\nu} &= g_V^e L^{U,\gamma}_{\mu\nu}, \\
L^{U,Z}_{\mu\nu} &= \left[(g_V^e)^2 + (g_A^e)^2\right] L^{U,\gamma}_{\mu\nu},
\end{aligned} \tag{2.10}$$

where g_V^e and g_A^e are the vector and axial couplings, respectively, of the electron to the Z boson. g_V^e can be written in terms of the weak mixing angle θ_W as $g_V^e = -\frac{1}{2} + 2\sin^2\theta_W$, while $g_A^e = -\frac{1}{2}$. The unpolarized hadronic tensor can be defined in

terms of the proton structure functions $F^p_{1,2,3}$ as:

$$W^U_{\mu\nu} = \Big(\frac{q_\mu q_\nu}{q^2} - g_{\mu\nu}\Big)F^p_1 + \Big(P^\mu + \frac{q^\mu}{2x_{\rm bj}}\Big)\Big(P^\nu + \frac{q^\nu}{2x_{\rm bj}}\Big)\frac{F^p_2}{P\cdot q} - i\epsilon_{\mu\nu\alpha\beta}\frac{q^\alpha P^\beta}{2P\cdot q}F^p_3, \quad (2.11)$$

where $\epsilon_{\mu\nu\alpha\beta}$ is the Levi-Civita tensor. Note the coefficients of F^p_1 and F^p_2 are symmetric under $\mu \leftrightarrow \nu$, while the coefficient of F^p_3 is anti-symmetric. The $F_{1,2}$ are parity-conserving structure functions, while F_3 is parity-violating. Performing the contraction in Eq. (2.8), one arrives at the spin-averaged NC DIS cross section

$$\frac{{\rm d}^2\sigma^{\rm NC}}{{\rm d}x_{\rm bj}{\rm d}y} = \frac{4\pi\alpha^2}{x_{\rm bj}yQ^2}\left[\Big(1 - y - \frac{x^2_{\rm bj}y^2M^2}{Q^2}\Big)F^p_2 + y^2x_{\rm bj}F^p_1 \mp \Big(y - \frac{y^2}{2}\Big)x_{\rm bj}F^p_3\right], \quad (2.12)$$

where the plus (minus) is taken for negative (positive) lepton scattering. Note that a sum over the γ, γZ, and Z channels has been performed such that the structure functions above are defined in terms of the individual channel structure functions $F^{p,\gamma}_{1,2,3}$, $F^{p,\gamma Z}_{1,2,3}$, and $F^{p,Z}_{1,2,3}$ as:

$$\begin{aligned} F^p_{1,2} &= F^{p,\gamma}_{1,2} - g^e_V\eta_{\gamma Z}F^{p,\gamma Z}_{1,2} + \big((g^e_V)^2 + (g^e_A)^2\big)\eta_Z F^{p,Z}_{1,2}, \\ F^p_3 &= -g^e_A\eta_{\gamma Z}F^{p,\gamma Z}_3 + 2g^e_Vg^e_A\eta_Z F^{p,Z}_3. \end{aligned} \quad (2.13)$$

For the HERA data used in this fit, the measured observable is the reduced cross section, defined as

$$\sigma^{\rm NC}_{\rm red} \equiv \frac{x_{\rm bj}yQ^2}{2\pi\alpha^2Y^+}\frac{{\rm d}^2\sigma^{\rm NC}}{{\rm d}x_{\rm bj}{\rm d}y} \approx F^p_2 - \frac{y^2}{Y^+}F^p_L \pm x_{\rm bj}\frac{Y^-}{Y^+}F^p_3, \quad (2.14)$$

where

$$Y^\pm \equiv 1 \pm (1-y)^2, \quad (2.15)$$

and

$$F^p_L(x_{\rm bj}, Q^2) \equiv \rho^2F^p_2 - 2x_{\rm bj}F^p_1, \quad (2.16)$$

where

$$\rho^2 \equiv 1 + \frac{4x^2_{\rm bj}M^2}{Q^2}, \quad (2.17)$$

and the $\frac{x_{\rm bj}^2 y^2 M^2}{Q^2}$ term in Eq. (2.12) has been neglected since at the relevant kinematics $x_{\rm bj}$ is small and $M^2 \ll Q^2$.

Expanding in powers of $\frac{M^2}{Q^2}$ and $\frac{\Lambda^2_{\rm QCD}}{Q^2}$ the NC structure functions can be written in terms of the spin-averaged PDFs q, g as:

$$\begin{aligned} F_1^p(x_{\rm bj},Q^2) &= \frac{1}{2}\sum_q e_q^2\big[C_{1,q}^{\rm DIS}\otimes q^+ + C_{1,g}^{\rm DIS}\otimes g\big] + \mathcal{O}\Big(\frac{[M^2,\Lambda^2_{\rm QCD}]}{Q^2}\Big), \\ F_2^p(x_{\rm bj},Q^2) &= x\sum_q e_q^2\big[C_{2,q}^{\rm DIS}\otimes q^+ + C_{2,g}^{\rm DIS}\otimes g\big] + \mathcal{O}\Big(\frac{[M^2,\Lambda^2_{\rm QCD}]}{Q^2}\Big), \\ F_3^p(x_{\rm bj},Q^2) &= \sum_q e_q^2\big[C_{3,q}^{\rm DIS}\otimes q^-\big] + \mathcal{O}\Big(\frac{[M^2,\Lambda^2_{\rm QCD}]}{Q^2}\Big), \end{aligned} \tag{2.18}$$

where the sum q runs over all quark flavors, e_q is the charge of the quark of flavor q, and $q^\pm \equiv q \pm \bar{q}$. The $1/Q^2$ corrections, which include target mass and higher twist corrections, are not relevant for the results discussed in this chapter. They will be discussed in Chapter 3 in relation to nuclear corrections. The DIS hard scattering coefficients, $C_{i,f}^{\rm DIS}$ with $i = 1,2,3$ and $f = q, g$, are expanded to NLO in the strong coupling constant $\alpha_s(\mu_R)$ as

$$C_{i,f}^{\rm DIS} = C_{i,f}^{\rm DIS,(0)} + \frac{\alpha_s(\mu_R)}{4\pi} C_{i,f}^{\rm DIS,(1)} + \mathcal{O}(\alpha_s^2),$$

with the coefficients taken from [49]. The hard scattering coefficients depend on the renormalization scale, while the PDFs depend on the factorization scale, both of which are taken to be $\mu_R = \mu_F = Q$ for the DIS process. For the neutron functions $F_{1,2,3}^n$, the same proton PDFs are used except with the switch $u \leftrightarrow d$ that is derived from isospin symmetry. Nuclear structure functions $F_{2,L,3}^A$ will be discussed in Chapter 3.

For the charged current process, which has been measured at HERA, the only channel is $W^\pm$ for which the leptonic tensor is given by

$$L_{\mu\nu}^{U,W} = 4(k_\mu k'_\nu + k'_\mu k_\nu - g_{\mu\nu} k\cdot k'). \tag{2.19}$$

Substitution into Eq. (2.8) then leads to

$$\frac{{\rm d}^2\sigma^{{\rm CC},W^\pm}}{{\rm d}x_{\rm bj}{\rm d}y} = \frac{8\pi\alpha^2}{x_{\rm bj} y Q^2}\eta_W\Big[(1-y-\frac{x_{\rm bj}^2 y^2 M^2}{Q^2})F_2^{p,W^\pm} + y^2 x_{\rm bj} F_1^{p,W^\pm}$$

$$\mp (y - \frac{y^2}{2}) x_{\rm bj} F_3^{p,W^\pm} \Big], \tag{2.20}$$

and for the reduced cross section

$$\sigma_{\rm red}^{{\rm CC},W^\pm} \equiv \frac{x_{\rm bj} y Q^2}{16\pi\alpha^2\eta_W} \frac{{\rm d}^2\sigma^{{\rm CC},W^\pm}}{{\rm d}x_{\rm bj}{\rm d}y} \approx \frac{Y^+}{4} F_2^{p,W^\pm} - \frac{y^2}{4} F_L^{p,W^\pm} \mp x_{\rm bj} \frac{Y^-}{4} F_3^{p,W^\pm}, \tag{2.21}$$

where $F_{1,2,3}^{W^\pm}$ are the CC structure functions and again the $\frac{x_{\rm bj}^2 y^2 M^2}{Q^2}$ term has been neglected. The CC structure functions can be expanded similarly to the NC structure functions as

$$\begin{aligned}
F_1^{p,W^\pm}(x_{\rm bj}, Q^2) &= \sum_q \left[C_{1,q}^{\rm DIS} \otimes q + C_{1,g}^{\rm DIS} \otimes g \right] + \mathcal{O}\Big(\frac{[M^2, \Lambda_{\rm QCD}^2]}{Q^2} \Big), \\
F_2^{p,W^\pm}(x_{\rm bj}, Q^2) &= 2x \sum_q \left[C_{2,q}^{\rm DIS} \otimes q + C_{2,g}^{\rm DIS} \otimes g \right] + \mathcal{O}\Big(\frac{[M^2, \Lambda_{\rm QCD}^2]}{Q^2} \Big), \\
F_3^{p,W^\pm}(x_{\rm bj}, Q^2) &= 2 \sum_q \left[I_q C_{3,q}^{\rm DIS} \otimes q \right] + \mathcal{O}\Big(\frac{[M^2, \Lambda_{\rm QCD}^2]}{Q^2} \Big),
\end{aligned} \tag{2.22}$$

where the summation is over $q = \bar{u}, d, s, \bar{c}, b, \bar{t}$ for W^+ and $q = u, \bar{d}, \bar{s}, c, \bar{b}, t$ for W^-. The factor I_q in F_3 is 1 for quarks and -1 for antiquarks. The DIS hard scattering kernels are the same as those in Eq. (2.18).

2.2.2 Drell-Yan Processes

Included in this analysis are also data from proton-proton, proton-antiproton, and proton-deuteron collisions. The first type of collision that we discuss is the Drell-Yan process given by $h_A h_B \to \gamma^*/ZX \to \ell\bar{\ell}X$. This process is shown in Fig. 2.2. In this analysis h_A denotes a proton while h_B may denote a proton, antiproton, or deuteron for the Drell-Yan processes. Furthermore, ℓ and $\bar{\ell}$ denote a lepton anti-lepton pair, while X denotes the undetected part of the final system. At leading order, a quark (q) and antiquark ($\bar{q}$) from the incoming particles annihilate to create a photon or Z boson.

The double differential Drell-Yan lepton-pair cross section can be written in terms of convolutions of the PDFs in the colliding hadrons with short-distance partonic cross

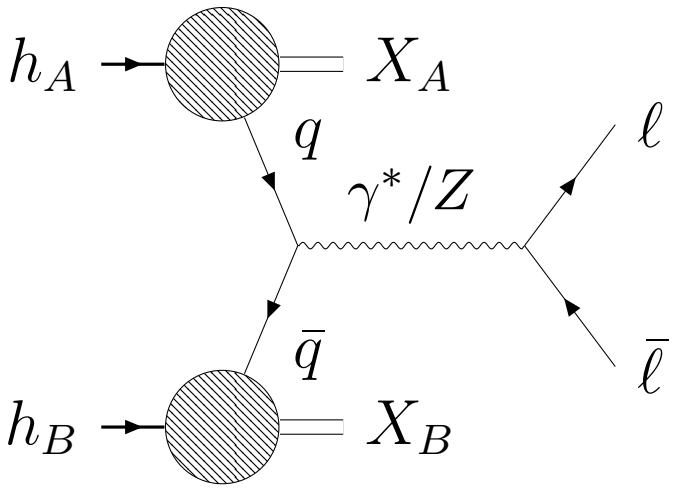

Figure 2.2: Drell-Yan process. The process $h_A h_B \to \gamma^*/ZX \to \ell\bar{\ell}X$ at LO in perturbative QCD.

sections $\hat{\sigma}_{ab}^{\mathrm{DY}}$ [101] as

$$\frac{\mathrm{d}^2\sigma^{\mathrm{DY}}}{\mathrm{d}M_{\ell\ell}^2\,\mathrm{d}Y} = \frac{4\pi\alpha^2}{9sM_{\ell\ell}^2}\sum_{ab}\int \mathrm{d}x_1 \int \mathrm{d}x_2 \times f_a(x_1,\mu_F)\, f_b(x_2,\mu_F)\, \hat{\sigma}_{ab}^{\mathrm{DY}}(x_1,x_2,s,M_{\ell\ell},\mu_R,\mu_F), \tag{2.23}$$

where x_1 (x_2) is the momentum fraction of the parton coming from h_A (h_B), α is the electromagnetic coupling, s is the invariant mass squared of the reaction, and μ_R and μ_F are the renormalization and factorization scales, respectively. We write the cross section as differential in the invariant mass of the lepton pair squared, $M_{\ell\ell}^2$, and the rapidity of the lepton pair in the COM frame, Y. For the Drell-Yan process, the "Feynman-x" variable

$$x_F \equiv x_1 - x_2, \tag{2.24}$$

is also commonly used. We note that at LO in α_s, one has $x_1 = \frac{M_{\ell\ell}}{\sqrt{s}}e^{Y}$ and $x_2 = \frac{M_{\ell\ell}}{\sqrt{s}}e^{-Y}$, and thus $x_F = 2\frac{M_{\ell\ell}}{\sqrt{s}}\cosh Y$.

The sum over the quark flavors a, b runs over all partonic channels that can contribute to the Drell-Yan process, for which the scale is set to $\mu_R = \mu_F = M_{\ell\ell}$. The partonic cross sections $\hat{\sigma}_{ab}^{\mathrm{DY}}$ are computed to NLO in the strong coupling $\alpha_s(\mu_R)$, with the NLO expressions taken from Ref. [101]. For pp scattering, the PDFs $f_{a,b}$ are those

in the proton. For proton-deuteron scattering, the x of the parton in the deuteron target is small enough that nuclear corrections can be neglected and it is reasonable to approximate the pD cross section by a simple sum of proton and neutron cross sections [102],

$$\sigma_{pD} \approx \sigma_{pp} + \sigma_{pn}, \tag{2.25}$$

with the PDFs in the neutron related to those in the proton through isospin symmetry (see Eq. (1.5)).

We now discuss specifically the fixed-target experiments performed by the NuSea and SeaQuest collaborations, defining h_A to be the target and h_B to be the beam. They provide the cross sections σ_{pp} and σ_{pD}. Alternatively (or in addition), they provide the ratio $\sigma_{pD}/2\sigma_{pp}$. The experiments were designed to be most sensitive to the $x_1 \gg x_2$ regime (with x_1 still small enough so that nuclear corrections are negligible), where, keeping only the up and down (anti)quarks, one has

$$\begin{aligned}
\sigma_{pp} &\propto \frac{4}{9}u(x_1)\bar{u}(x_2) + \frac{1}{9}d(x_1)\bar{d}(x_2), \\
\sigma_{pn} &\propto \frac{4}{9}u(x_1)\bar{d}(x_2) + \frac{1}{9}d(x_1)\bar{u}(x_2).
\end{aligned}$$

Using Eq. (2.25), one then has for the ratio

$$\frac{\sigma_{pD}}{2\sigma_{pp}} \approx \frac{1}{2}\frac{[1+\frac{1}{4}\frac{d(x_1)}{u(x_1)}]}{[1+\frac{1}{4}\frac{d(x_1)}{u(x_1)}\frac{\bar{d}(x_2)}{\bar{u}(x_2)}]}[1+\frac{\bar{d}(x_2)}{\bar{u}(x_2)}]. \tag{2.26}$$

Finally, with $d(x) \ll 4u(x)$, one has [94]

$$\left.\frac{\sigma_{pD}}{2\sigma_{pp}}\right|_{x_1 \gg x_2} \approx \frac{1}{2}\Big[1+\frac{\bar{d}(x_2)}{\bar{u}(x_2)}\Big]. \tag{2.27}$$

Thus the ratio is almost directly sensitive to the sea asymmetry $\bar{d}/\bar{u}$ (or $\bar{d}-\bar{u}$) and is an important tool for its extraction. The impact of including the Drell-Yan measurement for the sea asymmetry will be shown in Section 2.6.

Measurements of the W-lepton production process are also included in this analysis, which includes the processes $h_A h_B \to W^- X \to \ell\bar{\nu}_\ell X$ (shown in Fig. 2.3) and $h_A h_B \to W^+ X \to \bar{\ell}\nu_\ell X$. In these processes, h_A always denotes proton while h_B may denote a proton (if the data is from the LHC or RHIC) or an antiproton (if the data

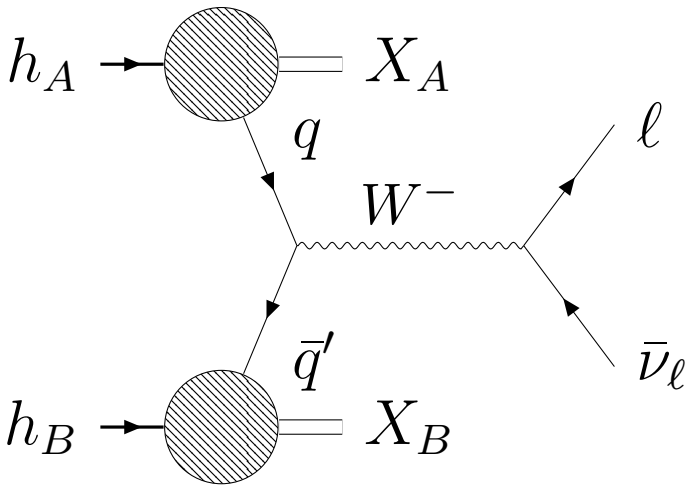

Figure 2.3: W boson production process. The W^- boson production process $pp \to W^- X \to \ell\ \bar{\nu}_\ell X$ at LO in perturbative QCD. The W^+ boson production process is similar, but with $\ell \to \bar{\ell}$ and $\bar{\nu}_\ell \to \nu_\ell$.

is from the Tevatron). ℓ ($\bar{\ell}$) denotes a lepton (anti-lepton), while ν_ℓ ($\bar{\nu}_\ell$) denotes the corresponding neutrino (anti-neutrino). At leading order, a quark and antiquark from the incoming particles annihilate to create a W boson.

We first focus specifically on the measurements from the Tevatron on reconstructed W and Z bosons, where h_B is an antiproton. For the Z boson production observable, the calculation is the same as that for Drell-Yan, except that the result is integrated over $M_{\ell\ell}$, generally from $60 \lesssim M_{\ell\ell} \lesssim 120$ GeV, in order to isolate the Z boson contribution. For the W boson production process the calculation is similar, with the channels adjusted appropriately and weighted with the CKM matrix elements. The result is also integrated over $0 < M_{\ell\ell} < \sqrt{s}$. The observable measured in this case is the W boson asymmetry

$$A_W(Y_W) = \frac{\mathrm{d}\sigma^{W^+}/\mathrm{d}Y_W - \mathrm{d}\sigma^{W^-}/\mathrm{d}Y_W}{\mathrm{d}\sigma^{W^+}/\mathrm{d}Y_W + \mathrm{d}\sigma^{W^-}/\mathrm{d}Y_W}, \tag{2.28}$$

where Y_W is the rapidity of the W boson and $\mathrm{d}\sigma^{W^\pm}/\mathrm{d}Y_W$ is the differential cross section for $W^\pm$ production. At LO and considering only the up and down quarks, one has

$$A_W(Y_W) \approx \frac{u(x_1)\bar{d}(x_2) + u(x_2)\bar{d}(x_1) - \bar{u}(x_1)d(x_2) - \bar{u}(x_2)d(x_1)}{u(x_1)\bar{d}(x_2) + u(x_2)\bar{d}(x_1) + \bar{u}(x_1)d(x_2) + \bar{u}(x_2)d(x_1)}.$$

At large Y_W, one has $x_1 \gg x_2$, and so:

$$A_W(Y_W) \xrightarrow[\text{large } Y_W]{} \frac{u(x_1)\bar{d}(x_2) - \bar{u}(x_2)d(x_1)}{u(x_1)\bar{d}(x_2) + \bar{u}(x_2)d(x_1)}.$$

Approximating now that $\bar{u} \approx \bar{d}$ at small x, one arrives at

$$A_W(Y_W) \xrightarrow[\text{large } Y_W]{} \frac{1 - d(x_1)/u(x_1)}{1 + d(x_1)/u(x_1)}. \tag{2.29}$$

Since large Y_W corresponds to large x_1, the W boson asymmetry provides strong constraints on the d/u ratio at large x.

In contrast to the Tevatron, the measurements from the LHC are differential in the outgoing lepton's kinematics. In this case, the double differential cross section is given by [103]

$$\begin{aligned}\frac{\mathrm{d}^2\sigma^W}{\mathrm{d}p_T^\ell\,\mathrm{d}\eta_\ell} = \frac{2}{p_T^\ell}\sum_{ab}\int_{VW}^{V}\mathrm{d}v\int_{VW/v}^{1}\mathrm{d}w \\ \times\, x_1 f_a(x_1,\mu_F)\, x_2 f_b(x_2,\mu_F)\,\hat{\sigma}_{ab}^W(v,w,s,\mu_R,\mu_F),\end{aligned} \tag{2.30}$$

which is differential in the outgoing lepton's pseudorapidity, η_ℓ, and its transverse momentum, p_T^ℓ. Labeling the momenta of the incoming hadrons by P_1 and P_2, we define the Mandelstam invariants $T \equiv (P_1 - p_\ell)^2$ and $U \equiv (P_2 - p_\ell)^2$, which then allows us to introduce the variables $V \equiv 1 + T/s$ and $W \equiv -U/(s+T)$, where s is the COM energy of the proton-proton collision. Defining v and w as the partonic analogues of V and W, we can write for the parton momentum fractions $x_1 = VW/vw$ and $x_2 = (1-V)/(1-v)$. The sum over the quark flavors a, b runs over all partonic channels that can contribute to W-lepton production, for which the renormalization and factorization scales are chosen to be the mass of the W boson, $\mu_R = \mu_F = M_W$. The partonic cross sections $\hat{\sigma}_{ab}^W$ are again computed at NLO in the strong coupling $\alpha_s(\mu_R)$, with the NLO expressions used here taken from Ref. [103]. Taking the ratio of the W^+ and W^- cross sections for pp reactions, at LO and considering only up and down (anti)quarks, one has

$$\frac{\sigma_{pp}^{W^+}}{\sigma_{pp}^{W^-}} \approx \frac{u(x_1)\bar{d}(x_2) + u(x_2)\bar{d}(x_1)}{d(x_1)\bar{u}(x_2) + d(x_2)\bar{u}(x_1)}. \tag{2.31}$$

This observable, in addition to the DY observable Eq. (2.27), also provides sensitivity

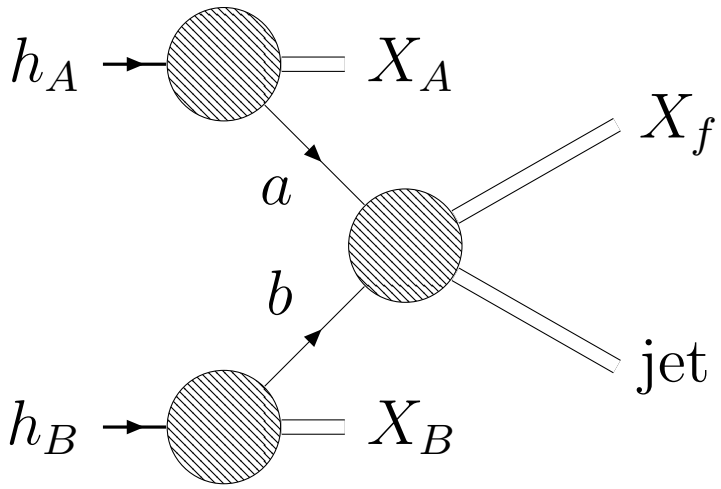

Figure 2.4: Jet production process. The process $h_A h_B \to \text{jet } X$ at leading order.

to the light antiquarks and thus the sea asymmetry.

2.2.3 Jet Production

Measurements of the jet production process $h_A h_B \to \text{jet X}$, shown in Fig. 2.4 are also included in this analysis. In this process, h_A always denotes a proton while h_B may denote a proton (if the data is from RHIC) or an antiproton (if the data is from the Tevatron). X denotes the undetected part of the final system. A quark or gluon (a and b) from the two incoming particles interact through a hard scattering, and one outgoing parton is the seed of a single jet. For jet production, the double differential cross section is given by [104]

$$\frac{\mathrm{d}^2\sigma^{\text{jet}}}{\mathrm{d}p_T^{\text{jet}}\,\mathrm{d}\eta_{\text{jet}}} = \frac{2}{p_T^{\text{jet}}} \sum_{ab} \int_{VW}^{V} \mathrm{d}v \int_{VW/v}^{1} \mathrm{d}w \times x_1 f_a(x_1,\mu_F)\, x_2 f_b(x_2,\mu_F)\, \hat{\sigma}_{ab}^{\text{jet}}(v,w,p^{\text{jet}},\mu_R,\mu_F;r), \tag{2.32}$$

which is differential in the outgoing jet's pseudorapidity, η_{jet}, and its transverse momentum, p_T^{jet}. Here p^{jet} is the momentum of the jet and r the jet radius. The variables V, W, v and w are defined in Section 2.2.2. Note that Eq. (2.32) is very similar to Eq. (2.30). The sum over the quark flavors a, b runs over all partonic channels that can contribute to jet production, for which the renormalization and factorization scales are chosen to be the jet's transverse momentum, $\mu_R = \mu_F = p_T^{\text{jet}}$. The partonic cross

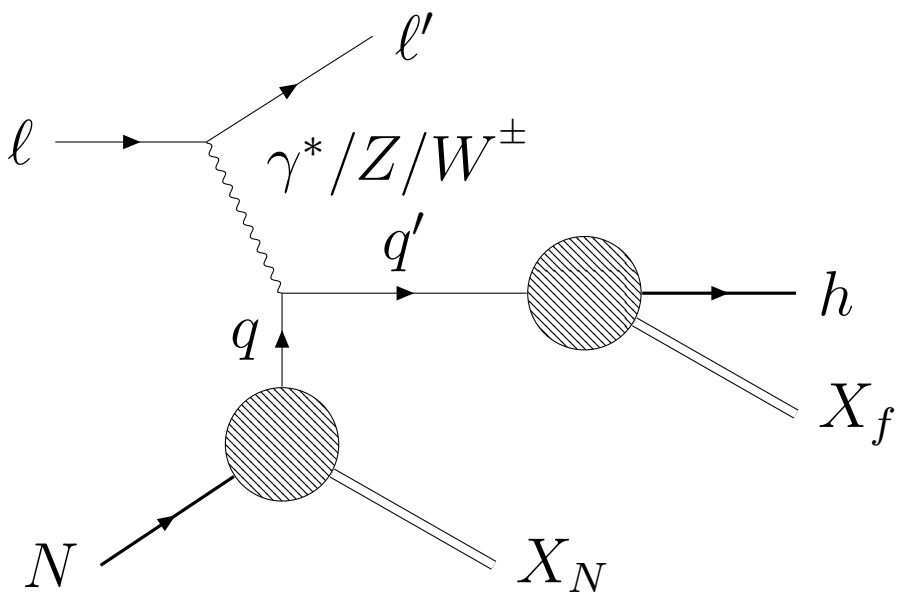

Figure 2.5: Semi-inclusive deep inelastic scattering process. The process $\ell N \to \ell' h X$ with a single photon exchange at leading order in perturbative QCD.

sections $\hat{\sigma}_{ab}^{\text{jet}}$ are again computed at NLO in the strong coupling $\alpha_s(\mu_R)$, with the NLO expressions used here taken from Refs. [104, 105]. We note that a, b can be a gluon, and thus the single jet production process is sensitive to the gluon distribution even at LO.

2.2.4 Semi-Inclusive Deep Inelastic Scattering

The semi-inclusive DIS (SIDIS) process $\ell N \to \ell' h X$ is shown in Fig. 2.5. It is identical to the DIS process (Subsection 2.2.1), except that in the final state, in addition to the scattered lepton, a hadron is detected with momentum P_h. In this process, we define the "fragmentation variable"

$$z_H \equiv \frac{P_h \cdot P}{q \cdot P}. \tag{2.33}$$

One can define structure functions analogously to Eq. (2.18):

$$F_1^{h(p)}(x_{\text{bj}}, z_H, Q^2) = \frac{1}{2} \sum_{q=u,\bar{u},\ldots,\bar{s}} e_q^2 \big[C_{1,qq}^{\text{SIDIS}} \otimes q \otimes D_q^h + C_{1,gq}^{\text{SIDIS}} \otimes q \otimes D_g^h$$

$$+C_{1,qg}^{\text{SIDIS}} \otimes g \otimes D_q^h\big] + \mathcal{O}\Big(\frac{[M^2, \Lambda_{\text{QCD}}^2]}{Q^2}\Big), \tag{2.34a}$$

$$F_2^{h(p)}(x_{\text{bj}}, z_H, Q^2) = x \sum_{q=u,\bar{u},\ldots,\bar{s}} e_q^2 \big[C_{2,qq}^{\text{SIDIS}} \otimes q \otimes D_q^h + C_{2,gq}^{\text{SIDIS}} \otimes q \otimes D_g^h$$

$$+C_{2,qg}^{\text{SIDIS}} \otimes g \otimes D_q^h\big] + \mathcal{O}\Big(\frac{[M^2, \Lambda_{\text{QCD}}^2]}{Q^2}\Big), \tag{2.34b}$$

where the sum q runs over all quark flavors. The SIDIS hard scattering coefficients, $C_{i,ff'}^{\text{SIDIS}}$ with $i = 1, 2$ and $ff' = qq, gq, qg$, are expanded to NLO in the strong coupling constant $\alpha_s(\mu_R)$ with the coefficients taken from [106]. The hard scattering coefficients depend on the renormalization scale, while the PDFs and FFs depend on the factorization scale, both of which are taken to be $\mu_R = \mu_F = Q$ for the SIDIS process (as in the DIS process). For the neutron functions $F_{1,2}^{n(h)}$, the same proton PDFs are used except with the switch $u \leftrightarrow d$ that is derived from isospin symmetry (see Eq. (1.5)). For the deuterium structure function we ignore nuclear effects and simply take the sum of the proton and neutron structure functions divided by $A = 2$. Analogous to the inclusive DIS case in Eq. (2.16), one also has $F_L^{p(h)} \equiv \rho^2 F_2^{p(h)} - 2x_{\text{bj}} F_1^{p(h)}$.

The observable measured in SIDIS experiments is the multiplicity M^h, defined as the differential cross section for hadron production normalized to the differential inclusive DIS cross section. Ignoring terms of order M^2/Q^2, it can be written in terms of the structure functions as [107–111]

$$\frac{\mathrm{d}M^h}{\mathrm{d}z_h} = x_{\text{bj}} \frac{Y^+ F_1^h + (1-y) F_L^h}{x_{\text{bj}} Y^+ F_1 + (1-y) F_L}, \tag{2.35}$$

with Y^+ defined as in Eq. (2.15).

2.2.5 Semi-Inclusive Annihilation

The semi-inclusive annihilation (SIA) process $e^+e^- \to hX$ is shown in Fig. 2.6. In this process two electrons annihilate to form a photon or Z boson which splits into a $q\bar{q}$ pair. The quarks then hadronize to form a detected hadron with momentum P_h and the undetected final system X. In this process, we define the "fragmentation variable"

$$z_h \equiv \frac{2P_h \cdot q}{s}, \tag{2.36}$$

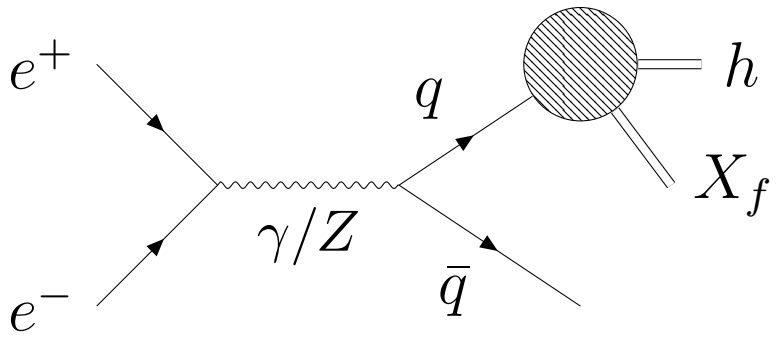

Figure 2.6: Semi-inclusive annihilation process. The process $e^+e^- \to hX$ at leading order in perturbative QCD.

where s is the COM energy of the colliding electrons. One can define structure functions analogously to Eq. (2.18):

$$F_1^h(z_h, Q^2) = \frac{1}{2}\sum_q e_q^2 \big[C_{1,q}^{\text{SIA}} \otimes D_{q^+}^h + C_{1,g}^{\text{SIA}} \otimes D_g^h\big] + \mathcal{O}\Big(\frac{\Lambda_{\text{QCD}}^2}{Q^2}\Big), \tag{2.37}$$

$$F_L^h(z_h, Q^2) = x\sum_q e_q^2 \big[C_{L,q}^{\text{SIA}} \otimes D_{q^+}^h + C_{L,g}^{\text{SIA}} \otimes D_g^h\big] + \mathcal{O}\Big(\frac{\Lambda_{\text{QCD}}^2}{Q^2}\Big), \tag{2.38}$$

where the sum q runs over all quark flavors, e_q is the charge of the quark of flavor q, and $Q^2 = s$. The SIA hard scattering coefficients, $C_{i,f}^{\text{SIA}}$ with $i = 1, L$ and $f = q, g$, are expanded to NLO in the strong coupling constant $\alpha_s(\mu_R)$ with the coefficients taken from [107]. The hard scattering coefficients depend on the renormalization scale, while the FFs depend on the factorization scale, both of which are taken to be $\mu_R = \mu_F = \sqrt{s}$ for the SIA process. The measured observable in SIA is given by [108–110]

$$\frac{\mathrm{d}\sigma^{\text{SIA}}}{\mathrm{d}z_h} = \sigma_0\Big[2F_1^h + F_L^h\Big], \tag{2.39}$$

where $\sigma_0 \equiv 4\pi\alpha^2/s$.

2.3 PDF Parametrization

Our PDF extraction procedure is based on Bayesian inferences using Monte Carlo techniques developed in previous JAM analyses [112–116]. We parametrize the PDFs at the input renormalization scale $\mu_0^2 = m_c^2$, where m_c is the mass of the charm quark. As the input scale decreases, the errors from the backward evolution increase due to the truncation of the perturbative series. This scale is chosen to be as small as possible while keeping these errors under control [113]. The PDFs are parametrized using a generic template function of the form

$$f(x, \mu_0^2) = \frac{N}{\mathcal{M}} x^\alpha (1-x)^\beta (1 + \gamma\sqrt{x} + \eta x), \tag{2.40}$$

where $\boldsymbol{a} = \{N, \alpha, \beta, \gamma, \eta\}$ is the set of parameters to be inferred, and $\mathcal{M} = \mathrm{B}[\alpha + 2, \beta + 1] + \gamma \mathrm{B}[\alpha + \frac{5}{2}, \beta + 1] + \eta \mathrm{B}[\alpha + 3, \beta + 1]$ normalizes the function to the second moment to maximally decorrelate the normalization and shape parameters (which would otherwise be correlated through the momentum sum rule Eq. (1.4d)).

To characterize the nucleon valence region and discriminate it from the sea components, we parametrize the light-quark and strange PDFs according to

$$\begin{aligned} u &= u_v + \bar{u}, \qquad & d &= d_v + \bar{d}, \\ \bar{u} &= S_1 + \bar{u}_0, \qquad & \bar{d} &= S_1 + \bar{d}_0, \\ s &= S_2 + s_0, \qquad & \bar{s} &= S_2 + \bar{s}_0, \end{aligned} \tag{2.41}$$

where the dependence on x and the scale μ_0^2 has been suppressed for convenience. The input quark distributions u_v, d_v, $\bar{u}_0$, $\bar{d}_0$, s_0, and $\bar{s}_0$, as well as the gluon distribution g, are parametrized individually as in Eq. (2.40). For the sea quark PDFs, the additional functions S_1 and S_2 are also parametrized via Eq. (2.40), and are designed to allow a more singular small-x behavior compared to the valence distributions by allowing the corresponding α parameter to more negative values. Letting $S_1 \neq S_2$ allows for different small-x behaviors for the light sea quarks and the strange quarks. This separation is designed so that the $\bar{u}_0$, $\bar{d}_0$, s_0, and $\bar{s}_0$ distributions account for the nonperturbative contributions to the sea, while the S_1 and S_2 distributions account for the perturbative contributions.

The normalization parameter N for the gluon distribution is fixed by the momentum sum rule, while the corresponding normalization parameters for u_v, d_v, and s_0

are fixed by the valence number sum rules (see Eq. (1.4c)). For the u_v, d_v, g, $\bar{u}_0$, and $\bar{d}_0$ distributions, the parameters γ and η in Eq. (2.40) are included to allow sufficient flexibility. We have verified that including these parameters for S_1, S_2, s_0, or $\bar{s}_0$ does not lead to significant changes to the final results, so for these distributions the γ and η parameters are set to zero. In total there are 33 fitted parameters for the PDFs, as well as 58 fitted normalization parameters. The parametrization of the single hadron FFs will be discussed in Chapter 4.

The scale dependence of the PDFs and FFs is determined according to the DGLAP evolution equations (Eq. (1.6)), with the PDFs, FFs, and α_s evolved according to their renormalization group equation (RGE) at next-to-leading logarithmic accuracy with the boundary condition $\alpha_s(M_Z) = 0.118$. For light and for heavy quarks, the PDFs and FFs are evolved using the zero-mass variable-flavor-number scheme. The values of the heavy quark mass thresholds for the evolution of the PDFs, FFs, and α_s are taken from the PDG values $m_c = 1.28$ GeV and $m_b = 4.18$ GeV in the $\overline{\text{MS}}$ scheme [11].

2.4 Mellin Space Techniques

By choosing parametrizations that can be analytically converted to Mellin space, one can use Mellin space techniques to greatly increase the efficiency of the analysis. These techniques are used to convert convolution integrals into matrix multiplications. This is relevant for the DGLAP evolution equation (see Eq. (1.6)) and for the calculation of hadron-hadron collision observables. In the latter case, it also allows one to isolate and pre-calculate parameter-independent integrals prior to the actual fitting procedure.

The N^{th} Mellin moment of a PDF $f_j(x, \mu_0)$ at the initial scale μ_0 is given by

$$\tilde{f}_j(N, \mu_0) = \int_0^1 \mathrm{d}x x^{N-1} f_j(x, \mu_0), \tag{2.42}$$

where N is also termed the "conjugate variable" to x. A similar expression holds for the FFs, with $x \to z$. With the PDFs converted to Mellin space, one can write the DGLAP evolution equation (Eq. (1.6)) as [117–121]

$$\frac{\partial \tilde{f}_i(x; \mu_R)}{\partial \ln \mu_R^2} = \sum_j \left[\tilde{P}_{ij} \otimes \tilde{f}_i \right] (x; \mu_R), \tag{2.43}$$

where $\tilde{P}_{ij}$ are the Mellin moments of the splitting kernels. Following the conventions of Ref. [121], we solve Eq. (2.43) by working in a flavor singlet and nonsinglet basis and defining:

$$\tilde{f}_{\pm 3} \equiv \tilde{f}_{u^\pm} - \tilde{f}_{d^\pm}, \tag{2.44a}$$

$$\tilde{f}_{\pm 8} \equiv \tilde{f}_{u^\pm} + \tilde{f}_{d^\pm} - 2\tilde{f}_{s^\pm}, \tag{2.44b}$$

$$\tilde{f}_{\pm 15} \equiv \tilde{f}_{u^\pm} + \tilde{f}_{d^\pm} + \tilde{f}_{s^\pm} - 3\tilde{f}_{c^\pm}, \tag{2.44c}$$

$$\tilde{f}_{\pm 24} \equiv \tilde{f}_{u^\pm} + \tilde{f}_{d^\pm} + \tilde{f}_{s^\pm} + \tilde{f}_{c^\pm} - 4\tilde{f}_{b^\pm}, \tag{2.44d}$$

$$\tilde{f}_{\pm 35} \equiv \tilde{f}_{u^\pm} + \tilde{f}_{d^\pm} + \tilde{f}_{s^\pm} + \tilde{f}_{c^\pm} + \tilde{f}_{b^\pm} - 5\tilde{f}_{t^\pm}, \tag{2.44e}$$

$$\tilde{f}_{\pm} \equiv \tilde{f}_{u^\pm} + \tilde{f}_{d^\pm} + \tilde{f}_{s^\pm} + \tilde{f}_{c^\pm} + \tilde{f}_{b^\pm} + \tilde{f}_{t^\pm}, \tag{2.44f}$$

where $\tilde{f}_{q^\pm} = \tilde{f}_q \pm \tilde{f}_{\bar{q}}$. To avoid the system of equations becoming degenerate, one only considers the equations up to $\tilde{f}_{\pm n_f^2 - 1}$, where n_f is the number of active flavors. Eq. (2.43) can then be expressed as

$$\frac{\partial \tilde{f}_{\pm i}(x;\mu_R)}{\partial \ln \mu_R^2} = \tilde{P}^{\pm}_{\rm NS} \tilde{f}_{\pm i}, \tag{2.45a}$$

$$\frac{\partial \tilde{f}_{-}(x;\mu_R)}{\partial \ln \mu_R^2} = \tilde{P}^{-}_{\rm NS} \tilde{f}_{-}, \tag{2.45b}$$

$$\frac{\partial}{\partial \ln \mu_R^2} \begin{pmatrix} \tilde{f}_+ \\ \tilde{f}_g \end{pmatrix} = \begin{pmatrix} \tilde{P}_{qq} & \tilde{P}_{qg} \\ \tilde{P}_{gq} & \tilde{P}_{gg} \end{pmatrix} \begin{pmatrix} \tilde{f}_+ . \\ \tilde{f}_g \end{pmatrix}. \tag{2.45c}$$

Explicit expressions for all of the splitting kernels in Mellin space can be found in Appendix B of Ref. [113]. The equations are easily computed in this basis, and the basis Eq. (2.44) can be inverted to get the evolved PDFs.

With the DGLAP evolution equations written in Mellin space, one can also write the observables as inverse Mellin transforms. Focusing on hadron-hadron collision processes, one can schematically write the cross-section as

$$\mathrm{d}\sigma = \sum_{ij} \mathcal{H}_{ij}(x_1, x_2, \mu_R) \otimes f_i(x_1, \mu_R) \otimes f_j(x_2, \mu_R), \tag{2.46}$$

where $\mathcal{H}_{ij}$ is the relevant partonic cross section for the process, x_1 and x_2 are the partonic momentum fractions of the two protons, and i and j run over the channels

that contribute to the process. The double convolution is given by

$$\big[C \otimes f \otimes g\big](x,z) \equiv \int_x^1 \frac{\mathrm{d}w}{w} \int_z^1 \frac{\mathrm{d}v}{v} C(w,v) f(\frac{x}{w}) g(\frac{z}{v}). \tag{2.47}$$

Since the analytical expressions of the Mellin moments of the partonic cross sections are not known, we follow the strategy of Ref. [106] and precalculate the Mellin moments numerically. One can then write the cross section as a 2D inverse Mellin transform with N being the conjugate variable to x_1 and M to x_2:

$$\begin{aligned}\mathrm{d}\sigma = \sum_{ijkl} \frac{1}{(2\pi i)^2} \int \mathrm{d}N \int \mathrm{d}M \tilde{f}_j(N,\mu_0) \tilde{f}_l(M,\mu_0) \\ \otimes \Big[x_1^{-N} x_2^{-M} \tilde{\mathcal{H}}_{ik}(N,M,\mu_R) U_{ij}^S(N,\mu_R,\mu_0) U_{kl}^S(M,\mu_R,\mu_0)\Big],\end{aligned} \tag{2.48}$$

where $\tilde{\mathcal{H}}_{ik}(N,M,\mu_R)$ are the moments of the partonic cross section. The solution of the DGLAP evolution is encoded in the matrix $U_{ij}^S(N,\mu_R,\mu_0)$ which evolves the moments of the PDFs $\tilde{f}_j(N,\mu_0)$ from μ_0 to μ_R. The superscript S denotes the space-like evolution of the PDFs, and the FFs would instead have a timelike (T) evolution. Note that, crucially, the quantity in brackets does not depend on any of the fitted parameters, and therefore only needs to be calculated once prior to the fit, massively reducing computation time. This strategy can also be applied to single convolutions such as in DIS, although the computational benefits are much smaller.

2.5 Data and Quality of Fit

The spin-averaged PDFs were extracted using F_2 data from fixed-target experiments on p, D, ^{3}He, and ^{3}H from BCDMS [122], NMC [123, 124], SLAC [125], and Jefferson Lab [126–128] including, in particular, the data from the MARATHON collaboration [129]. Reduced neutral and charged current proton cross sections from the combined H1/ZEUS analysis from HERA [130] are also included. The kinematic constraints $W^2 > 3.5$ GeV^2 and $Q^2 > m_c^2$ are applied to all of the DIS data. The cut on W^2 is placed to avoid the resonance region, and we find that reducing the cut any further significantly harms the description of the data. The cut on Q^2 is placed to match the input scale (see Section 2.3). In terms of hadron-hadron collision data, we include Drell-Yan di-muon data in pp and pD collisions from the Fermilab E866 [93, 94] and E906 experiments [95]. For weak vector boson mediated processes we use

reconstructed Z/γ^* cross sections and $W^\pm$ asymmetries from the Tevatron [131–134], as well as inclusive $W^\pm$-lepton asymmetries from CMS [135–138], LHCb [139,140], and STAR [96]. Finally, jet production data from the Tevatron [141,142] and STAR [143] are included with a cut on the jet transverse momentum of $p_T^{\rm jet} > 8$ GeV.

For semi-inclusive data, we include data on unpolarized SIDIS on deuterium targets from COMPASS [144,145] with the cuts $Q^2 > m_c^2$ and $W^2 > 10$ GeV^2, the latter allowing one to neglect higher twist, nuclear, and target mass corrections, which would be difficult to account for simultaneously with final-state hadron mass corrections [146]. An additional cut is placed on the fragmentation variable $0.2 < z_H < 0.8$ to ensure the applicability of the leading power formalism and avoid large-z threshold corrections [116]. Furthermore, SIA data is included from the TASSO [147–151] and ARGUS [152] collaborations at DESY; the TPC [153–155], HRS [156], SLD [157], and BABAR [158] collaborations at SLAC; the OPAL [159,160], ALEPH [161], and DELPHI [162] collaborations at CERN; and the TOPAZ [163] and Belle [164] at KEK. The SIA data are restricted to $0.2 < z_h < 0.9$. The kinematics (at LO in α_s) are shown for all experiments in Fig. 2.7.

The $\chi^2_{\rm red}$ values and fitted normalizations for the unpolarized DIS data are shown in Table 2.2, while Table 2.3 shows the values for the unpolarized hadron-hadron collision data. We note that when the experimental observable contains a ratio of cross sections, the normalization uncertainty may be zero or negligible. Thus the uncertainty is not provided and we do not include a fitted normalization parameter. Figs. (2.8)–(2.17) show the data and theory comparison for all of the inclusive data included in this analysis. The DIS data is shown in Fig. 2.8 for the neutral current proton data, Fig. 2.9 for the neutral current deuteron data, Fig. 2.10 for the neutral current $A = 3$ data, and Fig. 2.11 for the charged current proton data. Note that in Fig. 2.8 and Fig. 2.9 the scattering angle is given by $\theta = 2\sin^{-1}\sqrt{Q^2/(4E^2 - \frac{2EQ^2}{Mx_{\rm bj}})}$, where E is the energy in the lab frame ($E = 5.5$ GeV for Jefferson Lab Hall C) and M is the nucleon mass.

The DY process is presented in Fig. 2.12, where the ratio of pD to pp cross sections is shown, and in Fig. 2.13, where the absolute pp cross section is shown. For reconstructed W and Z boson production data, the comparison is shown in Fig. 2.14. The comparison for $W^\pm$-lepton asymmetries is shown in Fig. 2.15, while Fig. 2.16 shows the W^+/W^- ratio. Fig. 2.17 shows the comparison for the jet production data.

type	experiment	ref.	beam	target	observable	$N_{\rm dat}$	$\chi^2_{\rm red}$	fitted norm.
fixed target	BCDMS	[122]	μ	p	F_2	351	1.20	0.995(27)
	BCDMS	[122]	μ	D	F_2	254	1.08	1.015(30)
	NMC	[123]	μ	p	F_2	275	1.67	1.018(19)
	NMC	[124]	μ	D/p	F_2	189	0.89	0.998(11)
	SLAC	[125]	e	p	F_2	530	0.78	1.032(20)
	SLAC	[125]	e	D	F_2	541	0.84	1.030(18)
	JLab Hall C	[126]	e	p	F_2	92	1.19	1.037(18)
	JLab Hall C	[126]	e	D	F_2	91	0.87	1.003(18)
	JLab Hall C	[128]	e	^{3}He/D	F_2	13	0.31	1.011(18)
	MARATHON	[129]	e	D/p	F_2	7	1.06	1.018(6)
	MARATHON	[129]	e	^{3}He/^{3}H	F_2	21	0.67	1.006(12)
tagged	JLab BONuS	[127]	e	n/D	F_2	137	1.24	1.007(44)
collider	HERA (NC e^-)	[130]	e	p	$\sigma_{\rm red}$	159	1.68	—
	HERA (NC e^+)	[130]	e	p	$\sigma_{\rm red}$	945	1.41	—
	HERA (CC e^-)	[130]	e	p	$\sigma_{\rm red}$	42	1.31	—
	HERA (CC e^+)	[130]	e	p	$\sigma_{\rm red}$	39	1.12	—
Total						3686	**1.15**	

Table 2.2: χ^2 table: Unpolarized DIS. Summary of the $\chi^2_{\rm red}$ values for the inclusive DIS data used in this analysis, as well as their fitted normalizations.

For the latest W-lepton data from STAR, shown in Fig. 2.16, the description suffers slightly at high and at low pseudorapidities, leading to a $\chi^2_{\rm red}$ of 1.99 for these data. From Ref. [96], it is known that this is a common feature of most PDF extractions. This discrepancy is also partially due to some tension with the NuSea data, and the $\chi^2_{\rm red}$ improves to 1.54 upon its removal.

The $\chi^2_{\rm red}$ values and fitted normalizations for the unpolarized SIDIS data are shown in Table 2.4, while Table 2.5, Table 2.6, and Table 2.7 show the values for SIA pion data, kaon data, and unidentified hadron data respectively. The data and theory comparisons for SIDIS are shown in Figs. (2.18)–(2.20), while for SIA they are shown in Figs. (2.21)–(2.23).

type	experiment	ref.	process	variables	$\sqrt{s}$ (GeV)	observable	$N_{\rm dat}$	$\chi^2_{\rm red}$	fitted norm.
Drell-Yan	E866	[92]	$pp \to \gamma^* \to \mu^+\mu^-$	$M_{\ell\ell}, x_F$	39	$M^3_{\ell\ell}\mathrm{d}\sigma^{pp}$	184	1.12	1.119(75)
	E866	[94]	$pp, pD \to \gamma^* \to \mu^+\mu^-$	$M_{\ell\ell}, x_F$	39	$\mathrm{d}\sigma^{pD}/2\mathrm{d}\sigma^{pp}$	15	1.06	0.996(9)
	E906	[95]	$pp, pD \to \gamma^* \to \mu^+\mu^-$	$M_{\ell\ell}, x_F$	15	$\mathrm{d}\sigma^{pD}/2\mathrm{d}\sigma^{pp}$	6	0.96	0.998(11)
Z rapidity	CDF	[131]	$p\bar{p} \to Z/\gamma^*$	y_Z	1960	$\mathrm{d}\sigma$	28	1.16	0.994(60)
	D0	[132]	$p\bar{p} \to Z/\gamma^*$	y_Z	1960	$\mathrm{d}\sigma/\sigma$	28	0.94	—
W asymmetry	CDF	[133]	$p\bar{p} \to W$	y_W	1960	A_W	13	1.50	—
	D0	[134]	$p\bar{p} \to W$	y_W	1960	A_W	14	0.77	—
lepton rapidity	CMS	[135]	$pp \to W \to e\nu_e$	η_e	7000	A_e	6	0.80	—
	CMS	[135]	$pp \to W \to \mu\nu_\mu$	η_μ	7000	A_μ	6	0.12	—
	CMS	[136]	$pp \to W \to e\nu_e$	η_e	7000	A_e	11	0.94	—
	CMS	[137]	$pp \to W \to \mu\nu_\mu$	η_μ	7000	A_μ	11	0.35	—
	CMS	[138]	$pp \to W \to \mu\nu_\mu$	η_μ	8000	A_μ	11	1.42	—
	LHCb	[139]	$pp \to W \to \mu\nu_\mu$	η_μ	7000	A_μ	8	0.57	—
	LHCb	[140]	$pp \to W \to \mu\nu_\mu$	η_μ	8000	A_μ	8	0.30	—
	STAR	[96]	$pp \to W \to e\nu_e$	η_e	510	$\mathrm{d}\sigma^{W^+}/\mathrm{d}\sigma^{W^-}$	9	1.99	—
inclusive jets	D0	[141]	$p\bar{p} \to \mathrm{jet}$	$y_{\rm jet}, p_T^{\rm jet}$	1960	$\mathrm{d}\sigma$	110	1.04	0.765(87)
	CDF	[142]	$p\bar{p} \to \mathrm{jet}$	$y_{\rm jet}, p_T^{\rm jet}$	1960	$\mathrm{d}\sigma$	76	1.87	0.856(76)
	STAR MB	[143]	$pp \to \mathrm{jet}$	$y_{\rm jet}, p_T^{\rm jet}$	510	$\mathrm{d}\sigma/2\pi$	3	0.08	0.986(42)
	STAR HT	[143]	$pp \to \mathrm{jet}$	$y_{\rm jet}, p_T^{\rm jet}$	510	$\mathrm{d}\sigma/2\pi$	9	1.14	0.975(98)
Total							556	**1.19**	

Table 2.3: χ^2 table: Unpolarized pp collisions. Summary of the $\chi^2_{\rm red}$ values for the hadron-hadron collision data used in this analysis, as well as their fitted normalizations. All processes are inclusive and the undetected part of the final system X has been suppressed in the "process" column.

experiment	ref.	beam	target	hadron	$N_{\rm dat}$	$\chi^2_{\rm red}$	fitted norm.
COMPASS	[144]	μ	D	π^+	249	1.00	1.310(48)
COMPASS	[144]	μ	D	π^-	249	0.88	1.271(52)
COMPASS	[145]	μ	D	K^+	247	1.34	1.302(136)
COMPASS	[145]	μ	D	$K-$	247	1.29	1.205(109)
COMPASS	[144]	μ	D	$h+$	249	0.96	1.285(65)
COMPASS	[144]	μ	D	h^-	249	0.45	1.266(60)
Total					1490	**0.98**	

Table 2.4: χ^2 table: SIDIS. Summary of the $\chi^2_{\rm red}$ values for the SIDIS data on the multiplicity used in this analysis, as well as their fitted normalizations.

experiment	ref.	$\sqrt{s}$ (GeV)	N_{dat}	χ^2_{red}	fitted norm.
TASSO	[147]	12	1	1.17	1.111(73)
TASSO	[148]	14	4	3.05	0.988(58)
TASSO	[148]	30	4	0.82	1.024(32)
TASSO	[151]	34	5	1.13	1.026(49)
TASSO	[151]	44	3	1.58	0.992(39)
TPC	[154]	29	3	1.09	—
TPC	[155]	29	9	0.62	—
TPC(c)	[153]	29	3	0.51	—
TPC(b)	[153]	29	3	2.49	—
TOPAZ	[163]	58	1	0.09	—
SLD	[157]	91	15	1.54	0.999(9)
SLD(c)	[157]	91	15	1.24	1.009(3)
SLD(b)	[157]	91	15	0.75	1.000(3)
ALEPH	[161]	91	5	0.75	0.998(26)
OPAL	[159]	91	6	2.88	—
OPAL(c)	[160]	91	5	2.13	1.314(53)
OPAL(b)	[160]	91	5	2.03	1.276(72)
DELPHI	[162]	91	6	2.55	—
DELPHI(b)	[162]	91	6	1.09	—
BABAR	[158]	11	27	0.20	0.979(8)
BELLE	[164]	11	70	0.09	0.876(12)
ARGUS	[152]	10	20	1.80	1.012(11)
Total			231	**0.91**	

Table 2.5: χ^2 table: SIA pion. Summary of the χ^2_{red} values for the pion SIA data used in this analysis, as well as their fitted normalizations.

experiment	ref.	$\sqrt{s}$ (GeV)	$N_{\rm dat}$	$\chi^2_{\rm red}$	fitted norm.
TASSO	[148]	14	2	1.95	0.969(26)
TASSO	[148]	22	2	0.07	1.001(15)
TASSO	[151]	34	2	0.02	0.994(22)
TPC	[154]	29	2	4.46	—
TPC	[155]	29	7	0.60	—
TOPAZ	[163]	58	1	0.25	—
SLD	[157]	91	15	0.52	1.016(7)
SLD(c)	[157]	91	15	2.14	1.019(6)
SLD(b)	[157]	91	15	1.16	0.992(2)
ALEPH	[161]	91	5	0.09	1.003(23)
OPAL	[159]	91	6	0.41	—
OPAL(c)	[160]	91	5	3.41	1.310(44)
OPAL(b)	[160]	91	5	1.51	1.297(61)
DELPHI	[162]	91	6	1.76	—
DELPHI	[162]	91	6	0.40	—
DELPHI(b)	[162]	91	6	0.08	—
BABAR	[158]	11	28	0.48	0.967(17)
BELLE	[164]	11	70	0.18	0.963(24)
ARGUS	[152]	10	15	0.53	1.004(6)
Total			213	**0.70**	

Table 2.6: χ^2 table: SIA kaon. Summary of the $\chi^2_{\rm red}$ values for the kaon SIA data used in this analysis, as well as their fitted normalizations.

experiment	ref.	$\sqrt{s}$ (GeV)	$N_{\rm dat}$	$\chi^2_{\rm red}$	fitted norm.
TASSO	[151]	35	8	3.34	0.922(39)
TASSO	[151]	44	8	0.80	0.953(39)
TPC	[154]	29	10	1.83	—
SLD	[157]	91	15	1.89	1.002(10)
SLD(c)	[157]	91	15	0.99	1.007(6)
SLD(b)	[157]	91	15	0.86	1.010(5)
ALEPH	[161]	91	19	0.33	0.998(10)
OPAL	[159]	91	6	0.44	—
OPAL(c)	[160]	91	6	0.49	—
OPAL(b)	[160]	91	6	0.10	—
DELPHI	[162]	91	6	0.73	—
DELPHI(b)	[162]	91	6	0.67	—
Total			120	**1.07**	

Table 2.7: χ^2 table: SIA unidentified hadron. Summary of the $\chi^2_{\rm red}$ values for the unidentified hadron SIA data used in this analysis, as well as their fitted normalizations.

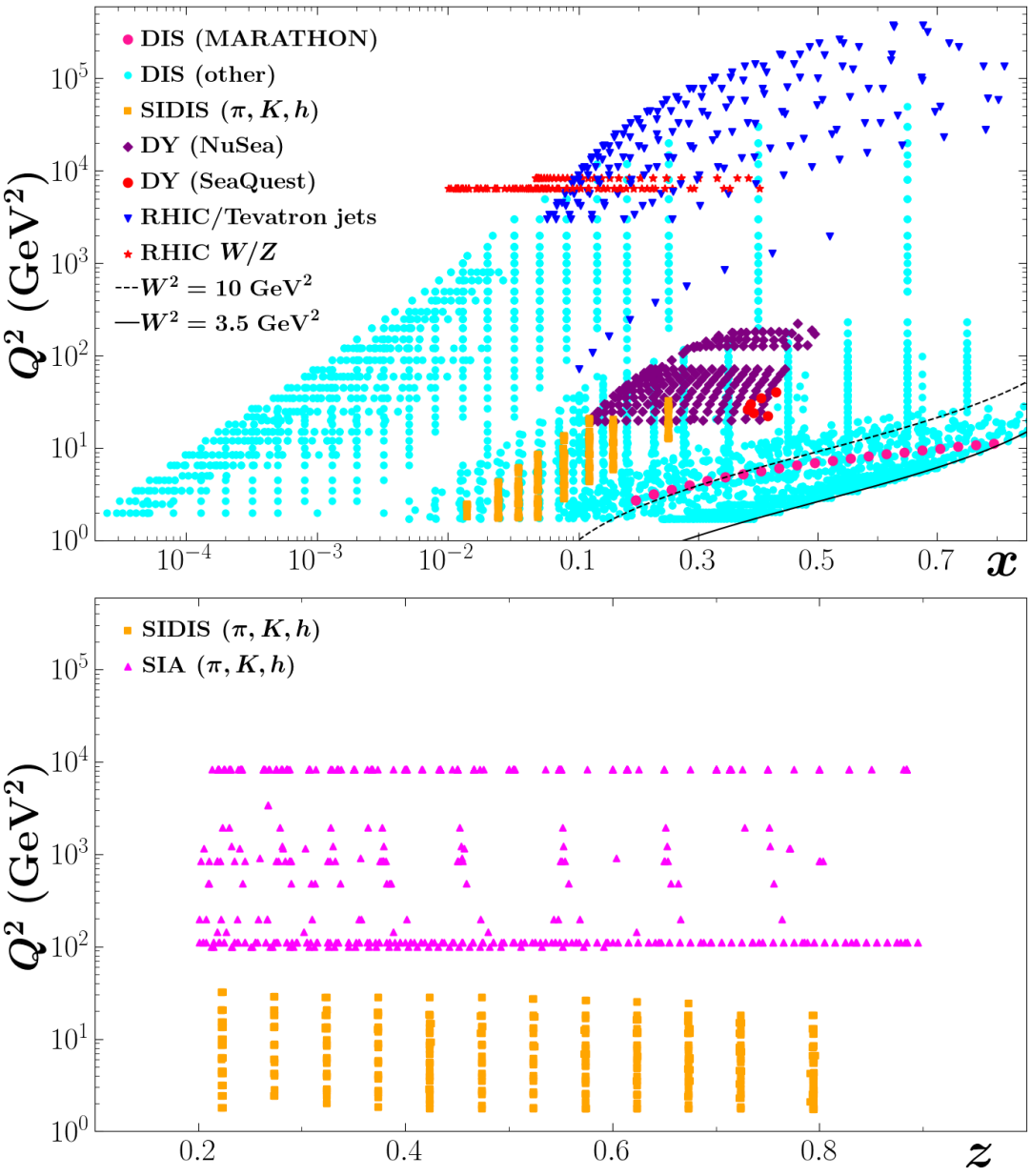

Figure 2.7: Kinematics of unpolarized datasets. Kinematic coverage of the datasets included in this analysis. The top panel shows the data as a function of x and Q^2. The variable x represents $x_{\rm bj}$ (Eq. (2.7)) for DIS and SIDIS and x_F (Eq. (2.24)) for vector boson and jet production, while the scale Q^2 represents the four-momentum transfer squared for DIS and SIDIS, the mass squared of the intermediate boson for vector boson production, and the transverse momentum squared for jet production. For the DIS data we isolate and show the MARATHON [126] (pink circles). For Drell-Yan we split the data into those from NuSea [93, 94] and SeaQuest [95]. Also indicated are the DIS cuts of $W^2 = 10$ GeV^2 (dashed black line) and $W^2 = 3.5$ GeV^2 (solid black line). The bottom panel shows the semi-inclusive data as a function of z and Q^2. The variable z represents z_H (Eq. (2.33)) for SIDIS, and z_h (Eq. (2.36)) for SIA.

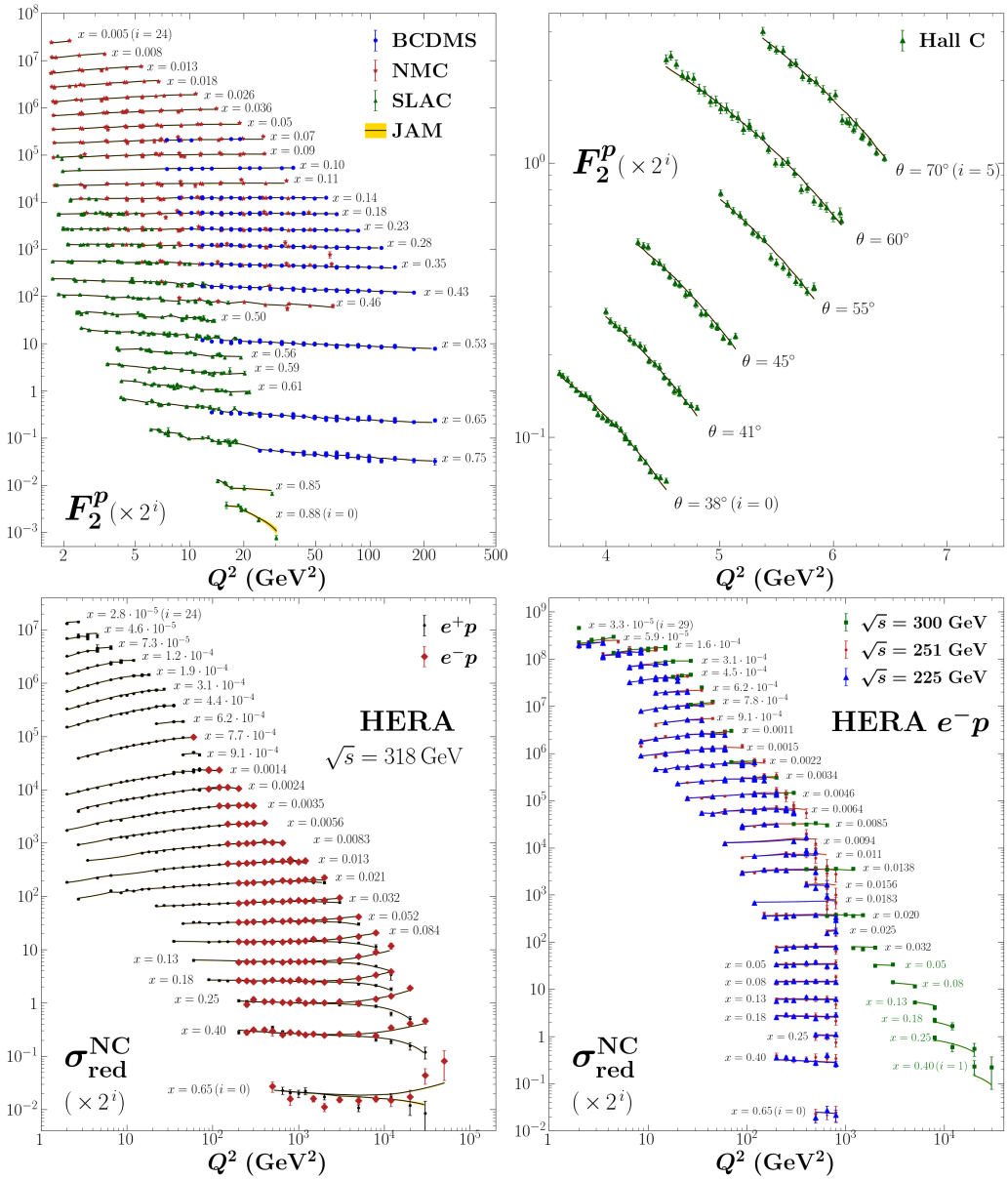

Figure 2.8: Data vs. Theory: Unpolarized DIS (proton). All proton DIS data included in this analysis, plotted as a function of Q^2 against the mean JAM result (black lines) with 1σ uncertainty bands in gold. Results are scaled by factors of 2^i for clarity. The top row shows F_2^p data from BCDMS [122], NMC [123], SLAC [125], and Hall C [126] plotted for different ranges of $x_{\rm bj}$ (left) or θ (right). The bottom row shows neutral current cross-section data from HERA [130] at $\sqrt{s} = 318, 300, 251$, and 225 GeV plotted for different ranges of $x_{\rm bj}$. For readability in this figure, $x_{\rm bj}$ has been shortened to x.

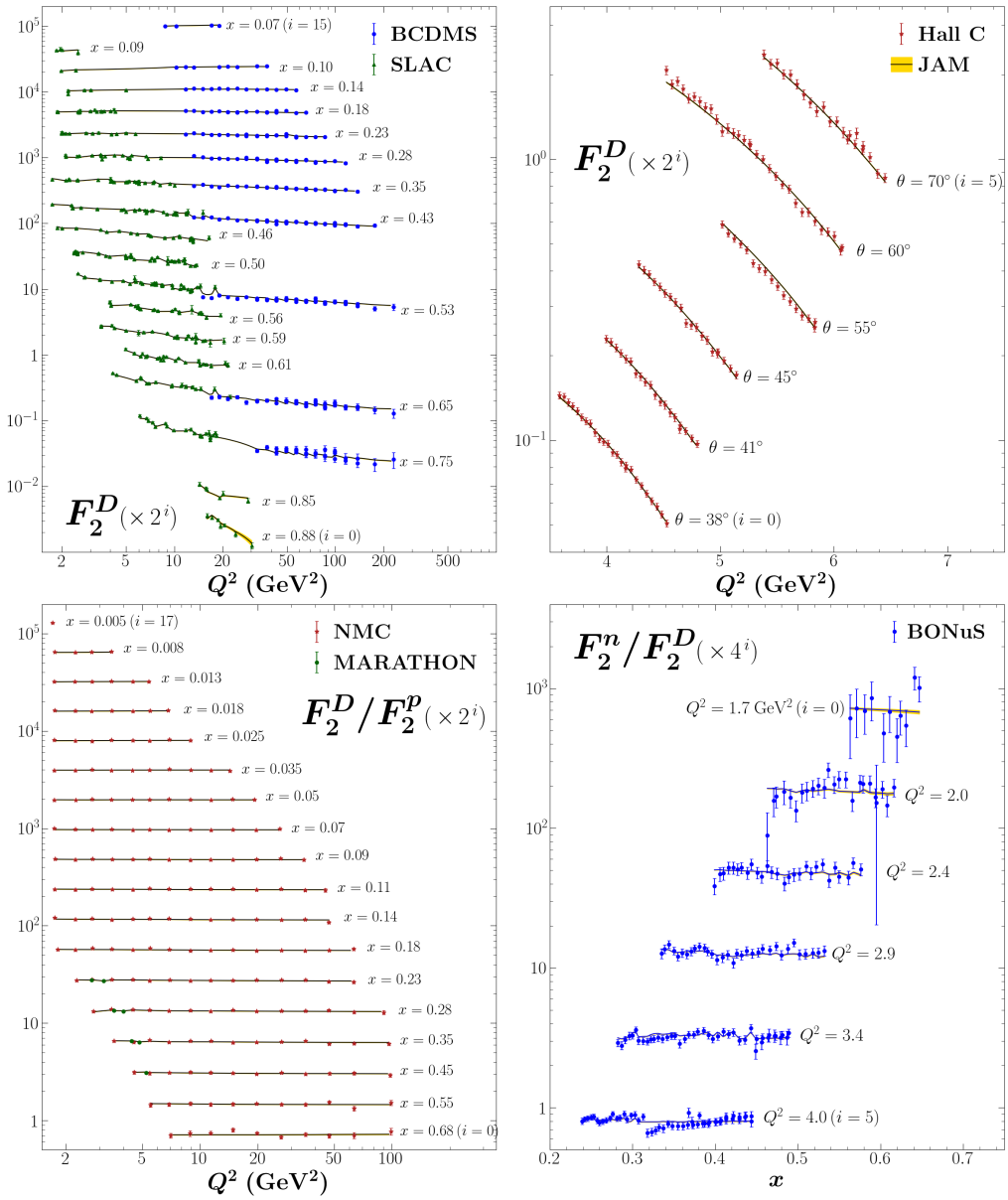

Figure 2.9: Data vs. Theory: Unpolarized DIS (deuteron). All deuteron DIS data included in this analysis, plotted against the mean JAM result (black lines) with 1σ uncertainty bands in gold. Results are scaled by factors of 2^i or 4^i for clarity. The top row shows F_2^D data from BCDMS [122], SLAC [125], and Hall C [126] plotted as a function of Q^2 for different ranges of $x_{\rm bj}$ (left) or θ (right). The bottom left panel shows F_2^D/F_2^p data from NMC [124] and MARATHON [129] plotted as a function of Q^2 for different ranges of $x_{\rm bj}$. The bottom right panel shows F_2^n/F_2^D data from BONuS [127] plotted as a function of $x_{\rm bj}$ with for different ranges of Q^2. For readability in this figure, $x_{\rm bj}$ has been shortened to x.

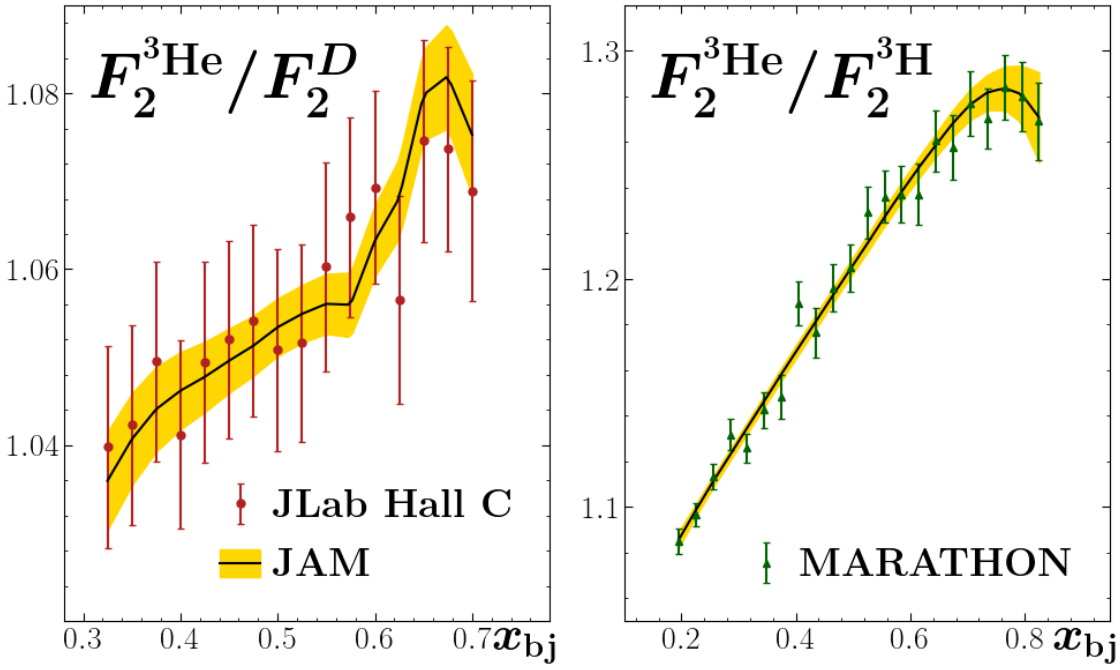

Figure 2.10: Data vs. Theory: Unpolarized DIS (helium and tritium). All $A = 3$ DIS data included in this analysis, plotted as a function of x_{bj} against the mean JAM result (black lines) with 1σ uncertainty bands in gold. The left panel shows $F_2^{^3\mathrm{He}}/F_2^D$ data from Hall C [128], while the right panel shows $F_2^{^3\mathrm{He}}/F_2^{^3\mathrm{H}}$ data from MARATHON [129].

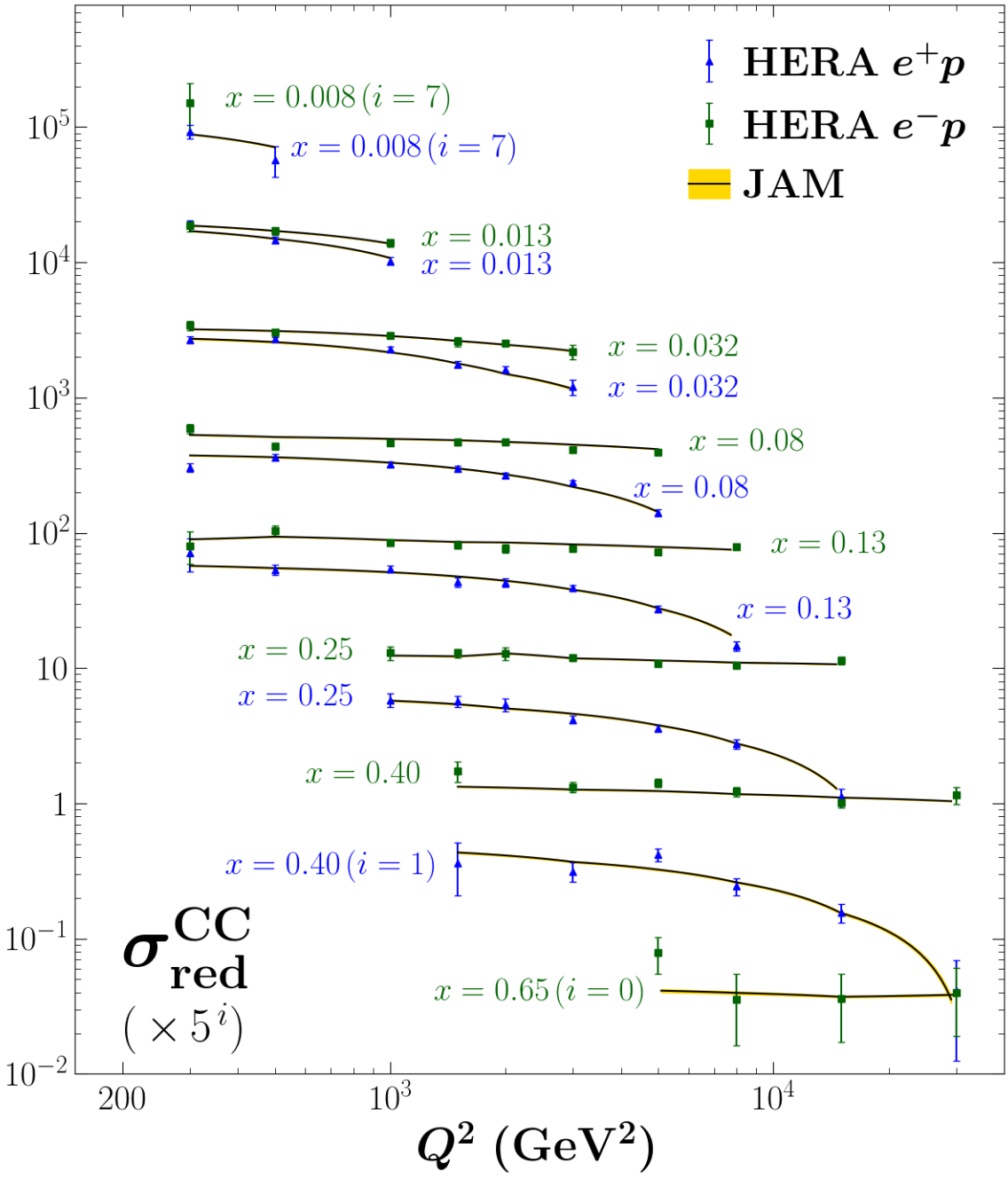

Figure 2.11: Data vs. Theory: Unpolarized DIS (charged current). All charged-current DIS data included in this analysis, plotted as a function of Q^2 for different ranges of x_{bj} against the mean JAM result (black lines) with 1σ uncertainty bands in gold. Results are scaled by factors of 5^i for clarity. For readability in this figure, x_{bj} has been shortened to x.

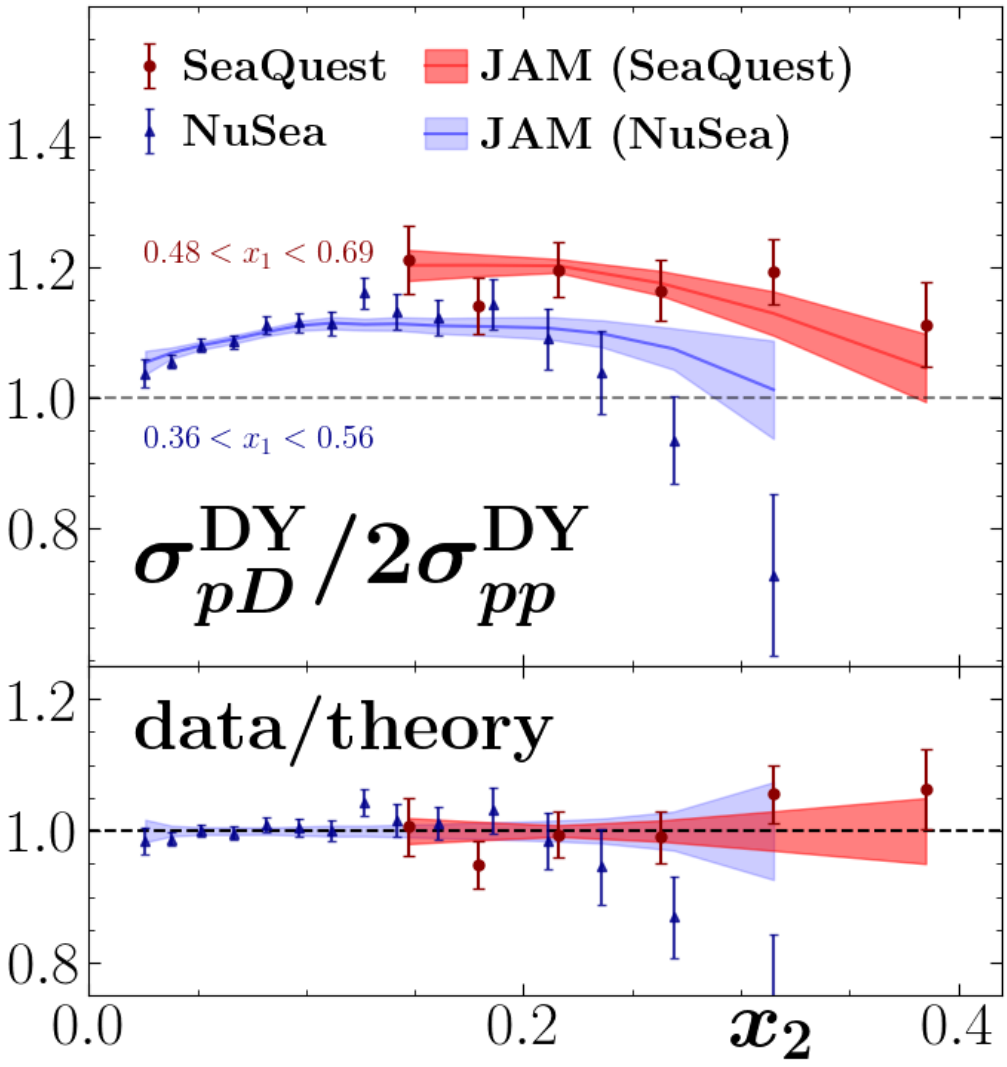

Figure 2.12: Data vs. Theory: Unpolarized Drell-Yan (ratios). Drell-Yan cross section ratio $\sigma_{pD}^{\rm DY}/2\sigma_{pp}^{\rm DY}$ (differential in $M_{\ell\ell}$ and x_F) from SeaQuest [95] (red circles) and NuSea [94] (blue triangles) compared with the JAM results at their respective kinematics (red and blue 1σ uncertainty bands), as a function of the target momentum fraction x_2, with the corresponding x_1 ranges indicated. The ratio of data to the average theory is illustrated in the lower panel with 1σ theoretical uncertainties at the SeaQuest and NuSea kinematics.

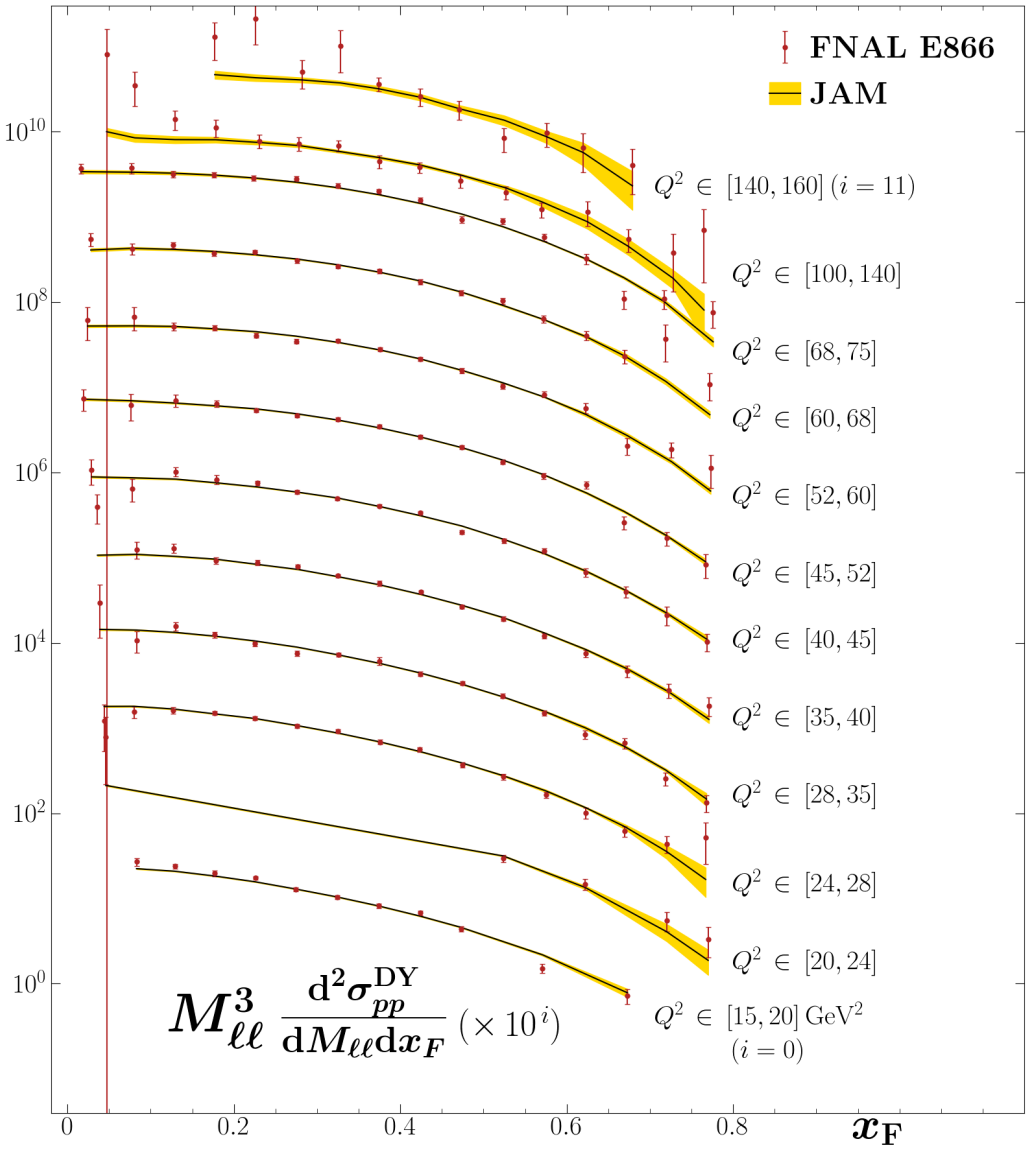

Figure 2.13: Data vs. Theory: Unpolarized Drell-Yan (cross section). Drell-Yan cross section from NuSea [94] (red circles) as a function of x_F compared with the mean JAM result (black lines) with 1σ uncertainty bands in gold. The Q^2 bins are indicated and results are scaled by factors of 10^i for clarity. The highest Q^2 bin $160 < Q^2 < 280$ GeV^2 has not been plotted. Note that the $Q^2 \in [75, 100]$ bin is absent in the experimental data to avoid the Υ resonance which lies at $Q^2 \approx 89$ GeV^2.

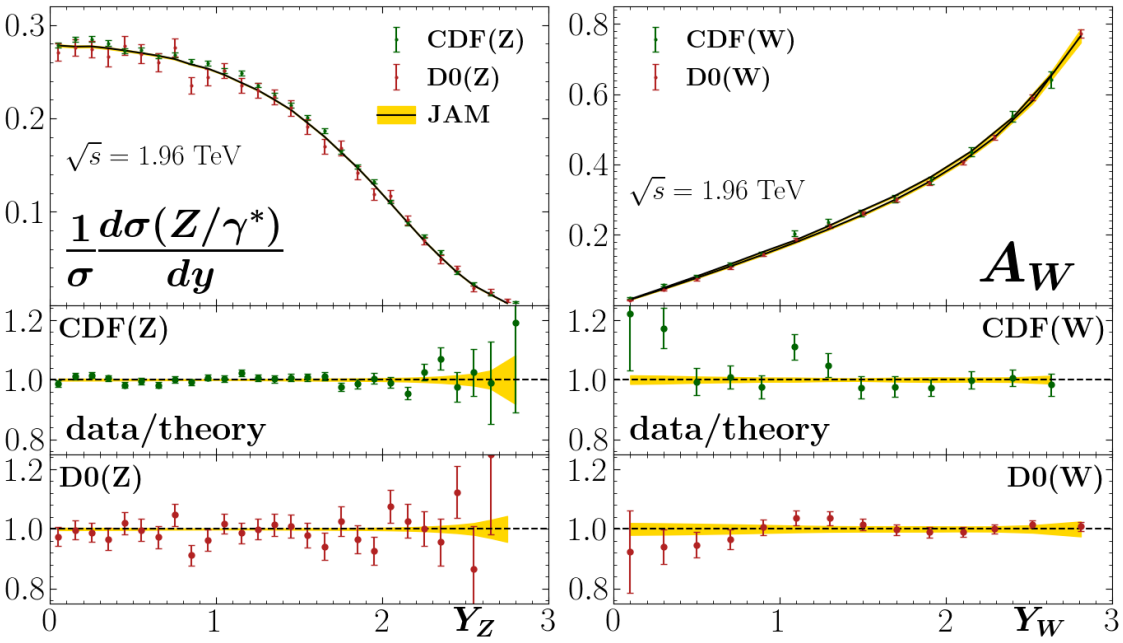

Figure 2.14: Data vs. Theory: Unpolarized W/Z boson production. Reconstructed Z boson normalized cross-sections and reconstructed W boson asymmetry data from Tevatron [131–134], plotted as a function of y_Z and y_W compared with the mean JAM result (black lines) with 1σ uncertainty bands in gold. The ratio of data to the average theory is illustrated in the lower panel with 1σ theoretical uncertainties (gold bands).

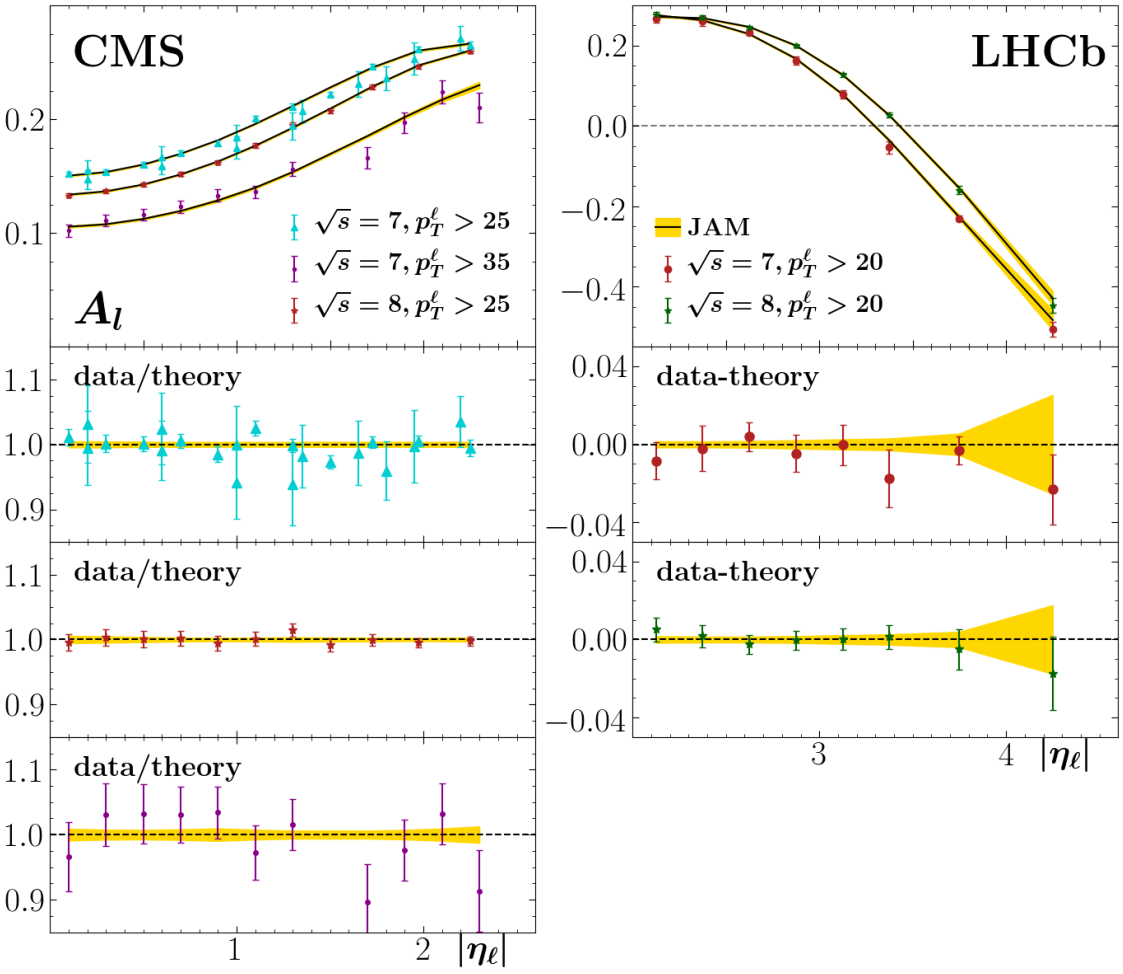

Figure 2.15: Data vs. Theory: Unpolarized W-lepton asymmetries. Inclusive $W^{\pm}$-lepton asymmetries from CMS [135–138] and LHCb [139, 140] plotted as a function of $|\eta_\ell|$ compared to the mean JAM result (black lines) with 1σ uncertainty bands in gold. The ratio of data to the average theory is illustrated in the lower panel with 1σ theoretical uncertainties (gold bands).

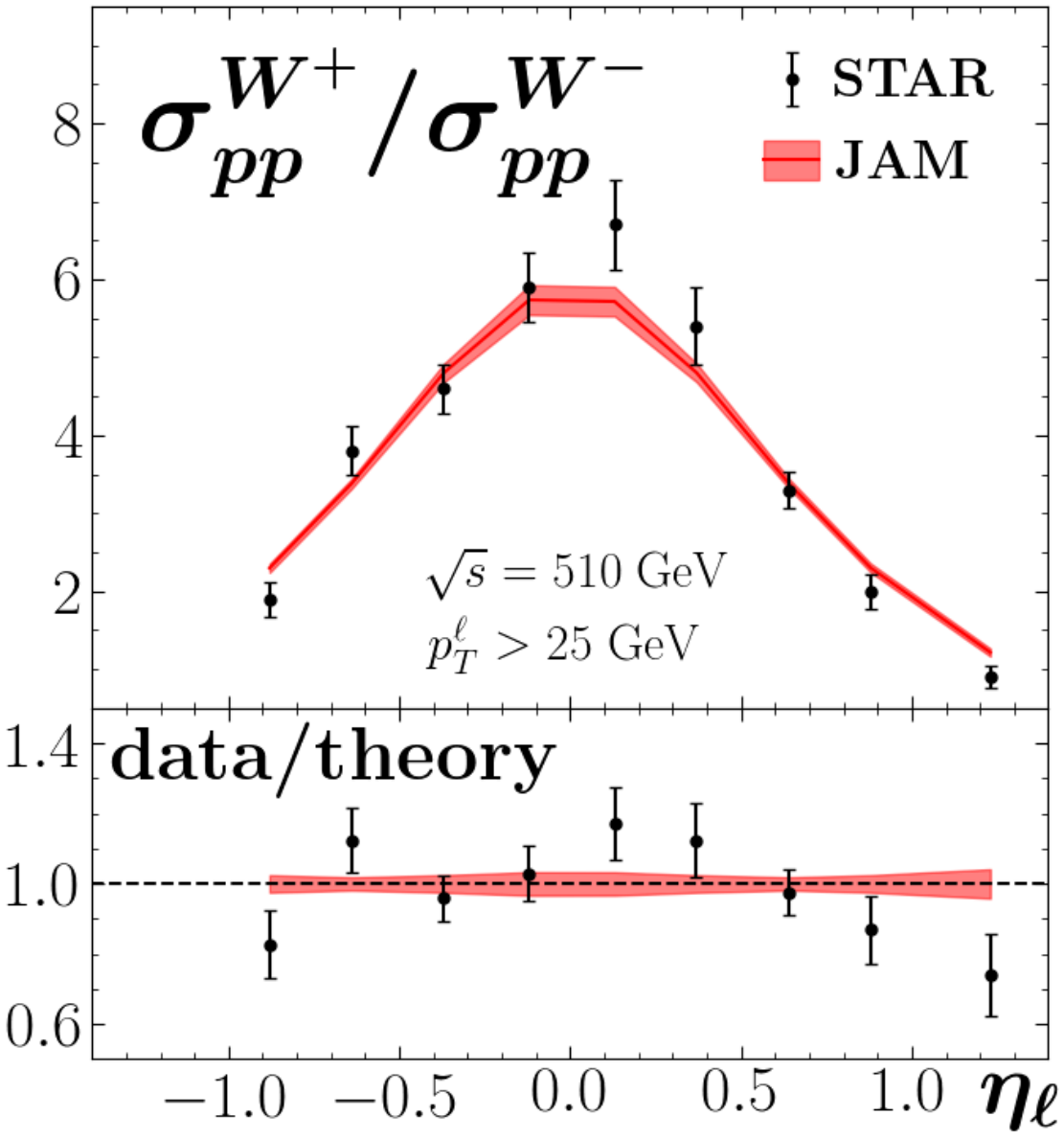

Figure 2.16: Data vs. Theory: Unpolarized W-lepton ratio. W-lepton cross section ratio $\sigma_{pp}^{W^+}/\sigma_{pp}^{W^-}$ from STAR [96] (black circles) compared with the JAM fit (red 1σ uncertainty band), as a function of the lepton pseudorapidity, η_ℓ. The ratio of data to average theory is shown in the lower panel together with the 1σ theoretical uncertainty (red band).

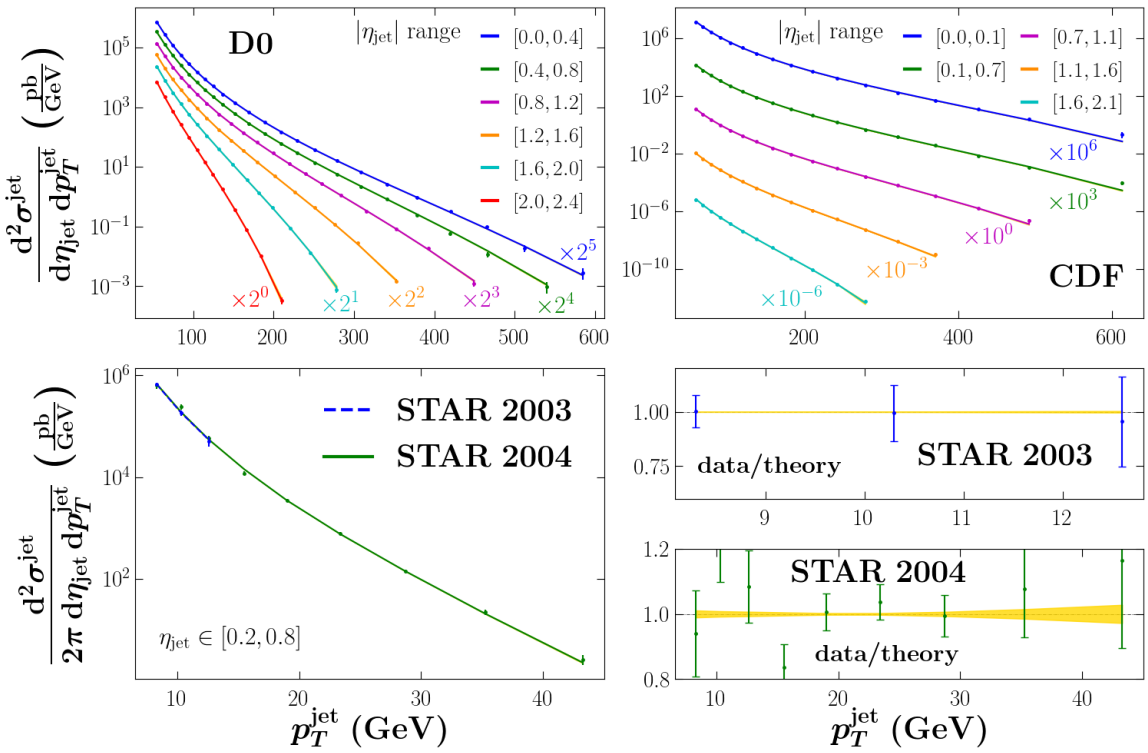

Figure 2.17: Data vs. Theory: Unpolarized jet production. Single jet production data in $p\bar{p}$ collisions from D0 [141] and CDF [142] and in pp collisions from STAR [143] plotted as a function of $p_T^{\rm jet}$ with approximately constant $\eta_{\rm jet}$ compared with the mean JAM fit (solid lines) with 1σ uncertainty bands in gold. The results are scaled by factors of 2^i or 10^i for clarity. The ratio of data to average theory for the STAR data is shown in the lower right panel together with the 1σ theoretical uncertainties (gold bands).

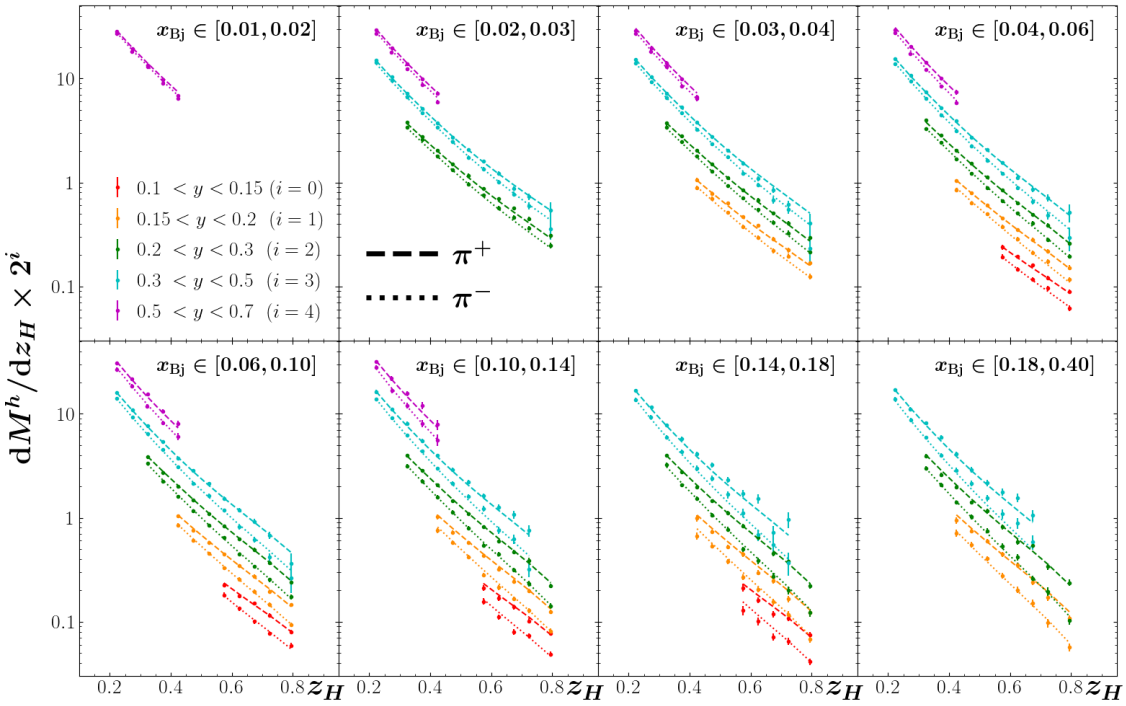

Figure 2.18: Data vs. Theory: Unpolarized SIDIS (pion). Pion multiplicities $\mathrm{d}M_h/\mathrm{d}z_H$ in SIDIS from COMPASS [144] plotted as a function of z_H for different ranges of $x_{\rm bj}$ and y against the mean JAM result. Each panel shows a different bin of $x_{\rm bj}$, while the colors indicate different bins of y. The solid JAM curves correspond to the π^+ multiplicities, while the dashed curves correspond to π^-. Results are scaled by a factor of 2^i for clarity.

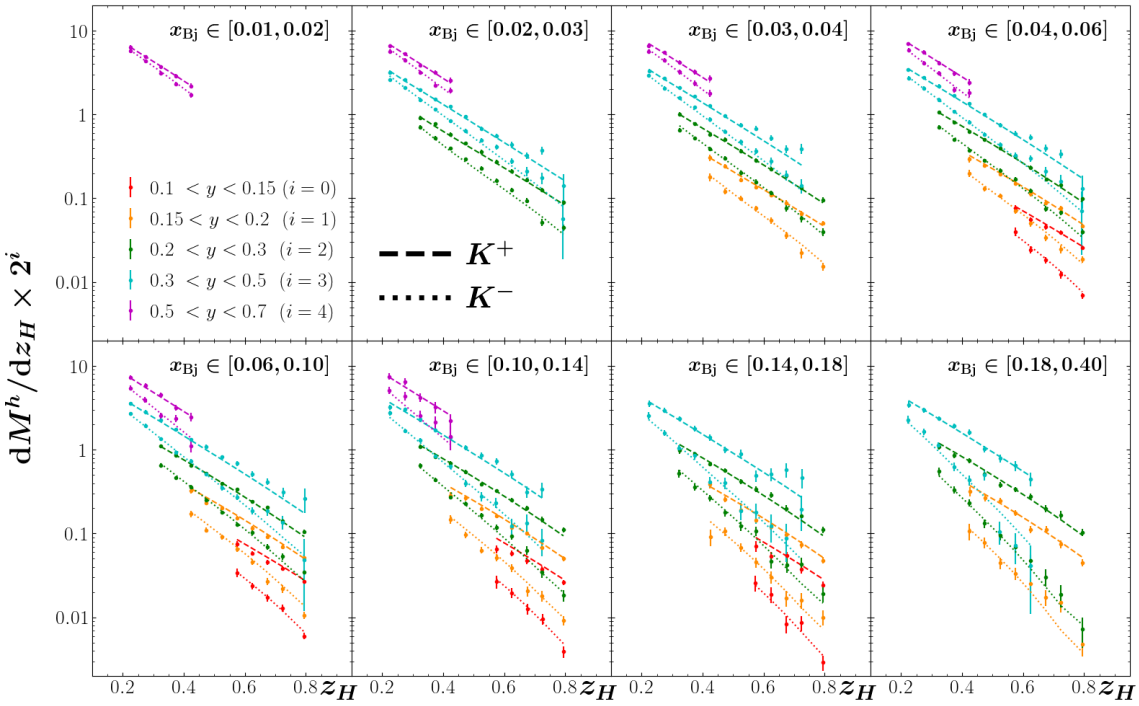

Figure 2.19: Data vs. Theory: Unpolarized SIDIS (kaon). Kaon multiplicities $\mathrm{d}M_h/\mathrm{d}z_H$ in SIDIS from COMPASS [145] plotted as a function of z_H for different ranges of $x_{\rm bj}$ and y against the mean JAM result. Each panel shows a different bin of $x_{\rm bj}$, while the colors indicate different bins of y. The solid JAM curves correspond to the K^+ multiplicities, while the dashed curves correspond to K^-. Results are scaled by a factor of 2^i for clarity.

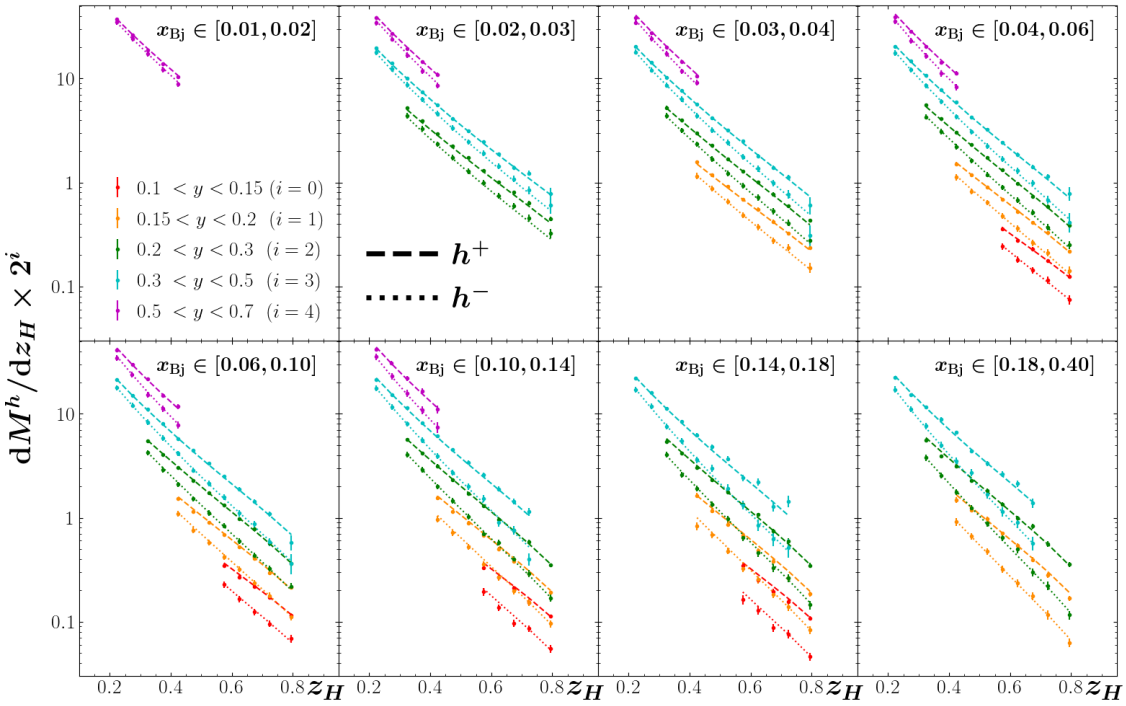

Figure 2.20: Data vs. Theory: Unpolarized SIDIS (hadron). Unidentified charged hadron multiplicities $\mathrm{d}M_h/\mathrm{d}z_H$ in SIDIS from COMPASS [144] plotted as a function of z_H for different ranges of $x_{\rm bj}$ and y against the mean JAM result. Each panel shows a different bin of $x_{\rm bj}$, while the colors indicate different bins of y. The solid JAM curves correspond to the h^+ multiplicities, while the dashed curves correspond to h^-. Results are scaled by a factor of 2^i for clarity.

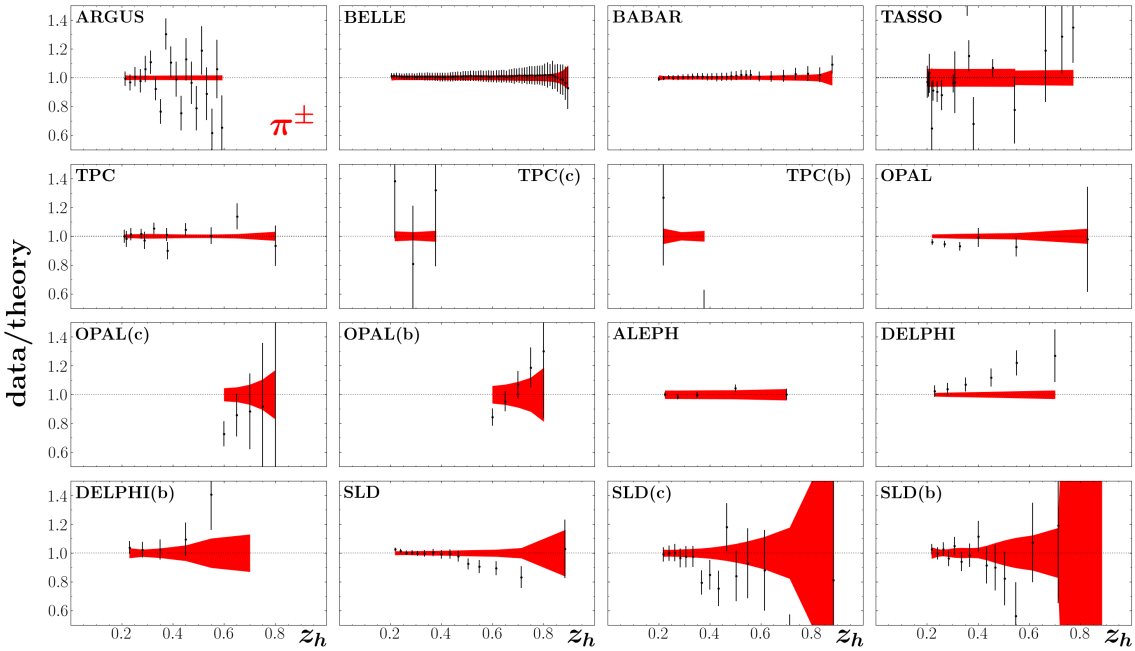

Figure 2.21: Data vs. Theory: SIA (pion). Data to theory ratios for SIA pion production cross sections as a function of z_h. The data points over theory from ARGUS [152], BELLE [164], BABAR [158], TASSO [147,148,151], TPC [153–155], OPAL [159, 160], ALEPH [161], DELPHI [162], and SLD [157] are shown in black, while the 1σ uncertainty band of the fitted result is shown in red. The labels (c) and (b) indicate that the data are for charm and bottom quark production respectively.

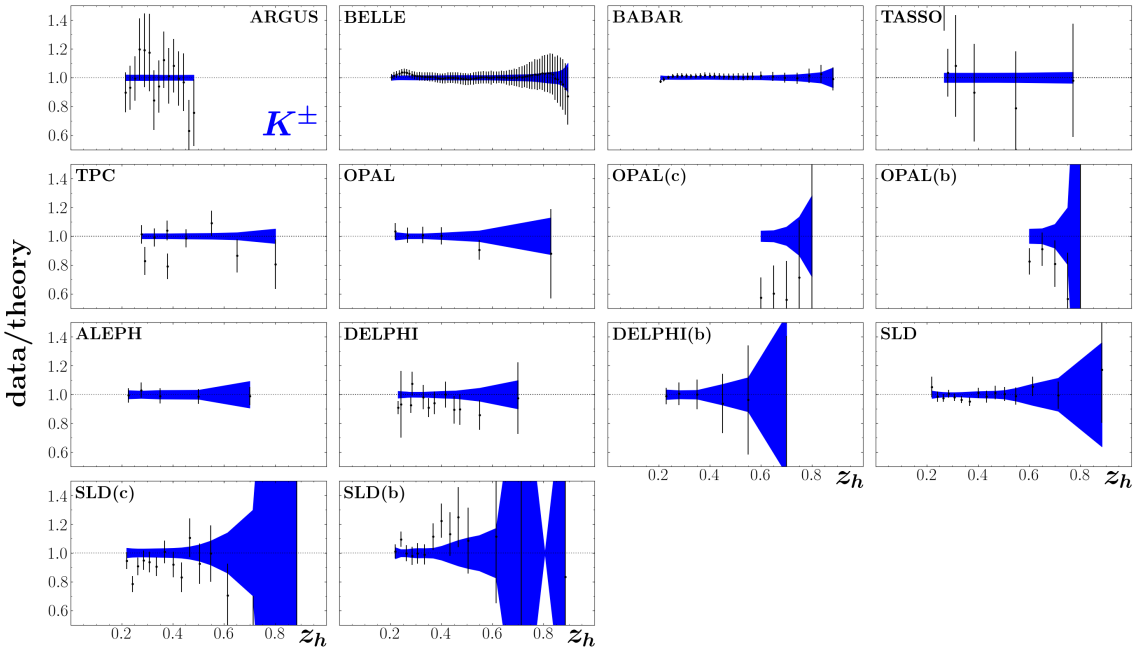

Figure 2.22: Data vs. Theory: SIA (kaon). Data to theory ratios for SIA kaon production cross sections as a function of z_h. The data points over theory from ARGUS [152], BELLE [164], BABAR [158], TASSO [148, 151], TPC [154, 155], OPAL [159, 160], ALEPH [161], DELPHI [162], and SLD [157] are shown in black, while the 1σ uncertainty band of the fitted result is shown in blue. The labels (c) and (b) indicate that the data are for charm and bottom quark production respectively.

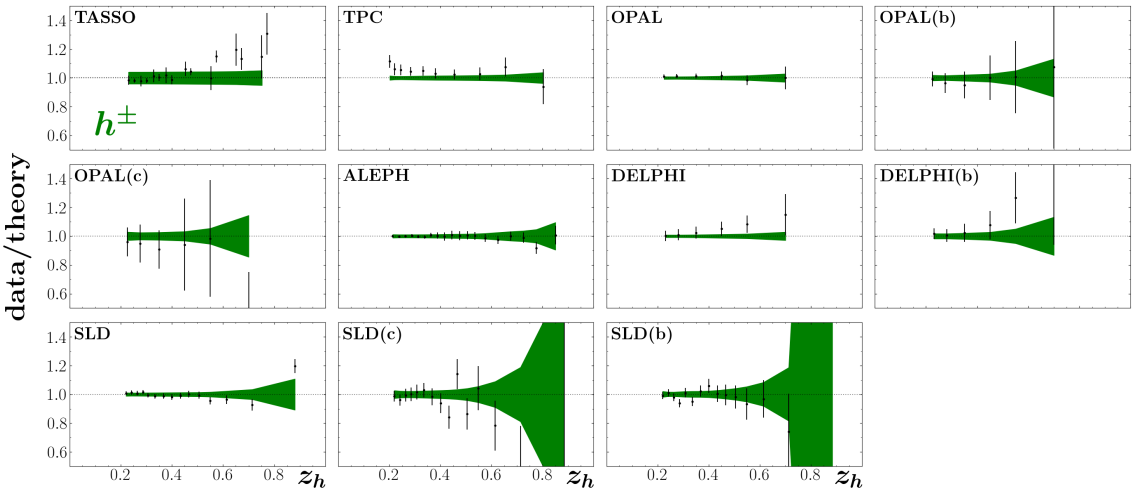

Figure 2.23: Data vs. Theory: SIA (hadron). Data to theory ratios for SIA unidentified charged hadron production cross sections as a function of z_h. The data points over theory from TASSO [151], TPC [154], OPAL [159, 160], ALEPH [161], DELPHI [162], and SLD [157] are shown in black, while the 1σ uncertainty band of the fitted result is shown in green. The labels (c) and (b) indicate that the data are for charm and bottom quark production respectively.

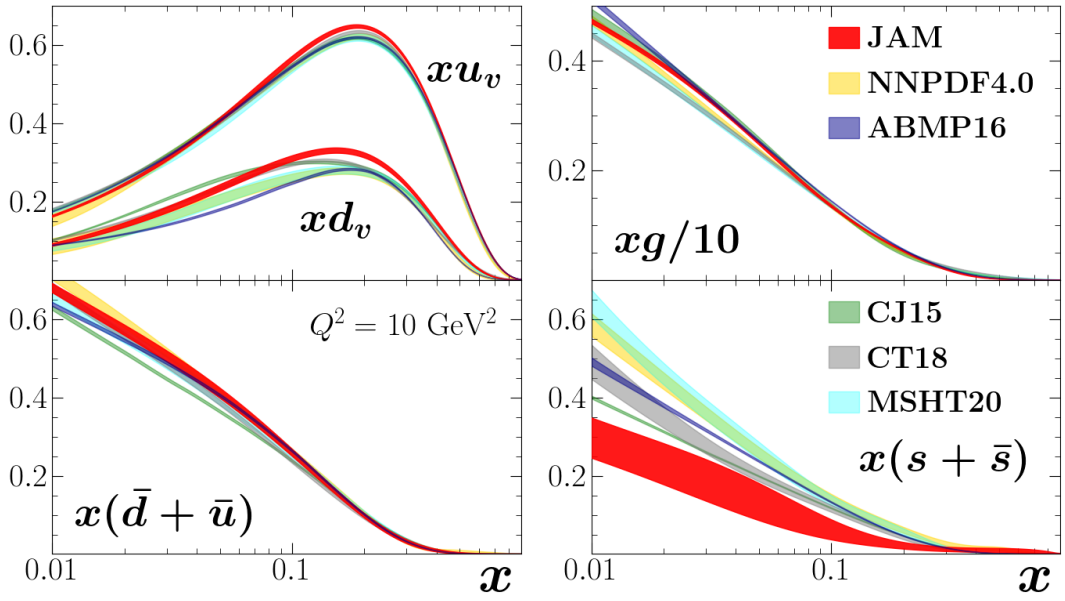

Figure 2.24: Extracted spin-averaged PDFs. Comparison of the JAM PDFs from this analysis (red bands) with NLO results from the NNPDF4.0 [165] (gold), ABMP16 [166] (blue), CJ15 [167] (green), CT18 [168] (gray), and MSHT20 [169] (cyan) parametrizations at the scale $Q^2 = 10$ GeV^2, with the bands representing 1σ uncertainty. Note that x multiplied by the PDFs are shown.

2.6 Extracted Spin-Averaged PDFs

This analysis [56] is based on more than 900 Monte Carlo samples, which we use to ensure the statistical convergence of the PDFs, from which the means and expectation values are then computed using Eqs. (2.4). In this section we show only the spin-averaged PDFs, except for the quark ratio d/u which is closely tied to the extraction of nuclear effects and will be discussed in Chapter 3. The FFs will be shown in Chapter 4, as they play a larger role in the extraction of helicity PDFs than spin-averaged PDFs due to the relative lack of precision in the polarized data. The resulting spin-averaged parton densities are displayed in Fig. 2.24, where we show the valence (u_v and d_v), gluon (g), light antiquark ($\bar{d} + \bar{u}$), and strange ($s + \bar{s}$) distributions at the scale[1] $Q^2 = 10$ GeV^2. The results are compared to several other groups [165–169].

In the valence sector, our results for both u_v and d_v are slightly larger than those from the other groups near the peaks of the distributions. At lower x, the u_v distribu-

[1]We note that here and in the following we often use Q and μ_R interchangeably for the renormalization scale.

tion agrees with the other groups, while the d_v PDF agrees best with the CT18 [168] and NNPDF4.0 [165] parametrizations. For the gluon PDF, our results are largely in agreement with other extractions, although at low x there are some differences with the ABMP16 [166] fit. The same is true for the sum of light sea quark distributions, $\bar{d}+\bar{u}$, but the disagreements are with CJ15 [167]. For the strange distribution $s+\bar{s}$, our result is somewhat suppressed at low x relative to the other extractions. We note the following three points regarding the size of the strange distribution and its potential correlations to the inputs of the analyses shown here. First, this analysis is the only analysis to include SIDIS data, which were found in previous JAM analyses [115, 116] to suppress the strange distribution. Secondly, this analysis, along with CJ15, does not include W/Z boson production data from ATLAS [170–176], which were found to enhance the strange quark distribution at low x. The other four analyses include this data. Finally, the only analysis that parameterizes the nonperturbative charm is NNPDF4.0, which may play a role in the shape of the strange quark distribution [177].

2.6.1 Spin-Averaged Sea Asymmetry

As mentioned previously, the latest SeaQuest and STAR W-lepton results were first included in Ref. [56], and we now discuss the light-quark sea asymmetry and its relation to these new datasets. We note that the SeaQuest collaboration recently released updated results [178] with two methods of extracting the cross-section ratios: the "Intensity-Extrapolation" (IE) method, used in their previous publication [95], and the new "Mass-Fit" (MF) method. They find complete agreement between these two methods, so we continue to use the IE-generated data here and do not consider the MF-generated data. The impact on the light-quark sea asymmetry from both the new STAR measurement [96] of W-lepton cross sections and the SeaQuest measurement [95] of Drell-Yan di-muon production is shown in Fig. 2.25. From the baseline analysis, which excludes these new datasets, the STAR cross section ratios are added first in order to assess their impact. While the STAR data do not lead to significant shifts in the central values of $\bar{d}/\bar{u}$, they do reduce the uncertainties somewhat, by up to $\approx 20\%$ at $x \approx 0.2$. This modest impact can be understood from the fact that the NuSea Drell-Yan measurements are more directly sensitive to $\bar{d}/\bar{u}$ (see Eqs. (2.27) and (2.31)), and already provide the bulk of the constraints on the ratio even when compared to high precision W-lepton and reconstructed W data from the Tevatron

and LHC. Since the STAR data overlap kinematically with the NuSea measurement, it is therefore difficult to improve on the extraction of $\bar{d}/\bar{u}$ using W-lepton production alone (see, in this context, Fig. 2.27 and the corresponding text below).

After adding the STAR data to the baseline, the SeaQuest Drell-Yan cross section ratios are then included. In this case, the SeaQuest data greatly reduce the uncertainties on the $\bar{d}/\bar{u}$ ratio, by up to $\approx 50\%$ at $x \gtrsim 0.3$. Moreover, the addition of the SeaQuest data also increases the $\bar{d}/\bar{u}$ ratio for $x \approx 0.2$, where it remains above unity up to values of $x \approx 0.4$. This is a direct consequence of the extended x range of the data compared with the earlier NuSea results, from $x \approx 0.3$ up to $x \approx 0.4$, with higher precision at the largest x values. This feature is also reflected in the $\bar{d} - \bar{u}$ difference remaining positive across the entire range of x probed.

With the new data from STAR and SeaQuest included, the final $\bar{d}/\bar{u}$ ratio and $x(\bar{d} - \bar{u})$ difference are shown in Fig. 2.26, compared with the corresponding distributions from several other groups [165–169] at $Q^2 = 10$ GeV^2 (see also Refs. [179, 180]). We note that of the other analyses shown in the plot, only the NNPDF4.0 analysis includes the new data from STAR and SeaQuest. Our results are in agreement, within errors, with those from ABMP16 [166] and CT18 [168] (except with ABMP16 at low $x \lesssim 0.04$), whose ratios also remain positive. Although there are differences in the shapes of the ratios, our results also largely agree within errors with those from NNPDF4.0 [165] and CJ15 [167]. The disagreement with CJ15 at high x results from their choice of parametrization that forces $\bar{d}/\bar{u} \to 1$ as $x \to 1$, and, more importantly, the fact that this fit predates the SeaQuest data, which pull the ratio upwards at large x. Similarly, the disagreement with MSHT20 [169] at high x is likely due to the lack of the SeaQuest data in the MSHT20 analysis.

It is instructive to examine the impact of individual datasets on the light-quark sea asymmetry, $x(\bar{d} - \bar{u})$, which we illustrate in the left panel of Fig. 2.27. Starting with inclusive DIS data only, and excluding data from the NMC experiment, the asymmetry is consistent with zero within large uncertainties. Upon the inclusion of the NMC data [123, 124], the errors are significantly reduced, and the asymmetry gives an indication of deviation from zero in the range $0.01 < x < 0.2$. When W-lepton, reconstructed W and Z boson, and jet production data from RHIC, Tevatron, and the LHC are included (but not the new STAR data [96]), the asymmetry becomes significantly larger, and more distinguishable from zero below $x = 0.3$. The new constraints come primarily from the high precision W and W-lepton asymmetry measurements

from the Tevatron and LHC, which are sensitive to $\bar{u}$ and $\bar{d}$ (see Eq. (2.31)). The further addition of the NuSea [94] Drell-Yan data greatly decreases the uncertainty, showing that these data still provide a strong constraint on the asymmetry even when compared to the Tevatron and LHC W and W-lepton asymmetries. Finally, the inclusion of the new SeaQuest [95] and STAR [96] data reduces the uncertainty on the asymmetry even further, while increasing the magnitude at $x \gtrsim 0.2$, as already seen in Fig. 2.25 except now displayed on a logarithmic x scale.

The impact of the various datasets on the antiquark asymmetry can also be represented in the form of the truncated moment, $\int_{0.01}^{1} \mathrm{d}x(\bar{d}-\bar{u})$, illustrated in Fig. 2.27 in the form of the normalized yield of the Monte Carlo replicas. We choose $x = 0.01$ for the lower limit as this is approximately the extent to which existing data provide information on the asymmetry (see Fig. 2.7). We choose $x = 1$ as the upper limit due to the fact that for $\bar{u}$ and $\bar{d}$ the β parameter in Eq. (2.40) is always significantly larger than zero, so that the distributions are very small at large x and there is no need to worry about introducing errors into the integral. Note that because of large uncertainties associated with the small-x extrapolation, estimates of the total moment are not as meaningful without additional constraints on the $x \to 0$ behavior. For the same combinations of datasets as described above, one observes that prior to the addition of the NMC data the truncated moments of the replicas can vary widely between -0.2 and $+0.15$. Once the NMC data are added, the moments become almost entirely positive, and the yield continues to contract as more data are added. Once all of the data are included, the truncated moments are tightly gathered around 0.1.

The shape and magnitude of the $\bar{d} - \bar{u}$ asymmetry has long been an intriguing puzzle for our understanding of the nonperturbative structure of the nucleon. A frequently used interpretation of the excess of $\bar{d}$ over $\bar{u}$ in the proton sea has been that associated with the pion-cloud model and the prevalence of the virtual $p \to n\pi^+$ dissociation [65]. Scattering from the π^+ component of the proton wave function then naturally enhances the $\bar{d}$ distribution, even though some of this will be cancelled by the subdominant $p \to \Delta^{++}\pi^-$ dissociation, which favors $\bar{u}$ over $\bar{d}$. As an illustration, in Fig. 2.28 we compare the inferred JAM asymmetry with $\bar{d}-\bar{u}$ at $x > 0$ calculated from a convolution of the $p \to$ baryon + π splitting functions and the valence PDF of the pion [181–183],

$$(\bar{d}-\bar{u})(x) \;=\; \left[\left(f_{n\pi^+} + f_{\Delta^0\pi^+} - f_{\Delta^{++}\pi^-}\right) \otimes \bar{q}_v^{\pi}\right](x), \tag{2.49}$$

where the convolution integral is defined as $f \otimes q = \int_0^1 \mathrm{d}y \int_0^1 \mathrm{d}z f(y)\, q(z)\, \delta(x - yz)$, and y is the light-cone fraction of the proton's momentum carried by the pion.

For the calculation shown in Fig. 2.28, the pion PDF is taken from the recent NLO analysis of pion-induced Drell-Yan and deep-inelastic leading neutron electroproduction data by Barry *et al.* [184]. The splitting functions are computed at one-pion loop order from chiral effective theory [181–183], using several different models for the ultraviolet regulators [185, 186]. For the regulator mass parameters, we use the values from the global analysis in Ref. [186], for which the integrated splitting function was found to be $\langle n \rangle_{\pi N} = 0.22$. Normalizing the various models of the regulator function to this value, the resulting band in Fig. 2.28 can be taken as a representation of the uncertainty on the calculated asymmetry. The uncertainty also includes the errors from the Monte Carlo analysis, although these are small compared to the variations in the models.

The agreement between the calculated asymmetry and the extracted JAM PDFs is generally good within the uncertainties, with the model giving a slightly harder x dependence compared with the global fit. Qualitative agreement with the extracted asymmetry is also found for various other nonperturbative models (chiral loop and other) discussed in Refs. [63,66–75,77–80]. Since these models are largely in agreement for the unpolarized sea asymmetry, the next step to discriminate between them is to look at the polarized sea asymmetry, where the differences between their predictions are more pronounced. See, in this context, Subsection 4.4.1.

2.7 Summary

In this chapter we have presented the results of a global QCD analysis of spin-averaged PDFs including the SeaQuest measurement of pp and pD Drell-Yan cross sections [95] and the latest STAR measurement of the W-lepton production process [96]. The inclusion of the SeaQuest data reduces the uncertainty on the $\bar{d}/\bar{u}$ ratio by up to 50% at high values of x, and strongly suggests that the $\bar{d} - \bar{u}$ asymmetry remains positive up to $x \approx 0.4$. The impact of the STAR data is less dramatic due to the lower sensitivity of the W-lepton production process to the sea asymmetry when compared to the Drell-Yan process, but the data still leads to a modest reduction of uncertainties in the $x \approx 0.2$ region. Although the new SeaQuest data indicate some tension with the earlier NuSea Drell-Yan measurement [93, 94] at large x, a good

simultaneous description of both datasets is still possible due to the larger relative uncertainties of the NuSea data at high x. There is some difficulty in describing the latest STAR data at high and low pseudorapidities, which is partially due to tension with the NuSea data, but this feature is common among most PDF extractions. The shape and magnitude of $\bar{d}-\bar{u}$ from the global analysis is consistent with expectations from nonperturbative models, such as those based on chiral symmetry breaking in QCD, all of which predict a positive asymmetry up to $x \sim 0.4 - 0.5$. We also examined in detail the impact of various datasets and processes on the asymmetry. Finally, we compared our extracted PDFs to those of other analyses and find some notable differences, particularly for the strange quark PDF. These differences may be explained by the inclusion of SIDIS data, the inclusion of Z boson production data from ATLAS, or the parameterization of the nonperturbative charm.

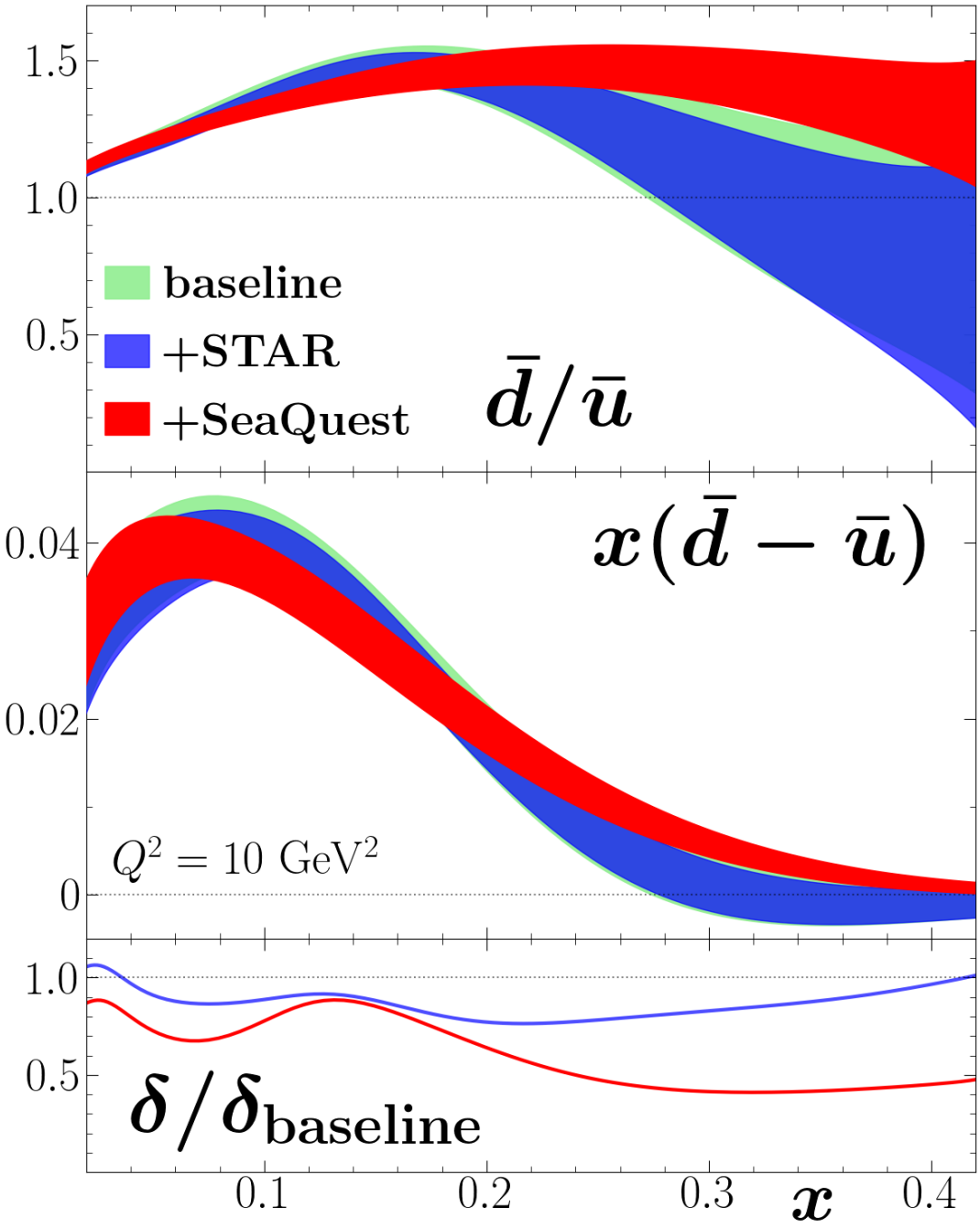

Figure 2.25: Impact of SeaQuest and STAR data on sea asymmetry. Impact on the $\bar{d}/\bar{u}$ ratio (top panel) and the asymmetry $x(\bar{d}-\bar{u})$ (middle panel) from the W-lepton data from STAR [96] (blue bands) and the Drell-Yan measurement from SeaQuest [95] (red bands) relative to the "baseline" (green bands) which contains all data except these. The STAR and SeaQuest data are added in succession. The uncertainty on $\bar{d}/\bar{u}$ for these two scenarios normalized to that of the baseline are shown in the bottom panel. All bands represent 1σ uncertainty.

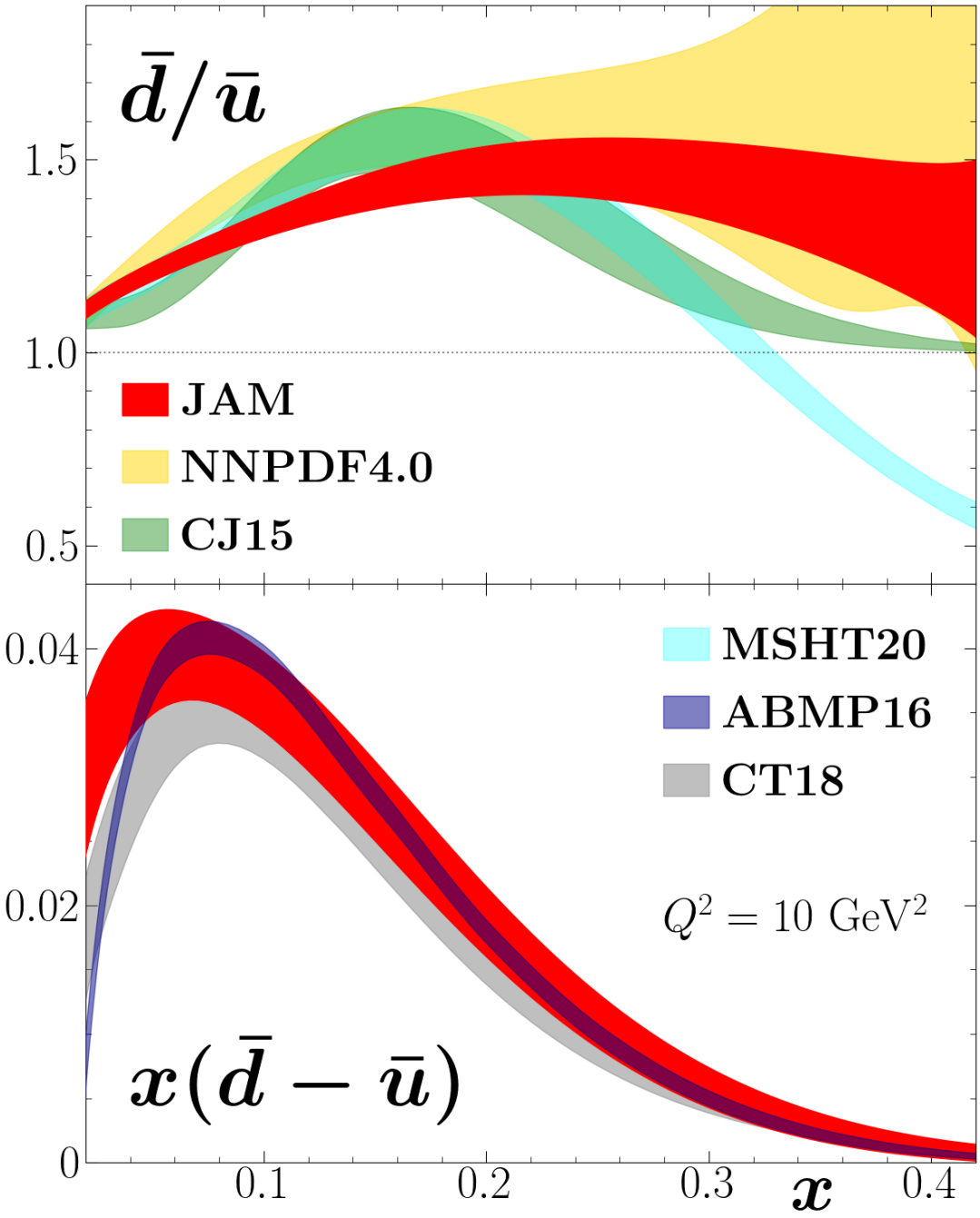

Figure 2.26: Extracted sea asymmetry. Comparison of the JAM $\bar{d}/\bar{u}$ and $x(\bar{d} - \bar{u})$ PDFs (red bands) with the NLO parametrizations from NNPDF4.0 [165] (gold), CJ15 [167] (green), MSHT20 [169] (cyan), ABMP16 [166] (blue), and CT18 [168] (gray) at the scale $Q^2 = 10$ GeV^2. All bands represent 1σ uncertainty.

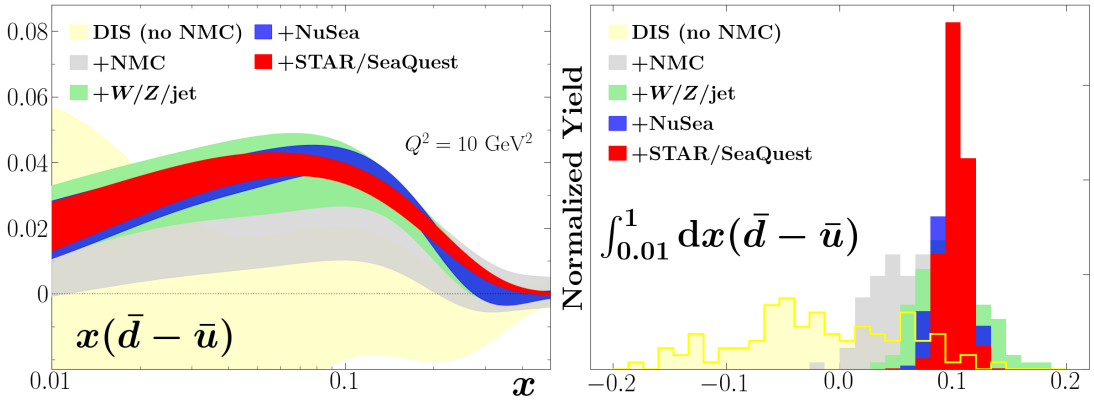

Figure 2.27: Impact of various datasets on sea asymmetry. Comparison of $x(\bar{d} - \bar{u})$ with different combinations of datasets (left panel). First only DIS data [122, 125–127, 130] excluding NMC are included in the fit (gold band). Then data are added successively, with NMC [123, 124] (gray); W-lepton and W asymmetries/Z boson production/jet production from RHIC [143], Tevatron [131–134, 141, 142], and the LHC [135–140] (green); NuSea [92, 94] (blue); and finally the SeaQuest Drell-Yan [95] and STAR W-lepton ratio [96] (red). For the same combinations of data, the normalized yield of the Monte Carlo replicas for the truncated moment $\int_{0.01}^{1} \mathrm{d}x(\bar{d} - \bar{u})$ is also shown (right panel).

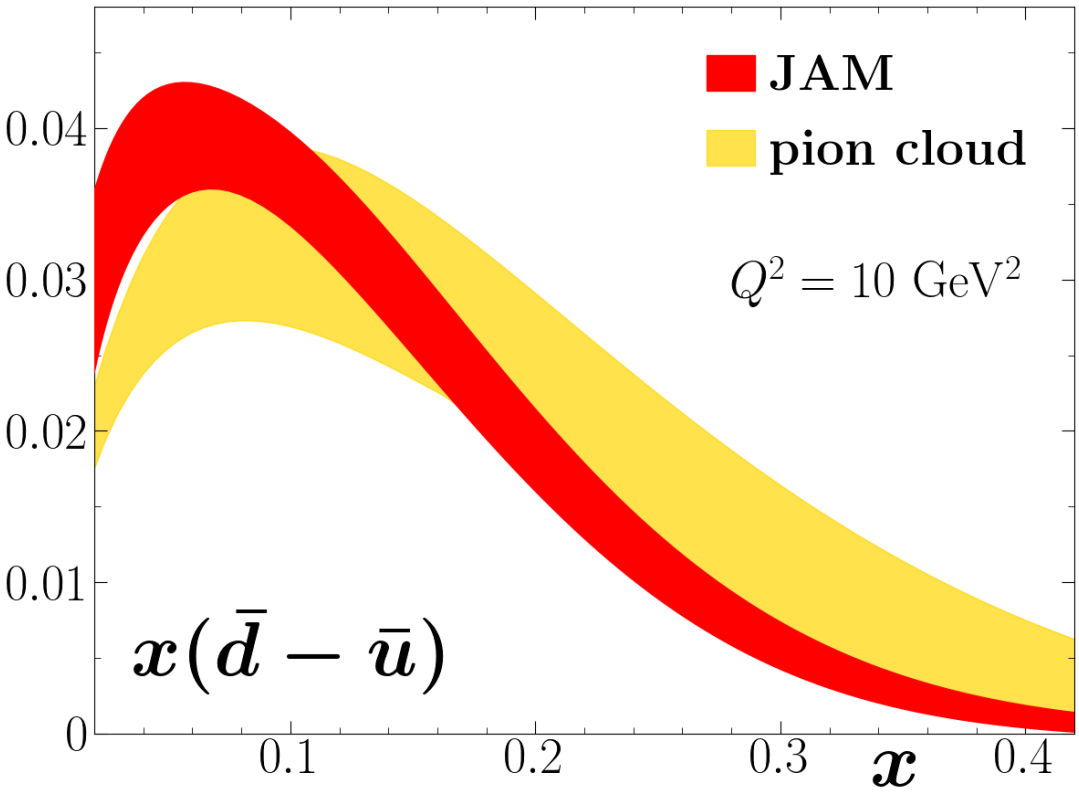

Figure 2.28: Extracted sea asymmetry compared to pion cloud model. Comparison of the extracted JAM $x(\bar{d} - \bar{u})$ distribution (red band) at $Q^2 = 10$ GeV^2 with results from nonperturbative calculations based on chiral symmetry breaking and the pion cloud [181, 182]. The JAM band represents 1σ uncertainty, while the pion cloud band includes model dependence associated with the shape of the ultraviolet regulator function [184–186].

CHAPTER 3

EXTRACTION OF NUCLEAR AND HIGHER TWIST EFFECTS

In this chapter we present the details of a simultaneous global QCD analysis of nuclear and higher twist effects in DIS in the context of spin-averaged PDFs. The general details of such an analysis were discussed in Section 2.1. We note that the analysis in Chapter 2 also contained nuclear and higher twist effects, but the related discussion is reserved for this chapter. We focus in particular on the inclusion of data from the MARATHON experiment and its use in extracting "isovector" nuclear effects, which will be defined below (see Eq. (3.5)).

The scattering of leptons off light nuclear targets has been used to access the structure of the neutron and thus the structure of the down quark, where knowledge at large x has remained elusive [187, 188]. The traditional method of extracting neutron structure from inclusive deuteron DIS has been shown to be handicapped by significant uncertainties in the nuclear corrections at high x [167, 189, 190].

The MARATHON experiment was designed to exploit the mirror symmetry of $A = 3$ nuclei to extract the neutron to proton structure function ratio from the ratio of ^{3}He and ^{3}H cross sections, where nuclear effects are expected to largely cancel [191, 192]. Differences between the free nucleon and nuclear structure functions were first observed by the European Muon Collaboration (EMC) [193], which is now referred to as the "nuclear EMC effect."

The results from the MARATHON experiment that was performed at Jefferson Lab Hall A were presented in Ref. [129]. The experiment measured the ^{3}He/^{3}H cross section ratio in the range of $x_{\rm bj}$ values between 0.195 and 0.825 and Q^2 between 2.7

and 11.9 GeV^2, with the D/p ratio taken over a smaller $x_{\rm bj}$ range as a systematic check. Using as an input the deuteron EMC ratio

$$R(D) \equiv F_2^D/(F_2^p + F_2^n), \tag{3.1}$$

or the super-ratio of the ^{3}He and ^{3}H EMC ratios,

$$\mathcal{R} \equiv R(^3\mathrm{He})/R(^3\mathrm{H}), \tag{3.2}$$

where

$$R(^3\mathrm{He}) \equiv F_2^{^3\mathrm{He}}/(2F_2^p + F_2^n), \tag{3.3}$$

$$R(^3\mathrm{H}) \equiv F_2^{^3\mathrm{H}}/(F_2^p + 2F_2^n), \tag{3.4}$$

the neutron to proton ratio F_2^n/F_2^p can be directly extracted using the measured F_2^D/F_2^p or $F_2^{^3\mathrm{He}}/F_2^{^3\mathrm{H}}$ ratios, respectively.

In the MARATHON analysis [129], the model calculation of Kulagin and Petti (KP) [194] was used to extract F_2^n/F_2^p from $\mathcal{R}$. Assuming that all EMC ratios for $A = 2$ and $A = 3$ nuclei cross unity at $x = 0.31$, an overall normalization of 1.025 was applied to the ^{3}He/^{3}H data to match the ratio F_2^n/F_2^p extracted from F_2^D/F_2^p at $x = 0.31$ using the KP model. While the unity crossing is approximately established empirically from measurements of the EMC effect in heavy nuclei, F_2^A/F_2^D [195, 196], it has not been demonstrated experimentally for light nuclei, with $A \leq 3$. The MARATHON data are unique in their ability to provide information on isovector effects, with n/p values for ^{3}He and ^{3}H ranging from $\frac{1}{2}$ to 2.

In the following we do not assume any prior knowledge of $\mathcal{R}$. In contrast, by assuming the KP model [194] for the nuclear corrections, the MARATHON analysis [129] introduces significant bias into the extracted F_2^n/F_2^p ratio and underestimates the true uncertainties associated with the model dependence of the super-ratio. In particular, while the KP model assumes that the off-shell modifications of bound protons and neutrons are equal and identical for all nuclei [194], this analysis allows a data-driven identification of explicit isospin dependent nuclear effects in $A = 3$ systems. This "isovector" EMC effect can be directly quantified by comparing the PDFs in the

proton bound in ^{3}He and in ^{3}H and defining the quantity

$$\Delta_3^q \equiv \frac{q_{p/^3\mathrm{H}} - q_{p/^3\mathrm{He}}}{q_{p/^3\mathrm{H}} + q_{p/^3\mathrm{He}}}, \tag{3.5}$$

for $q = u$ and d quarks. We note that one can also write (see Eq. (3.15)):

$$\Delta_3^u = \frac{u_{p/^3\mathrm{H}} - d_{n/^3\mathrm{H}}}{u_{p/^3\mathrm{H}} + d_{n/^3\mathrm{H}}}, \quad \Delta_3^d = \frac{d_{n/^3\mathrm{He}} - u_{p/^3\mathrm{He}}}{d_{n/^3\mathrm{He}} + u_{p/^3\mathrm{He}}}, \tag{3.6}$$

which illustrates that we are comparing, within the same nucleus, the difference of up and down quarks within protons and neutrons, respectively, where they are each the dominant valence quark. Hence the term "isovector effect."

Ref. [57] was the first global QCD analysis to include the MARATHON data on F_2^D/F_2^p and $F_2^{^3\mathrm{He}}/F_2^{^3\mathrm{H}}$, as well as the Jefferson Lab Hall C data on $F_2^{^3\mathrm{He}}/F_2^D$ [128]. In this chapter we discuss the results of this analysis, including the first extraction of an isovector nuclear effect. We also show and discuss results for EMC ratios, the F_2^n/F_2^p structure function ratio, the d/u PDF ratio, and extracted higher twist functions.

3.1 Corrections in Deep Inelastic Scattering

The DIS process was discussed in Subsection 2.2.1. Here we discuss corrections to the DIS structure functions that are relevant at low $W^2 \lesssim 10$ GeV^2. These include target mass corrections, higher twist corrections, and nuclear corrections.

3.1.1 Target Mass Corrections

One can consider target mass corrections (TMCs) to the structure functions that take into account the $\mathcal{O}\left(\frac{M^2}{Q^2}\right)$ corrections, as well as potentially higher order terms in the expansion. The topic of TMCs for DIS off the nucleon has been reviewed in Refs. [197, 198], while Ref. [199] provides a recent comprehensive study for nuclear targets. Here we are considering two approaches for dealing with TMCs. The Aivazis-Olness-Tung (AOT) [200] method of TMCs factorizes the structure functions in terms of the so-called Nachtmann variable given by [201]

$$x_\mathrm{N} = \frac{2x_\mathrm{bj}}{1+\rho},$$

where ρ was defined in Eq. (2.17). Note that as $M^2/Q^2 \to 0$, $\rho \to 1$ and $x_N \to x_{\rm bj}$. The AOT structure functions are derived by expanding the hadronic tensor twice. First it is expanded under the Massless Target Approximation (MTA) in terms of the "massless" structure functions $F_i^{p(0)}$, given in Eq. (2.18). Expanding it without approximating the mass to be zero, it can be written in terms of the AOT structure functions $F_2^{p,\rm AOT}$. The two hadronic tensors must be the same regardless of how they are expanded, and relating the two leads to the following expressions for the AOT structure functions in terms of the massless structure functions [198]

$$\begin{aligned}
F_2^{p,\rm AOT}(x_{\rm N},Q^2) &= \frac{1+\rho}{2\rho^2}F_2^{p(0)}(x_{\rm N},Q^2) + \mathcal{O}\Big(\frac{\Lambda^2_{\rm QCD}}{Q^2}\Big), \\
F_L^{p,\rm AOT}(x_{\rm N},Q^2) &= \frac{\rho-1}{2}F_2^{p(0)}(x_{\rm N},Q^2) + F_L^{p(0)}(x_{\rm N},Q^2) + \mathcal{O}\Big(\frac{\Lambda^2_{\rm QCD}}{Q^2}\Big), \qquad (3.7)\\
F_3^{p,\rm AOT}(x_{\rm N},Q^2) &= \frac{1}{\rho}F_3^{p(0)}(x_{\rm N},Q^2) + \mathcal{O}\Big(\frac{\Lambda^2_{\rm QCD}}{Q^2}\Big).
\end{aligned}$$

Note that the massless structure functions are now evaluated at x_N. The expressions for the neutron structure functions are the same with $p \to n$. It is seen that as $\rho \to 1$ one recovers $F_i^{\rm AOT} \to F_i^{(0)}$, with $i = 2, L, 3$.

On the other hand, the Georgi-Poltizer (GP) [202] approach to TMCs uses the operator product expansion to get the following relations between the corrected structure functions and the massless ones [203],

$$\begin{aligned}
F_2^{p,\rm GP}(x_{\rm N},Q^2) &= \frac{(1+\rho)^2}{4\rho^3}F_2^{p(0)}(x_{\rm N},Q^2) \\
&+ \frac{3x(\rho^2-1)}{2\rho^4}\Big[h_2^p(x_{\rm N},Q^2) + \frac{\rho^2-1}{2x\rho}g_2^p(x_{\rm N},Q^2)\Big] + \mathcal{O}\Big(\frac{[M^2,\Lambda^2_{\rm QCD}]}{Q^2}\Big), \\
F_L^{p,\rm GP}(x_{\rm N},Q^2) &= \frac{(1+\rho)^2}{4\rho}F_L^{p(0)}(x_{\rm N},Q^2) \qquad (3.8)\\
&+ \frac{x(\rho^2-1)}{\rho^2}\Big[h_2^p(x_{\rm N},Q^2) + \frac{\rho^2-1}{2x\rho}g_2^p(x_{\rm N},Q^2)\Big] + \mathcal{O}\Big(\frac{[M^2,\Lambda^2_{\rm QCD}]}{Q^2}\Big), \\
F_3^{p,\rm GP}(x_{\rm N},Q^2) &= \frac{1+\rho}{2\rho^2}F_3^{p(0)}(x_{\rm N},Q^2) + \frac{\rho^2-1}{2\rho^3}h_3^p(x_{\rm N},Q^2) + \mathcal{O}\Big(\frac{[M^2,\Lambda^2_{\rm QCD}]}{Q^2}\Big),
\end{aligned}$$

with analogous expressions for the neutron structure functions and where the functions h_2^p, h_3^p, and g_2^p (which should not be confused with the spin-dependent g_2 struc-

ture function) are given by:

$$\begin{aligned} h_2^p(x_{\rm N},Q^2) &= \int_{x_{\rm N}}^1 {\rm d}u \frac{F_2^{p(0)}(u,Q^2)}{u^2}, \\ g_2^p(x_{\rm N},Q^2) &= \int_{x_{\rm N}}^1 {\rm d}u (u-x_{\rm N}) \frac{F_2^{p(0)}(u,Q^2)}{u^2}, \\ h_3^p(x_{\rm N},Q^2) &= \int_{x_{\rm N}}^1 {\rm d}u \frac{F_3^{p(0)}(u,Q^2)}{u}. \end{aligned}$$

It is again seen that as $\rho \to 1$ one recovers $F_i^{\rm GP} \to F_i^{(0)}$, with $i = 2, L, 3$. Note that the AOT and GP structure functions are different in how they account for the $\mathcal{O}\left(\frac{M^2}{Q^2}\right)$ corrections [198].

3.1.2 Higher Twist Corrections

After applying the TMCs, any residual $1/Q^2$ corrections are taken into account by parametrizing them as "higher twist" effects. These can be applied to the structure functions multiplicatively [167] or additively [166] as

$$F_{\rm TMC+HT}^N \equiv F_{\rm TMC}^N * (1 + C_{\rm HT}^N/Q^2), \quad \text{(Multiplicative)} \tag{3.9a}$$

$$F_{\rm TMC+HT}^N \equiv F_{\rm TMC}^N + C_{\rm HT}^N/Q^2. \quad \text{(Additive)} \tag{3.9b}$$

We note that $F_{\rm TMC+HT}^N$ are the quantities that are measured in the experiments, as the experimental observables naturally include all target mass and higher twist corrections. Given sufficient flexibility in the parameterization, both methods should lead to the same result as it is simply a parameterization choice. Both methods are tested in this analysis, and the higher twist is parametrized as

$$C_{\rm HT}^N(x) = h_N x^{a_N} (1-x)^{b_N} (1 + c_N x), \tag{3.10}$$

where $N = p, n$ and $b_N = 0$ when using the multiplicative parameterization [167]. The higher twist function for the proton and neutron structure functions are parameterized and fit independently in this analysis, but we also test the case where $C_{\rm HT}^p = C_{\rm HT}^n$. In the kinematic region where higher twist corrections are relevant, one has $y \ll 1$ and so $Y^- \approx 0$ (see Eq. (2.15)). From Eq. (2.14), it can be seen then that $\sigma_{\rm red} \approx F_2$, and so the higher twist corrections for $F_{L,3}$ are not nearly as relevant. When using the

multiplicative parametrization, the same $C_{\rm HT}^N$ is used for all structure functions $F_{2,L,3}^N$. However when using the additive parametrization, the function is only used for F_2^N and set for zero for $F_{L,3}^N$. The two cases are treated differently as assuming the same absolute higher twist contributions to the three structure functions is clearly wrong, while assuming the same percentage contribution is more reasonable. In theory, TMCs and higher twist effects should be taken into account for the CC structure functions as well. However the CC structure functions are only measured at kinematics where $W^2 > 10$ GeV^2 and we do not include their TMCs or higher twist effects in this analysis.

3.1.3 Nuclear Corrections

In this analysis the nuclear structure functions $F_{2,L,3}^A$ are also included, with $A = D, {}^3\text{He}, {}^3\text{H}$. As a first approximation the nuclear structure functions can be written as $F_i^A = \sum_N F_i^N$, where N sums over the nucleons in A. However, nuclear effects, such as Fermi motion, binding, and the fact that the nucleons are off-shell, become increasingly important as x increases (off-shell effects will be addressed in Subsection 3.1.4).

In order to take these effects into account, the nuclear impulse approximation is used where, at $x \gtrsim 0.2$, the scattering takes place incoherently from individual off-shell nucleons in the nucleus [204–207]. In the weak binding approximation [206,207], appropriate for $A \leq 3$, the nucleus is approximated as a system of weakly bound nucleons with four-momentum $p^\mu = (M + \epsilon, \mathbf{p})$, where ϵ is the off-shell energy, $\mathbf{p}$ is the three-momentum, and one has $|\mathbf{p}|, |\epsilon| \ll M$. The variable $y \equiv \frac{M_A}{M}\frac{p\cdot q}{P\cdot q} = \frac{p_0+\rho p_z}{M}$, where M_A is the mass of A, is the light-cone fraction of the nuclear momentum carried by the interacting nucleon and ranges from $0 < y < M_A/M$. The nucleon's virtuality $v(p^2) \equiv (p^2 - M^2)/M^2$ measures the extent that the nucleon is off-shell. The nuclear structure function can be expanded around $v = 0$ to order v such that the leading term represents the on-shell part of the structure function and the second term of the expansion the off-shell correction.

Focusing first on the on-shell part of the structure functions, to order $\mathbf{p}^2/M^2$ they can be written in the weak binding approximation as a convolution with on-shell

smearing functions [167, 194, 208, 209]

$$
\begin{aligned}
F_2^{A(\mathrm{on})} &= \sum_N \int_x^{M_A/M} \mathrm{d}y \Big[f_{22}^{N/A}(y,\rho) F_2^N(\frac{x}{y}, Q^2) \Big], \\
F_L^{A(\mathrm{on})} &= \sum_N \int_x^{M_A/M} \mathrm{d}y \Big[f_{LL}^{N/A}(y,\rho) F_L^N(\frac{x}{y}, Q^2) + f_{L2}^{N/A}(y,\rho) F_2^N(\frac{x}{y}, Q^2) \Big], \\
F_3^{A(\mathrm{on})} &= \sum_N \int_x^{M_A/M} \mathrm{d}y \Big[f_{33}^{N/A}(y,\rho) F_3^N(\frac{x}{y}, Q^2) \Big].
\end{aligned}
\tag{3.11}
$$

Here the nucleon structure functions have had target mass and higher twist corrections applied. The smearing functions are given by

$$
f_{ij}^{N/A}(y,\rho) = \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \mathcal{F}_0^{N/A}(\epsilon, \mathbf{p}) \big(1 + \frac{\rho p_z}{M}\big) C_{ij} \delta\big(y - 1 - \frac{\epsilon + \rho p_z}{M}\big), \tag{3.12}
$$

where $\mathcal{F}_0^N/A$ is the nonrelativistic nucleon spectral function for the nucleon N in nucleus A and the coefficients are

$$
\begin{aligned}
C_{22} &= \frac{1}{\rho^2}\Big[1 + \frac{\rho^2 - 1}{2y^2 M^2}(2p^2 + 3\mathbf{p}_\perp^2)\Big], \\
C_{LL} &= 1, \\
C_{L2} &= (\rho^2 - 1)\frac{\mathbf{p}_\perp^2}{y^2 M^2}, \\
C_{33} &= 1.
\end{aligned}
$$

The diagonal smearing functions f_{ii}^N are normalized such that integrating over y with $\rho = 1$ yields unity. Note that at finite Q^2 the longitudinal structure function F_L^A receives contributions also from the F_2^N nucleon structure functions, while F_2^A and F_3^A are always "diagonal" [209].

3.1.4 Off-Shell Corrections

As only F_2 has been measured in the region where off-shell corrections are relevant, we consider only off-shell corrections to F_2 and use the notation $f_{N/A}^{(\mathrm{on/off})}$ as the on-shell and off-shell smearing functions for F_2, with the on-shell smearing function corresponding to $f_{22}^{N/A}$ in Eq. (3.12). The off-shell smearing function is the same as Eq. (3.12), except with an extra factor of the nucleon virtuality $v(p^2)$ in the integrand. The off-shell corrections can be formulated at the nucleon level as in [209], but this

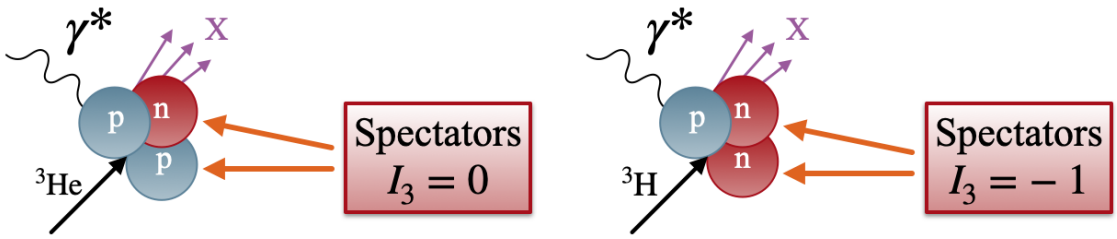

Figure 3.1: Off-shell spectator model. Illustration of the DIS process off of a proton in ^{3}He (left) and ^{3}H (right). For ^{3}He the spectator system is a proton and neutron with third component of isospin $I_3 = 0$. For ^{3}H, the spectator system is two neutrons with $I_3 = -1$.

formulation with isospin dependence introduces explicit charge symmetry breaking, which one ultimately would want to test [210]. On the other hand, by formulating the off-shell corrections at the quark level one can ensure that charge symmetry is respected. The off-shell corrections to the structure functions can be formulated at the quark level within the impulse approximation [57]. The PDFs can be written as a sum of on-shell and off-shell contributions as

$$q_A(x,Q^2) \;=\; \sum_N q_{N/A}(x,Q^2) = \sum_N \left[q^{(\rm on)}_{N/A}(x,Q^2) + q^{(\rm off)}_{N/A}(x,Q^2)\right], \qquad (3.13)$$

where $q_{N/A}$ refers to the PDF of a quark q in a nucleus N, as modified within a nucleus A. One can expand the PDFs in terms of the nucleon virtuality $v(p^2)$ to obtain

$$q_{N/A}(p^2) = q_N + v(p^2)\delta q_{N/A} + ...,$$

where q_N is the on-shell part of the PDF and $\delta q_{N/A}$ the off-shell part. In the weak binding approximation, these two terms can be written as convolutions with smearing functions:

$$\begin{aligned} q^{(\rm on)}_{N/A}(x,Q^2) &= f^{(\rm on)}_{N/A} \otimes q_N(x,Q^2), \\ q^{(\rm off)}_{N/A}(x,Q^2) &= f^{(\rm off)}_{N/A} \otimes \delta q_{N/A}(x,Q^2), \end{aligned}$$

The convolution here is the same as in Eq. (1.7) except that the upper limit is equal to the number of protons and neutrons in the nucleus A.

The off-shell PDFs in Eq. (3.14) cannot be calculated, and so are parametrized

and extracted from data. Since nuclear effects are only relevant at $x \gtrsim 0.2$, only the valence quarks u and d are considered. With $q = u, d$, $N = p, n$, and $A = D, {}^3\mathrm{He}, {}^3\mathrm{H}$, this leads to twelve functions in total. This number is reduced thanks to isospin symmetry, which requires that for deuterium

$$\delta u_{p/D} = \delta d_{n/D}, \quad \delta d_{p/D} = \delta u_{n/D}. \tag{3.14}$$

Similarly, for ${}^3\mathrm{He}$ and ${}^3\mathrm{H}$ one has the relations

$$\delta u_{p/{}^3\mathrm{He}} = \delta d_{n/{}^3\mathrm{H}}, \quad \delta d_{p/{}^3\mathrm{He}} = \delta u_{n/{}^3\mathrm{H}}, \tag{3.15a}$$

$$\delta u_{p/{}^3\mathrm{H}} = \delta d_{n/{}^3\mathrm{He}}, \quad \delta d_{p/{}^3\mathrm{H}} = \delta u_{n/{}^3\mathrm{He}}, \tag{3.15b}$$

reducing the number of functions to six. We emphasize that the relations above are model-independent and hold with an accuracy that is sufficient for our purpose here.

To further reduce the number of functions, we rely on a spectator model, shown in Fig. 3.1 for ${}^3\mathrm{He}$ and ${}^3\mathrm{H}$. In the following, we assume that the photon (we consider only NC here) interacts with a proton inside of the nucleus. Beginning with the deuteron, we have the two off-shell functions

$$\delta u \equiv \delta u_{p/D}\,, \qquad \delta d \equiv \delta d_{p/D}\,. \tag{3.16}$$

For each quark flavor, we decompose the function into its isoscalar and isovector exchange contributions

$$\delta u = \delta u^s_{p/D} + \delta u^v_{p/D}\,, \tag{3.17}$$

$$\delta d = \delta d^s_{p/D} + \delta d^v_{p/D}\,. \tag{3.18}$$

The (magnitude of the) strength of the interaction between the active quark and the spectator neutron is the same for up quarks and down quarks, for both the isoscalar and isovector exchange. However, we use that there are approximately twice as many (valence) up quarks in the proton than (valence) down quarks. This leads to

$$\delta u^s_{p/D} = 2\delta d^s_{p/D}\,, \qquad \delta u^v_{p/D} = -2\delta d^v_{p/D}\,. \tag{3.19}$$

Note the minus sign in the case of the isovector exchange contribution, which is due to the fact that the third component of isospin I_3 has a reversed sign for up quarks

and down quarks. With these relations, one also has

$$\frac{1}{2}(\delta u + 2\delta d) = \delta u^s_{p/D}\,, \tag{3.20}$$

that is, this linear combination does not contain a contribution from the isovector exchange.

We consider now the case for ^{3}He. If we again focus on quarks inside the proton, the spectator system is given by one proton and one neutron. While the total isospin I in this case can be 0 or 1, the third component of the isospin is always $I_3 = 0$. Therefore, we do not consider any contribution from isovector exchange and, by the same reasoning we use above for the deuteron, we find

$$\delta u_{p/^3\mathrm{He}} = \delta u^s_{p/^3\mathrm{He}} = 2\delta d^s_{p/^3\mathrm{He}} = 2\delta d_{p/^3\mathrm{He}}. \tag{3.21}$$

Compared to the deuteron, this leads to one additional off-shell function. In order to simplify our analysis, we relate the off-shell function in ^{3}He to the scalar exchange part in the deuteron according to

$$\delta u_{p/^3\mathrm{He}} = \delta u^s_{p/D} = \frac{1}{2}(\delta u + 2\delta d). \tag{3.22}$$

Finally, we consider ^{3}H, where the spectator system is given by two neutrons with third component of isospin $I_3 = -1$. In analogy to the deuteron we can write

$$\delta u_{p/^3\mathrm{H}} = \delta u^s_{p/^3\mathrm{H}} + \delta u^v_{p/^3\mathrm{H}}\,, \tag{3.23}$$

$$\delta d_{p/^3\mathrm{H}} = \delta d^s_{p/^3\mathrm{H}} + \delta d^v_{p/^3\mathrm{H}}\,, \tag{3.24}$$

as well as

$$\delta u^s_{p/^3\mathrm{H}} = 2\delta d^s_{p/^3\mathrm{H}}\,, \qquad \delta u^v_{p/^3\mathrm{H}} = -2\delta d^v_{p/^3\mathrm{H}}\,. \tag{3.25}$$

This leaves us with two additional off-shell functions, which we relate to the ones discussed above through

$$\delta u^s_{p/^3\mathrm{H}} = \delta u_{p/^3\mathrm{He}} = \delta u^s_{p/D}\,, \qquad \delta u^v_{p/^3\mathrm{H}} = \delta u^v_{p/D}, \tag{3.26}$$

from which one finds immediately the relations

$$\delta u_{p/^3\mathrm{H}} = \delta u\,, \qquad \delta d_{p/^3\mathrm{H}} = \delta d\,. \tag{3.27}$$

Through this model, we have now reduced the six functions down to two: δu and δd. This modeling was done assuming that the boson interacts with a proton, but one can also assume that the photon interacts with a neutron. In that case, an analogous derivation leads to the same relations between the six functions.

To preserve the number of valence quarks in the nucleons bound in a nucleus, the off-shell functions must satisfy

$$\int_0^1 \mathrm{d}x\ \delta u(x) = \int_0^1 \mathrm{d}x\ \delta d(x) = 0. \tag{3.28}$$

The two off-shell functions are parametrized with the template

$$\delta q(x, \mu_0^2) = \frac{N}{\mathcal{M}}\, x^\alpha (1-x)^\beta (1+\eta x), \tag{3.29}$$

where $\mathcal{M} = \mathrm{B}[\alpha+2, \beta+1] + \eta \mathrm{B}[\alpha+3, \beta+1]$ normalizes the function to the second moment to decorrelate the normalization and shape parameters. All quark flavors are set to zero at the input scale except for δu and δd, which evolve with Q^2 in the same way as the on-shell PDFs. The η parameters are fixed by the sum rules Eq. (3.28), while the N, α, and β parameters are inferred from the data.

Note that because the off-shell functions δq are convoluted with the off-shell smearing functions, $f_{N/A}^{(\mathrm{off})}$, their contribution to the total PDF is generally one or two orders of magnitude smaller. The size of the off-shell smearing functions, averaged over y and at $\rho = 1$, are shown in Table 3.1. Note that isospin symmetry relates the neutron smearing functions to the proton smearing functions through $f_{n,D} = f_{p/D}$, $f_{n,{}^3\mathrm{He}} = f_{p,{}^3\mathrm{H}}$, and $f_{n,{}^3\mathrm{H}} = f_{p,{}^3\mathrm{He}}$. It is seen that the off-shell effects in ^{3}He and ^{3}H are larger than those in D, due to the fact that $v(p^2)$ is generally larger in the former nuclei. Thus the MARATHON measurements [129] of $F_2^{^3\mathrm{He}}$ and $F_2^{^3\mathrm{H}}$ at high x are critical for the extraction of off-shell effects. As the baseline for our analysis, we use the Paris [211] wavefunction for $A = 2$ and KPSV [212] for $A = 3$. We have tested the other wavefunctions and find that all of the quantities of interest are largely unaffected by this choice.

3.2 Results

The data used in this analysis, as well as their χ^2_{red} values and normalizations, are summarized in Section 2.5. The resulting fits to the MARATHON F_2^D/F_2^p and $F_2^{^3\mathrm{He}}/F_2^{^3\mathrm{H}}$

	Paris [211]	AV18 [213]	CD-Bonn [214]	WJC1 [215]	WJC2 [216]	KPSV [212]	SS [217]	Average
$\langle f^{(\text{off})}_{p/D} \rangle$	-4.3%	-4.5%	-3.7%	-6.1%	-4.9%	—	—	-4.7%
$\langle f^{(\text{off})}_{p/^3\text{He}} \rangle$	—	—	—	—	—	-6.8%	-5.6%	-6.2%
$\langle f^{(\text{off})}_{p/^3\text{H}} \rangle$	—	—	—	—	—	-9.5%	-8.0%	-8.8%

Table 3.1: Off-shell smearing functions relative to on-shell. Average off-shell proton smearing function values for D, ^{3}He, and ^{3}H. For D, the values are calculated using the Paris [211], AV18 [213], CD-Bonn [214], WJC1 [215], and WJC2 [216] smearing functions, while for ^{3}He and ^{3}H they are calculated using the KPSV [212] and SS [217] smearing functions. The average over all groups is also shown.

data are shown in Fig. 3.2 both without and without the off-shell corrections. For D/p we are able to fit the data well in both cases with a fitted normalization of 1.018(6). For the ^{3}He/^{3}H ratio, the description of the high-x data improves with the inclusion of off-shell corrections, with the χ^2_{red} increasing significantly from 0.31 to 1.29 when the off-shell corrections are not included. This dataset displays by far the largest sensitivity to off-shell corrections, and thus is critical for the extraction of these effects. We repeat that in Ref. [129] a normalization of 1.025(7) was included for this dataset based on results from the KP model [194], which assumes that $R(D)$ and $\mathcal{R}$ are unity at $x = 0.31$. To avoid this model bias, we remove this normalization from the data and instead allow the global fit to determine the normalization. Our fitted value of 1.006(12) is in disagreement with the value from the KP model.

The result for the super-ratio $\mathcal{R}$ in Eq. (3.2) is shown in Fig. 3.3. Our analysis shows that it is consistent with unity until $x_{\text{bj}} \approx 0.6$, at which point it dips and reaches a mean value of 0.96 at $x_{\text{bj}} = 0.825$. The uncertainties on the super-ratio range from $\pm 0.4\%$ at low x_{bj} up to $\pm 3.5\%$ at the highest x_{bj}. Without the MARATHON data the uncertainties on $\mathcal{R}$ (not shown in Fig. 3.3) vary between 1.3% and 6.5%. This improvement demonstrates that the ^{3}He/^{3}H data provide a significant amount of information on the super-ratio. Our results disagree with the KP model [194], which predicts a rise to $\mathcal{R} = 1.01$ at $x_{\text{bj}} = 0.825$ [129]. They also suggest that the uncertainties from the KP model, which are an order of magnitude smaller than our extraction even after the inclusion of the MARATHON data, are significantly underestimated.

The result for the F_2^n/F_2^p structure function ratio is shown in Fig. 3.4. Related to the disagreements in the super-ratio, we also find differences between our result for F_2^n/F_2^p and the extraction in Ref. [129] based on the KP model. We find that while the MARATHON data lowers the central value at low x_{bj} for the n/p ratio, the central value is still well above the KP model extraction. At high x_{bj} values the

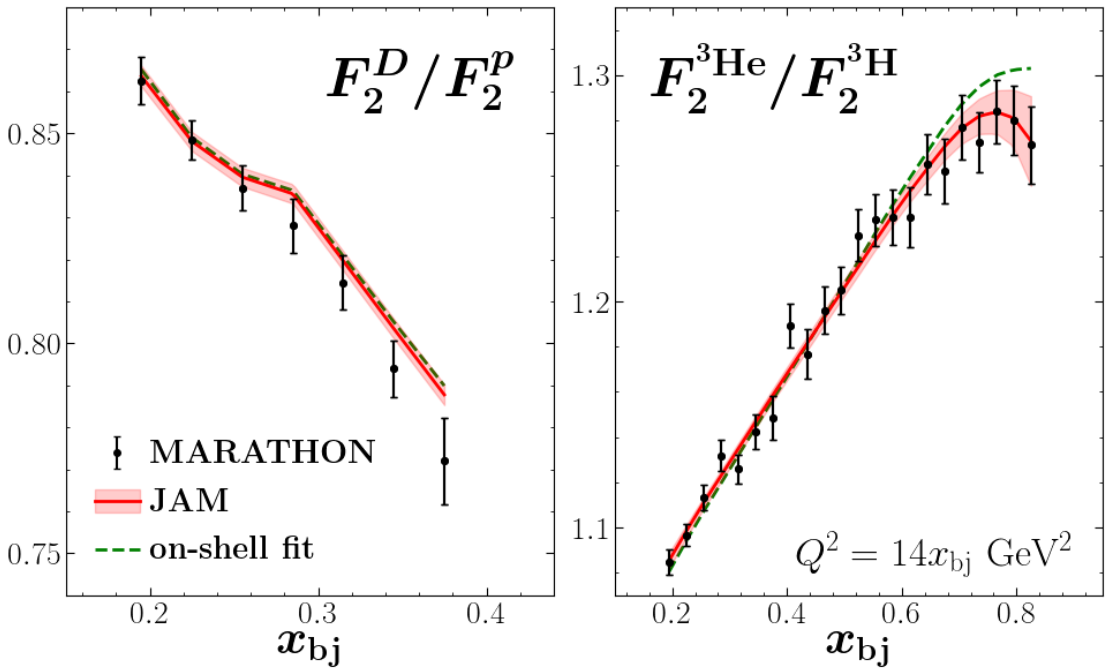

Figure 3.2: Data vs. Theory: MARATHON . Ratios F_2^D/F_2^p (left) and $F_2^{^3\mathrm{He}}/F_2^{^3\mathrm{H}}$ (right) from MARATHON [129] (black circles) as a function of x_{bj} at the experimental kinematics $Q^2 = 14x_{\mathrm{bj}}$ GeV^2 compared with the full JAM fit (red solid lines and 1σ uncertainty bands) and with an on-shell fit (green dashed lines) which sets the off-shell corrections to zero.

disagreements are smaller and the inclusion of the MARATHON data brings our result slightly closer to the KP model extraction.

The d/u quark ratio is shown in Fig. 3.5, where it is seen that the inclusion of the MARATHON data has a very small impact. The small changes for d/u at high x combined with the large differences between the on-shell and off-shell fits at high x (see Fig. 3.2) illustrate an important point: Due to the strong constraints placed on the d/u ratio by vector boson production data, and in particular the W asymmetry data from CDF [133] and D0 [134], the high-x MARATHON data primarily provide new information on nuclear effects, such as the off-shell corrections, which are most relevant in that region.

For the deuteron EMC ratio $R(D)$ (Eq. (3.1)) shown in Fig. 3.6, in the intermediate-x_{bj} region our result is generally in agreement with the CJ15 extraction [167], while at high x_{bj} it is between the CJ15 the AKP17 [189] fits. Notably, we do not see a strong indication for a unity crossing at $x_{\mathrm{bj}} = 0.31$, as was assumed in Ref. [129]. The

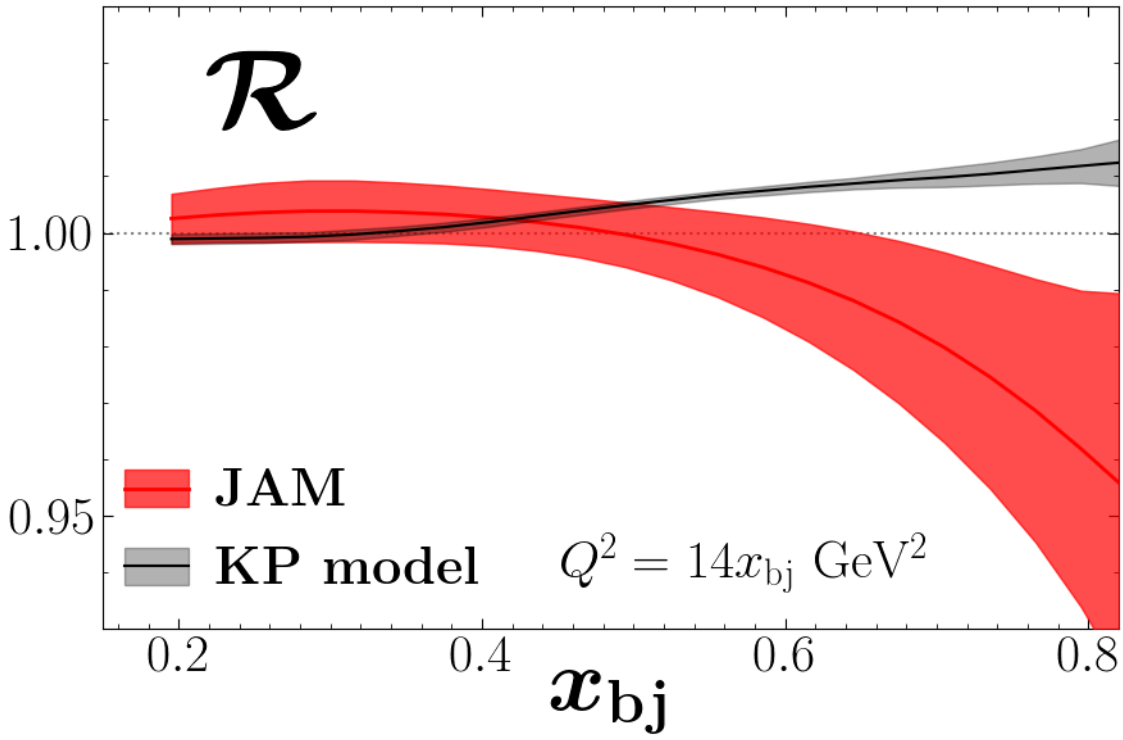

Figure 3.3: Super-ratio $\mathcal{R}$. Result for the super-ratio $\mathcal{R}$ (Eq. (3.2)) as a function of $x_{\rm bj}$ including the MARATHON data (red band) at $Q^2 = 14x_{\rm bj}$ GeV^2. The result is compared with the KP model input (gray band) used to extract the F_2^n/F_2^p ratio in [129].

inclusion of the MARATHON D/p data reduces the ratio in the range $0.2 < x_{\rm bj} < 0.4$.

The impact of the MARATHON data on the off-shell corrections δu and δd is shown in Fig. 3.7. In particular, whereas in the KP model [129, 194] the proton and neutron off-shell effects are set equal, in our analysis we allow flavor dependence of the effects to be determined from the global fit. Indeed, we find that while the $\delta u/u$ ratio is consistent with zero, for the d quark the $\delta d/d$ ratio is enhanced at large values of x. It is important to note that δq does not directly show the contribution to the PDF, as it must first be convoluted with the (small) off-shell smearing function.

Since ^{3}He and ^{3}H are mirror nuclei, the ratio Δ_3^q would vanish if the nuclear corrections were purely isoscalar. Instead, the behavior in Fig. 3.7 indicates some deviations from zero for Δ_3^d at $x \gtrsim 0.4$ and for Δ_3^u at $x \gtrsim 0.6$. The fact that the Δ_3^q are nonzero and of opposite sign for u and d quarks suggests the presence of an isovector component to the EMC effect. This effect is not taken into account in

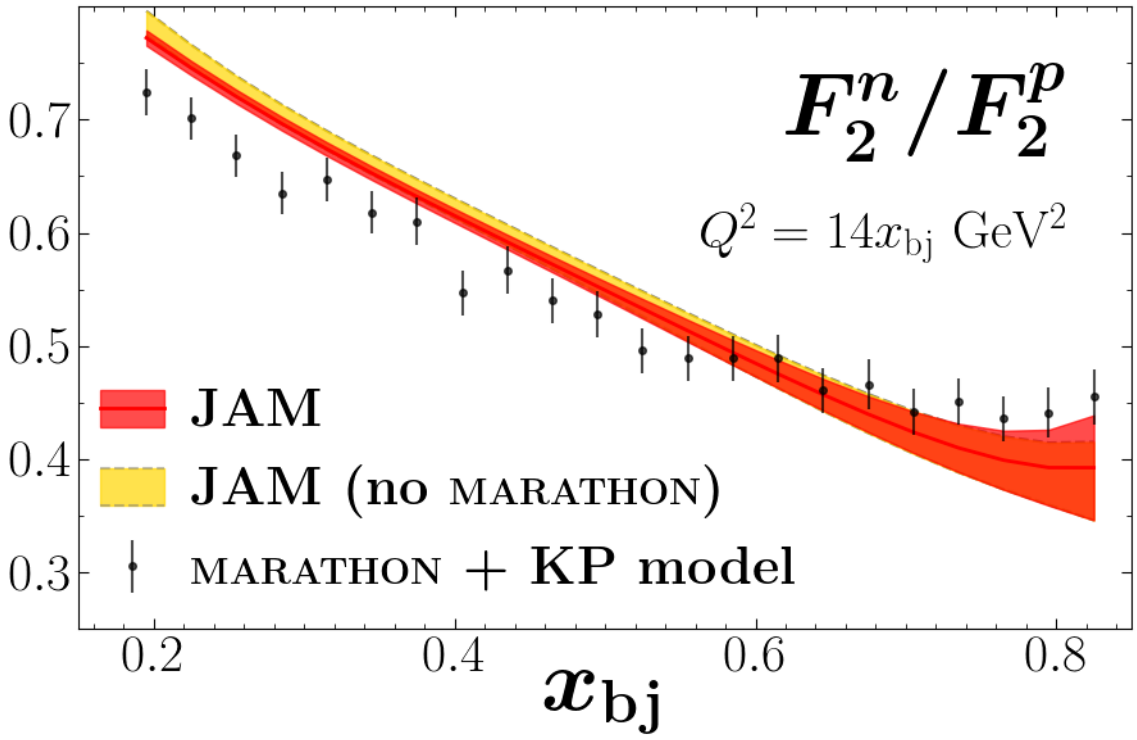

Figure 3.4: F_2^n/F_2^p structure function ratio. Result for the F_2^n/F_2^p ratio as a function of $x_{\rm bj}$ including the MARATHON data (red band) and without the MARATHON data (yellow band) at $Q^2 = 14x_{\rm bj}$ GeV^2. The results are compared to the KP model results of Ref. [129].

standard nuclear PDF analyses [218–221] which assume $u^{p/A} = d^{n/A}$, and thus may impact not only all nuclear PDF fits, but also numerical calculations that utilize nuclear PDFs in quark-gluon plasma simulations in heavy-ion collisions or neutrino-nucleus interactions in high-energy astrophysics.

Finally, we show the extracted higher twist functions in Fig. 3.8 using AOT TMCs [198,200] and the multiplicative HT parameterization Eq. (3.9a). We show the results at a fixed $W^2 = 3.0$ GeV^2, corresponding to the minimum cut on W^2, and extend the $x_{\rm bj}$ range down to $x_{\rm bj} = 0.43$ where $Q^2 = 1.60$ GeV^2, corresponding to the minimum cut on Q^2. Since higher twist effects are most relevant at low Q^2, this allows us to show them at their largest while respecting the W^2 cut on our analysis. Interestingly, at $x_{\rm bj} < 0.65$, we see differences between the proton and neutron $C_{\rm HT}^N$ functions, showing that the data is capable of distinguishing between the two functions when they are not forced to be equal. We also see differences between the additive higher

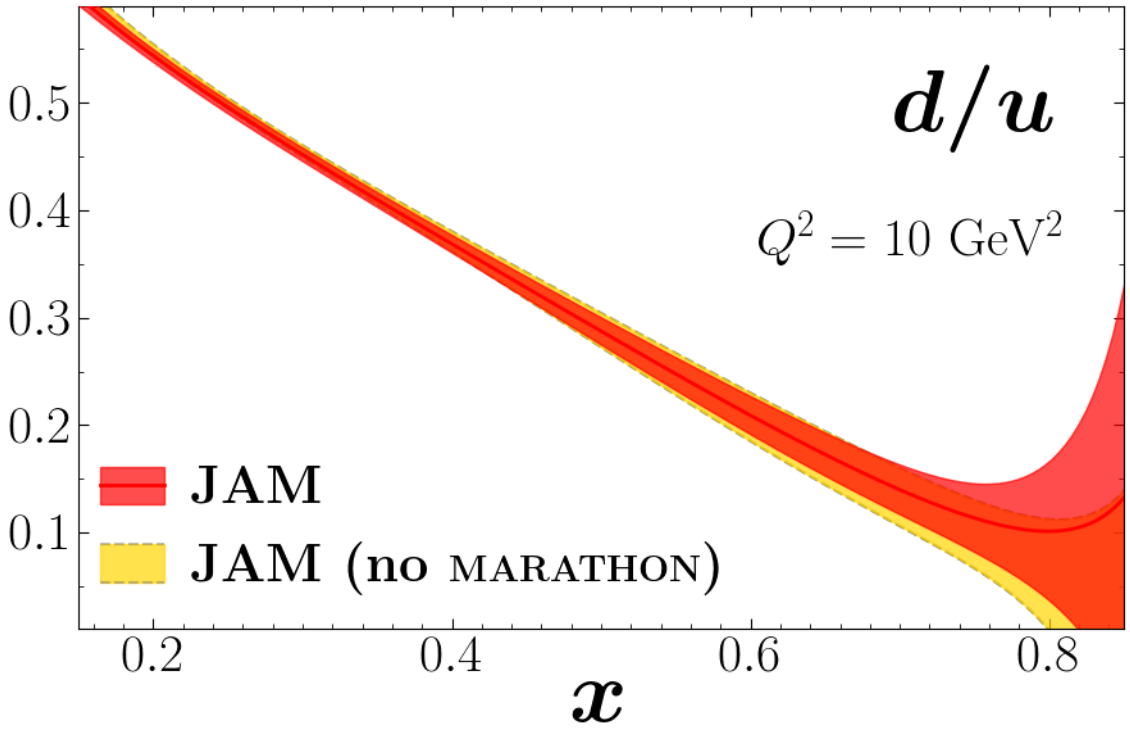

Figure 3.5: d/u quark ratio. Result for the d/u ratio as a function of x including the MARATHON data (red band) and without the MARATHON data (yellow band) at $Q^2 = 10$ GeV^2.

twist contributions to the F_2 structure functions, given by $F_{2,\mathrm{TMC}}^N \times C_{\mathrm{HT}}^N/Q^2$, between the proton and neutron throughout most of the range of x_{bj}. We have tested other scenarios, such as using GP TMCs, using the additive parameterization Eq. (3.9b), and setting $C_{\mathrm{HT}}^p = C_{\mathrm{HT}}^n$. Generally, the description of the data remains very similar, as do the PDFs and off-shell functions. When using GP TMCs, there are some changes in the higher twist functions to compensate for the difference in taking into account the M^2/Q^2 corrections. The higher twists remain similar when using the additive parameterization. We do find that the higher twists depend on the choice of W^2 cut. Thus we refrain from making conclusive statements about the higher twist contributions. We do conclude that all other results are stable regardless of how one chooses to implement TMCs or higher twists.

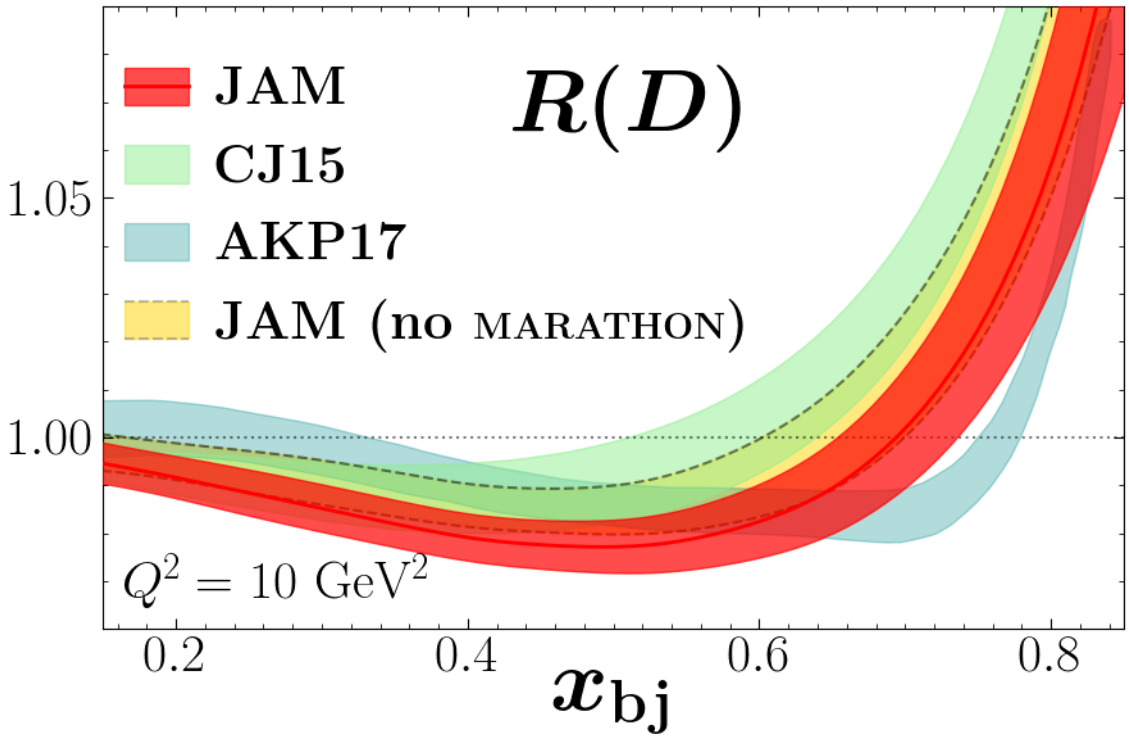

Figure 3.6: Deuteron EMC ratio $R(D)$. Result for the deuteron EMC ratio $R(D)$ (Eq. (3.1)) as a function of $x_{\rm bj}$ including the MARATHON data (red band) and without the MARATHON data (yellow band) at $Q^2 = 10$ GeV^2. The results are compared with that from CJ15 [167] (green band) and AKP17 [189] (light blue band).

3.3 Summary

In this chapter we have presented the results of a global QCD analysis of spin-averaged PDFs and nuclear PDFs including the latest data from the MARATHON collaboration on helium and tritium targets [129]. Our findings were the first indication of an isovector effect in nuclear structure functions, and demonstrate the power of combining the MARATHON ^{3}He/^{3}H data with a global QCD analysis to provide simultaneous information on PDFs and nuclear effects in $A \leq 3$ nuclei. Our extraction shows disagreement with the KP model when it comes to the EMC ratios for deuterium, helium, and tritium as well as the F_2^n/F_2^p ratio. We found that the MARATHON ^{3}He/^{3}H data is particularly sensitive to off-shell corrections at high $x_{\rm bj}$, making it vital for their extraction. We ultimately find a non-zero isovector nuclear effect, which may impact nuclear PDF analyses and all numerical calculations that utilize them. We

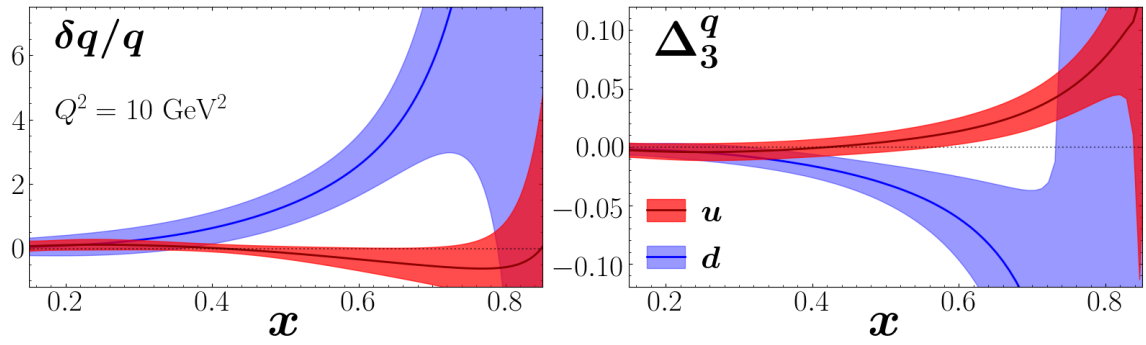

Figure 3.7: Isovector effect. Ratio of off-shell to on-shell PDFs $\delta q/q$ (left) and the difference between proton valence quarks in ^{3}He and ^{3}H normalized to the sum, Δ_3^q (right), as a function of x for valence u (red bands) and d (blue bands) quarks, at $Q^2 = 10$ GeV^2.

have also studied the impact of TMCs and higher twist contributions. While it is difficult to make conclusive statements on the higher twist contributions, we find that our results for the PDFs and off-shell functions are stable regardless of choices relating to TMCs and higher twists.

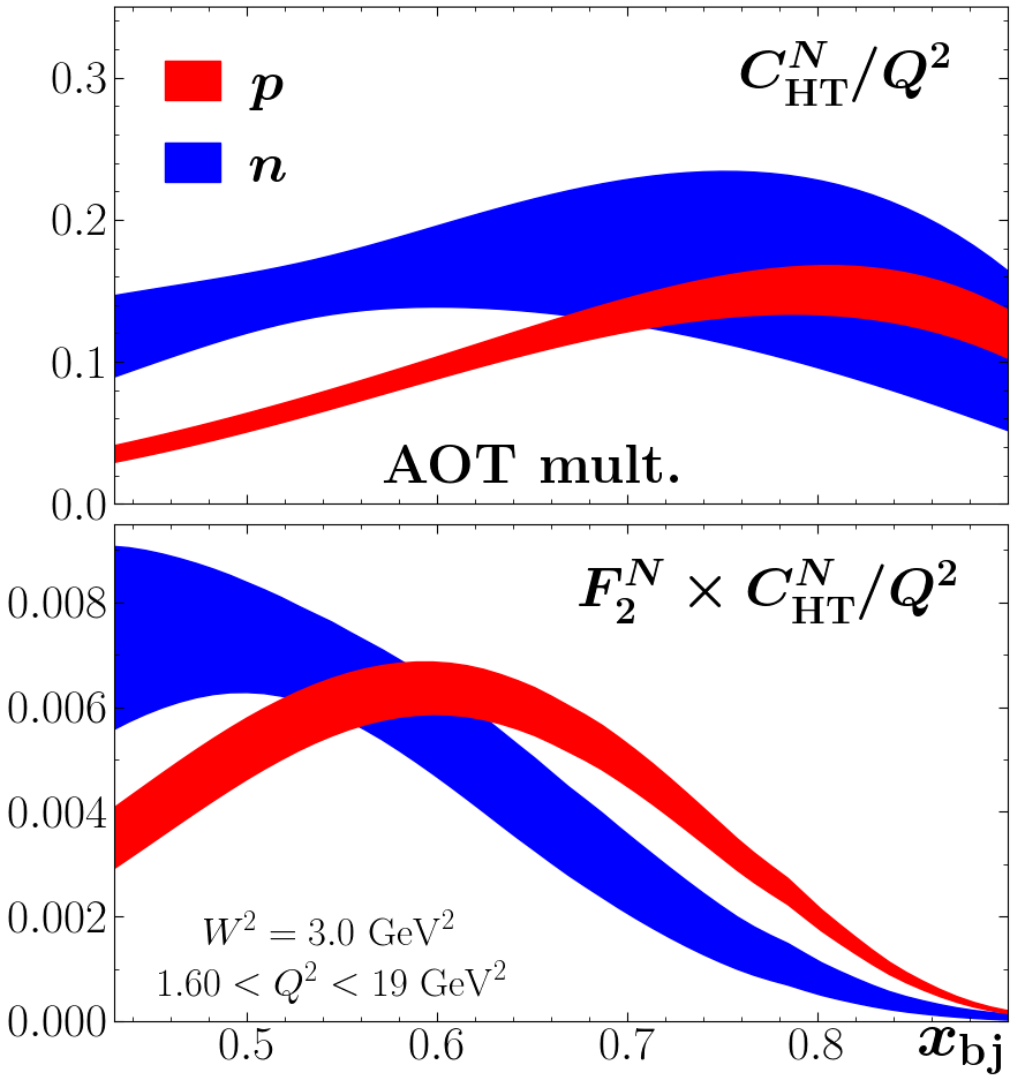

Figure 3.8: Extracted higher twist functions. They are shown as a function of $x_{\rm bj}$ for the proton (red bands) and neutron (blue bands). The top panel shows the multiplicative higher twist function C_{HT}^N divided by Q^2, while the bottom panel shows the additive contribution to the F_2 structure function $F_{2,\rm TMC}^N \times C_{\rm HT}^N/Q^2$. The results are shown at fixed $W^2 = 3.0$ GeV^2, with Q^2 varying from 1.6 GeV^2 at the lowest value of $x_{\rm bj}$ shown (corresponding to the minimum Q^2 cut on this analysis) to 19 GeV^2 at the highest value of $x_{\rm bj}$. Results were extracted using AOT TMCs [198, 200] and the multiplicative HT parameterization Eq. (3.9a).

CHAPTER 4

HELICITY PARTON DISTRIBUTION FUNCTIONS

In this chapter we present the details of a simultaneous global QCD analysis of helicity PDFs, spin-averaged PDFs, and FFs. The general details of such an analysis were discussed in Section 2.1. We focus in particular on the polarized sea asymmetry $\Delta\bar{u}-\Delta\bar{d}$ and its extraction from W-lepton production. While the total light quark contributions to the helicity are well determined from polarized inclusive deep-inelastic scattering (DIS) data [222–235], and jet production in polarized pp collisions [143, 236–242] provides constraints on the gluon helicity [243, 244], far less is known about the polarization of the antiquark sea. There have been some intriguing hints of a polarized antiquark asymmetry, $\Delta\bar{u}-\Delta\bar{d}$, from polarized semi-inclusive DIS (SIDIS) measurements [245–248], in analogy with the spin-averaged $\bar{u}-\bar{d}$ asymmetry inferred from unpolarized DIS and Drell-Yan measurements [87, 88, 93–95] (see also Chapter 2). Various nonperturbative model calculations have also been performed [77, 249–252], some of which predict large positive $\Delta\bar{u}-\Delta\bar{d}$ asymmetries [250–252].

Recently more probes of antiquark polarization have been possible through W-lepton production in polarized pp collisions. In particular, the STAR [253–255] and PHENIX [256, 257] collaborations at RHIC have used polarized pp collisions at COM energy $\sqrt{s} = 510$ GeV to measure the longitudinal single–spin asymmetry. Combined with the DIS observables, these asymmetries provide a vital new handle on the extraction of the polarized antiquark distributions $\Delta\bar{u}$ and $\Delta\bar{d}$.

Previous global analyses [243, 258–260] have sought to extract the asymmetry under various assumptions and with different methods for estimating uncertainties.

De Florian *et al.* (DSSV) [258] extracted a positive $\Delta\bar{u} - \Delta\bar{d}$ asymmetry from spin-dependent data with fixed input for unpolarized PDFs and fragmentation functions (FFs), assuming PDF positivity and SU(3) symmetry for axial-vector charges within errors. The impact of the latter assumptions was examined in a simultaneous analysis of spin PDFs and FFs by the JAM collaboration [114], who found polarized light antiquark and strange PDFs consistent with zero when the constraints were relaxed. The Monte Carlo analysis by the NNPDF collaboration [260] generated prior samples from the DSSV fit [258], thus inheriting the corresponding biases. The NNPDF analysis also used a reweighting procedure involving χ^2-based weights, which is inconsistent with the Gaussian likelihood used in the generation of the replicas [261].

Instead of relying on reweighting prescriptions and assumptions about PDF positivity or flavor symmetry, here we present a new simultaneous global QCD analysis of unpolarized and polarized PDFs and FFs, including for the first time STAR A_L^W data, along with data on inclusive and semi-inclusive polarized lepton-nucleon DIS and jet production in polarized pp collisions [244]. The Monte Carlo analysis allows us to more reliably quantify the uncertainties on all distributions, and examine the interplay between the sea asymmetry and parametrizations of FFs. The simultaneous determination of both types of PDFs also provides the first self-consistent extraction of the antiquark polarization ratios $\Delta\bar{u}/\bar{u}$ and $\Delta\bar{d}/\bar{d}$.

Ref. [58] was the first simultaneous analysis of all of spin-averaged PDFs, helicity PDFs, and FFs, and was also the first to include the latest W-lepton production data from STAR. Such a simultaneous fit is potentially important in order to properly quantify the errors on the helicity PDFs, given that the spin-averaged PDFs and the FFs enter into the observables we consider here that constrain them. In this chapter we show the resulting polarized sea asymmetry and the impact of the new data. In addition, results will be shown for the many other datasets in this analysis, which include data on inclusive and semi-inclusive DIS and jet production. The resulting gluon helicity distribution will also be discussed, along with the FFs compared to previous analyses.

4.1 Processes

In this analysis we include polarized inclusive data on DIS, W-lepton production, and jet production. Further, we include polarized semi-inclusive DIS data. In the

subsections below we summarize each of these processes and the relevant equations.

4.1.1 Polarized Deep Inelastic Scattering

The basics of the DIS process were summarized in Section 2.2.1 [11], with the LO diagram shown in Fig. 2.1. Here we discuss the case where both the lepton and target are polarized. Since no polarized charged current (CC) data exist, only the neutral current (NC) case is discussed below. Analogous to Eq. (2.8), the cross section for polarized DIS is given by:

$$\frac{\mathrm{d}^2\Delta\sigma}{\mathrm{d}x_{\mathrm{bj}}\mathrm{d}y} = \frac{2\pi y\alpha^2}{Q^4}\sum_j \eta_j L^{\mu\nu}_{P,j} W^P_{\mu\nu}, \tag{4.1}$$

where $L^{\mu\nu}_{P,j}$ and $W^P_{\mu\nu}$ are the polarized leptonic and hadronic tensors. The summation is again over $j = \gamma, \gamma Z, Z$ for the NC process, with the factors η_j given in Eq. (2.9). For polarized electrons and NC, the polarized leptonic tensors are given by

$$\begin{aligned} L^{P,\gamma}_{\mu\nu} &= -2i\lambda\epsilon_{\mu\nu\alpha\beta}k^\alpha k'^\beta, \\ L^{P,\gamma Z}_{\mu\nu} &= e\lambda g^e_A L^{P,\gamma}_{\mu\nu}, \\ L^{P,Z}_{\mu\nu} &= 2e\lambda g^e_V g^e_A L^{P,\gamma}_{\mu\nu}, \end{aligned} \tag{4.2}$$

where $e = \pm 1$ and $\lambda = \pm 1$ are the charge (in units of the elementary charge) and helicity of the incoming lepton. The polarized hadronic tensor is given by

$$W^P_{\mu\nu} = i\epsilon_{\mu\nu\alpha\beta}\frac{q^\alpha}{P\cdot q}\Big[S^\beta g_1(x,Q^2) + \big(S^\beta - \frac{S\cdot q}{P\cdot q}P^\beta\big)g_2(x,Q^2)\Big], \tag{4.3}$$

which is defined in terms of the structure functions $g_{1,2}$. Here S^β is the nucleon spin four-vector, with the conventions $S^2 = -M^2$ and $S\cdot P = 0$. Note that the coefficients of g_1 and g_2 are anti-symmetric under $\mu \leftrightarrow \nu$.

In polarized DIS experiments, the measured observables are generally the longitudinal and transverse double spin asymmetries [112]

$$A_\parallel = \frac{\sigma^{\uparrow\Uparrow} - \sigma^{\downarrow\Uparrow}}{\sigma^{\downarrow\Uparrow} + \sigma^{\uparrow\Uparrow}}, \quad A_\perp = \frac{\sigma^{\uparrow\Rightarrow} - \sigma^{\downarrow\Rightarrow}}{\sigma^{\downarrow\Rightarrow} + \sigma^{\uparrow\Rightarrow}}.$$

Here $\uparrow$ ($\downarrow$) denotes the spin of the lepton along (opposite to) the beam direction, $\Uparrow$ denotes the spin of the target along the beam direction, and $\Rightarrow$ denotes the spin of the

target perpendicular to the beam direction. The cross section $\sigma^{\uparrow\Uparrow}$ ($\sigma^{\downarrow\Uparrow}$) is calculated by taking $\lambda = 1$ ($\lambda = -1$) in Eq. (4.3), S^μ along the beam direction in Eq. (4.3), and contracting the resulting tensors in Eq. (4.1). Similarly, for $\sigma^{\uparrow\Rightarrow}$ one takes S^μ to be perpendicular to the beam direction in Eq. (4.3). One then arrives at the result

$$A_\parallel = D\,(A_1 + \eta A_2), \quad A_\perp = d\,(A_2 - \zeta A_1). \tag{4.4}$$

Defining $\gamma^2 \equiv \rho^2 - 1 \equiv \frac{4x^2M^2}{Q^2}$, the kinematic variables are given by

$$\begin{aligned} D &= \frac{y\,(2-y)(2+\gamma^2 y)}{2(1+\gamma^2)y^2 + \big[4(1-y) - \gamma^2 y^2\big](1+R)}, \\ d &= \frac{\sqrt{4(1-y) - \gamma^2 y^2}}{2-y} D, \\ \eta &= \gamma \frac{4(1-y) - \gamma^2 y^2}{(2-y)(2+\gamma^2 y)}, \\ \zeta &= \gamma \frac{2-y}{2+\gamma^2 y}. \end{aligned} \tag{4.5}$$

The virtual photoproduction asymmetries A_1 and A_2 can be written in terms of ratios of the spin-dependent and spin-averaged structure functions,

$$A_1 = \frac{(g_1 - \gamma^2 g_2)}{F_1}, \qquad A_2 = \gamma \frac{(g_1 + g_2)}{F_1}. \tag{4.6}$$

Finally, R is the ratio of longitudinal to transverse photoproduction cross sections and is given in terms of the spin-averaged structure functions by

$$R = \frac{(1+\gamma^2)F_2 - 2xF_1}{2xF_1} = \frac{F_L}{2xF_1}. \tag{4.7}$$

Note that a sum over the γ, γZ, and Z channels has been performed such that $g_{1,2}$ above are defined in terms of the individual channel structure functions $g^\gamma_{1,2}$, $g^{\gamma Z}_{1,2}$, and $g^Z_{1,2}$ as

$$g_{1,2} \;=\; g^\gamma_{1,2} - g^e_V \eta_{\gamma Z} g^{\gamma Z}_{1,2} + \big((g^e_V)^2 + (g^e_A)^2\big)\eta^2_{\gamma Z} g^Z_{1,2}.$$

Neglecting terms of $\mathcal{O}\Big(\frac{[M^2, \Lambda^2_{\rm QCD}]}{Q^2}\Big)$, the NC proton structure function g^p_1 can be

written at leading twist ($\tau = 2$) in terms of the helicity PDFs $\Delta q, \Delta g$ as:

$$g_1^{p(\tau 2)}(x, Q^2) = \frac{1}{2} \sum_q e_q^2 \big[\Delta C_{1,q}^{\rm DIS} \otimes \Delta q^+ + 2 \Delta C_{1,g}^{\rm DIS} \otimes \Delta g \big], \tag{4.8}$$

where the sum q runs over all quark flavors and $\Delta q^+ \equiv \Delta q + \Delta \bar{q}$. The symbol $\otimes$ is defined in Eq. (1.7). The polarized DIS hard scattering coefficients, $\Delta C_{1,f}^{\rm DIS}$ with $f = q, g$, are expanded to NLO in the strong coupling constant

$$\Delta C_{i,f}^{\rm DIS} = \Delta C_{i,f}^{\rm DIS,(0)} + \frac{\alpha_s(Q^2)}{4\pi} \Delta C_{i,f}^{\rm DIS,(1)} + \mathcal{O}(\alpha_s^2),$$

with the coefficients taken from [49]. The hard scattering coefficients depend on the renormalization scale, while the PDFs depend on the factorization scale, both of which are taken, as in the unpolarized case, to be $\mu_R = \mu_F = Q$ for the DIS process. For the neutron functions $g_1^{n(\tau 2)}$, the same proton PDFs are used except with the switch $u \leftrightarrow d$ that is derived from isospin symmetry.

In the following analysis, we will use a cut on the invariant mass squared of $W^2 > 10$ GeV^2 and assume that $\gamma \ll 1$, from which it follows that $\eta \to 0$. The asymmetries then approach:

$$A_\parallel \to \frac{y(2-y)}{y^2 + 2(1-y)(1+R)} \frac{g_1}{F_1}, \quad A_1 \to \frac{g_1}{F_1}, \quad A_\perp, A_2 \to 0. \tag{4.9}$$

In the analysis we fit only data on $A_\parallel$ and A_1 using the exact expressions in Eq. (4.4) and Eq. (4.6), and set $g_2 = 0$ as the terms containing it are kinematically suppressed at the kinematics that we consider. We also neglect target mass and higher twist corrections to g_1, which are most significant in the region below $W^2 = 10$ GeV^2. These effects were explored in a previous JAM analysis [112], and will be explored again in a future analysis [262]. Nuclear effects for g_1, on the other hand, can still be relevant with $W^2 > 10$ GeV^2.

Nuclear corrections for the unpolarized structure functions were discussed in Subsection 3.1.3. The polarized nuclear structure functions can be defined analogously to the unpolarized ones as in Eq. (3.12), leading to [263–265]

$$g_i^A(x, Q^2) = \sum_N \int_x^{M_A/M} \frac{{\rm d}y}{y} \Delta f_{ij}^{N/A}(y, \rho) g_j^N(\frac{x}{y}, Q^2). \tag{4.10}$$

Here $\Delta f_{ij}^{N/A}(y,\rho)$ are the polarized smearing functions defined analogously to Eq. (3.12). In the limit of zero nuclear binding and $\rho \to 1$, one can write the nuclear structure functions as weighted sums of the nucleon structure functions:

$$g_1^D = P_1^{p/D} g_1^p + P_1^{n/D} g_1^n, \tag{4.11a}$$

$$g_1^{^3\mathrm{He}} = 2P_1^{p/^3\mathrm{He}} g_1^p + P_1^{n/^3\mathrm{He}} g_1^n, \tag{4.11b}$$

where $P_1^{N/A}$ are the effective polarizations defined as

$$P_1^{N/A} = \int \mathrm{d}y \Delta f_{ii}^{N/A}(y, \rho = 1). \tag{4.12}$$

One can calculate for the deuteron [263] $P_1^{p/D} = P_1^{n/D} = 0.913$ using the Paris spectral functions [211], or $P_1^{p/D} = P_1^{n/D} = 0.940$ using the CD-Bonn spectral functions [214]. For helium, one has $P_1^{p/^3\mathrm{He}} = -0.028$ and $P_1^{n/^3\mathrm{He}} = 0.851$ using the KPSV spectral functions [212], or $P_1^{p/^3\mathrm{He}} = -0.021$ and $P_1^{n/^3\mathrm{He}} = 0.884$ from the SS spectral functions [217].

From this one approximation one sees that, especially for ^{3}He, nuclear corrections cannot be neglected even in the $W^2 > 10$ GeV^2 region where $\rho \approx 1$ is a reasonable approximation. In this analysis we use Eq. (4.10) to calculate g_1^A, which preserves the x dependence (contained in ρ) of the nuclear effects unlike the approximation in Eq. (4.11). We ignore off-shell effects as they are likely to be smaller than the errors on the currently available polarized data [265] and negligible in the $W^2 > 10$ GeV^2 region.

4.1.2 Polarized W-lepton Production

The W-lepton production process is shown in Fig. 2.3. For polarized W-lepton production, data is only available from RHIC, where h_A and h_B both denote protons that are longitudinally polarized. The measured observable is the single longitudinal spin asymmetry, defined as

$$A_L^{W^\pm} \equiv \frac{\mathrm{d}\sigma^{++} + \mathrm{d}\sigma^{+-} - (\mathrm{d}\sigma^{-+} + \mathrm{d}\sigma^{--})}{\mathrm{d}\sigma^{++} + \mathrm{d}\sigma^{+-} + \mathrm{d}\sigma^{-+} + \mathrm{d}\sigma^{--}} \equiv \frac{\mathrm{d}\Delta\sigma^W}{\mathrm{d}\sigma^W}, \tag{4.13}$$

where $\mathrm{d}\sigma^{++}$ ($\mathrm{d}\sigma^{--}$) is the cross section with both protons polarized parallel (anti-parallel) to the proton's momentum and $\mathrm{d}\sigma^{+-}$ ($\mathrm{d}\sigma^{-+}$) is the cross section where h_A is

polarized parallel (anti-parallel) to its momentum and h_B anti-parallel (parallel). We note that in contrast to the other observables (DIS discussed above and jet production discussed below), the observable here is a (longitudinal) *single* spin asymmetry. Such an asymmetry exists due to the fact that the weak force violates parity. Double spin asymmetries also exist for W-lepton production [255], but we do not include them here due to their large uncertainties which lead to them having no impact on the helicity PDFs [260]. The denominator was defined in Eq. (2.30), with the numerator defined analogously [103]

$$\begin{aligned}\frac{\mathrm{d}^2\Delta\sigma^W}{\mathrm{d}p_T^\ell\,\mathrm{d}\eta_\ell} = \frac{2}{p_T^\ell}\sum_{ab}\int_{VW}^{V}\mathrm{d}v\int_{VW/v}^{1}\mathrm{d}w \\ \times\, x_1\Delta f_a(x_1,\mu_F)\,x_2 f_b(x_2,\mu_F)\,\Delta\hat{\sigma}_{ab}^W(v,w,s,\mu_R,\mu_F).\end{aligned} \tag{4.14}$$

It is written as differential in the outgoing lepton's pseudorapidity, η_ℓ, and its transverse momentum, p_T^ℓ. The variables V, W, v and w are defined in Section 2.2.2. The sum over the quark flavors a, b runs over all partonic channels that can contribute to W-lepton production, for which the renormalization and factorization scales are chosen to be the mass of the W boson, $\mu_R = \mu_F = M_W$. The partonic cross sections $\Delta\hat{\sigma}_{ab}^W$ are computed at NLO in the strong coupling $\alpha_s(\mu_R)$, with the NLO expressions used here taken from Ref. [103].

At LO in α_s and considering only up and down (anti)quarks, the asymmetries can be approximated as

$$A_L^{W^+} \propto \frac{\Delta\bar{d}(x_1)u(x_2) - \Delta u(x_1)\bar{d}(x_2)}{\bar{d}(x_1)u(x_2) + u(x_1)\bar{d}(x_2)}, \tag{4.15a}$$

$$A_L^{W^-} \propto \frac{\Delta\bar{u}(x_1)d(x_2) - \Delta d(x_1)\bar{u}(x_2)}{\bar{u}(x_1)d(x_2) + d(x_1)\bar{u}(x_2)}, \tag{4.15b}$$

showing that they are sensitive to the light antiquarks $\Delta\bar{u}$ and $\Delta\bar{d}$ individually. In particular, at large backwards rapidity one has $x_2 \gg x_1$. Since at large x the quarks dominate over the antiquarks, the asymmetries approach

$$A_L^{W^+} \to \frac{\Delta\bar{d}(x_1)}{\bar{d}(x_1)}, \quad A_L^{W^-} \to \frac{\Delta\bar{u}(x_1)}{\bar{u}(x_1)}, \tag{4.16}$$

and thus provide direct sensitivity to the polarized antiquarks. Combined with DIS and SIDIS observables, these asymmetries provide a vital extra handle on the ex-

traction of the polarized antiquarks. We note that the PHENIX collaboration measures the $A_L^{W^\pm/Z}$ asymmetry, which includes contributions from the Z, γ, and γ/Z channels as well. These contributions are added to the numerator (denominator) of the asymmetry, using Eq. (4.14) (Eq. (2.30)) with $\Delta\hat{\sigma}_{ab}^{W} \to \Delta\hat{\sigma}_{ab}^{Z},\ \Delta\hat{\sigma}_{ab}^{\gamma},\ \Delta\hat{\sigma}_{ab}^{\gamma Z}$ ($\hat{\sigma}_{ab}^{W} \to \hat{\sigma}_{ab}^{Z},\ \hat{\sigma}_{ab}^{\gamma},\ \hat{\sigma}_{ab}^{\gamma Z}$) and adjusting the quark channels appropriately. The NLO expressions for $\Delta\hat{\sigma}_{ab}^{Z}$, $\Delta\hat{\sigma}_{ab}^{\gamma}$, $\Delta\hat{\sigma}_{ab}^{\gamma Z}$, $\hat{\sigma}_{ab}^{Z}$, $\hat{\sigma}_{ab}^{\gamma}$, and $\hat{\sigma}_{ab}^{\gamma Z}$ are also taken from Ref. [103].

4.1.3 Polarized Jet Production

The jet production process is shown in Fig. 2.4. For polarized jet production, data is only available from RHIC, where h_A and h_B both denote protons that are longitudinally polarized. The measured observable is the double longitudinal spin asymmetry, defined as

$$A_{LL}^{\text{jet}} \equiv \frac{\mathrm{d}\sigma^{++} - \mathrm{d}\sigma^{+-}}{\mathrm{d}\sigma^{++} + \mathrm{d}\sigma^{+-}} \equiv \frac{\mathrm{d}\Delta\sigma^{\text{jet}}}{\mathrm{d}\sigma^{\text{jet}}}, \tag{4.17}$$

where $\mathrm{d}\sigma^{++}$ is the cross section with both protons polarized parallel to the proton's momentum and $\mathrm{d}\sigma^{+-}$ is the cross section where h_A is polarized parallel to its momentum and h_B anti-parallel. The denominator was defined in Eq. (2.32), while the numerator is defined analogously [104]

$$\begin{aligned}\frac{\mathrm{d}^2\Delta\sigma^{\text{jet}}}{\mathrm{d}p_T^{\text{jet}}\,\mathrm{d}\eta_{\text{jet}}} = \frac{2}{p_T^{\text{jet}}} \sum_{ab} \int_{VW}^{V} \mathrm{d}v \int_{VW/v}^{1} \mathrm{d}w \\ \times\, x_1\Delta f_a(x_1,\mu_F)\, x_2\Delta f_b(x_2,\mu_F)\, \Delta\hat{\sigma}_{ab}^{\text{jet}}(v,w,p^{\text{jet}},\mu_R,\mu_F;r).\end{aligned} \tag{4.18}$$

It is written as differential in the outgoing jet's pseudorapidity, η_{jet}, and its transverse momentum, p_T^{jet}. Here p^{jet} is the momentum of the jet and r the jet radius. The variables V, W, v and w are defined in Section 2.2.2. The sum over the quark flavors a, b runs over all partonic channels that can contribute to jet production, for which the renormalization and factorization scales are chosen to be the jet's transverse momentum, $\mu_R = \mu_F = p_T^{\text{jet}}$. The partonic cross sections $\Delta\hat{\sigma}_{ab}^{\text{jet}}$ are again computed at NLO in the strong coupling $\alpha_s(\mu_R)$, with the NLO expressions used here taken from Refs. [104, 105]. We note that a, b can be a gluon, and thus the single jet production process is sensitive to the gluon helicity even at LO. Thus the polarized jet production process is very sensitive to the gluon helicity and is vital for its extraction.

4.1.4 Polarized Semi-Inclusive Deep Inelastic Scattering

The SIDIS process is shown in Fig. 2.5. In the case of having both the target and incoming lepton polarized, and ignoring target mass effects, one can define the virtual photoproduction asymmetry A_1^h for polarized SIDIS analogous to the polarized DIS asymmetry A_1 (see Eq. (4.6)):

$$A_1^h(x,z,Q^2) = \frac{g_1^h(x,z,Q^2)}{F_1^h(x,z,Q^2)}. \tag{4.19}$$

The unpolarized structure function F_1^h was defined in Eq. (2.34b) and g_1^h is defined analogously to g_1 (see Eq. (4.8)):

$$\begin{aligned} g_1^{h(p)}(x,z,Q^2) = \frac{1}{2}\sum_{q=u,\bar{u},\ldots,\bar{s}} e_q^2\big[&\Delta C_{1,qq}^{\text{SIDIS}} \otimes \Delta q \otimes D_q^h \\ &+ \Delta C_{1,gq}^{\text{SIDIS}} \otimes \Delta q \otimes D_g^h \\ &+ \Delta C_{1,qg}^{\text{SIDIS}} \otimes \Delta g \otimes D_q^h\big] + \mathcal{O}\Big(\frac{[M^2, \Lambda_{\text{QCD}}^2]}{Q^2}\Big), \end{aligned} \tag{4.20}$$

where the sum q runs over all quark flavors and e_q is the charge of the quark of flavor q. The polarized SIDIS hard scattering coefficients, $C_{1,ff}^{\text{SIDIS}}$ with $f = qq, gq, qg$, are expanded to NLO in the strong coupling constant $\alpha_s(Q^2)$ with the coefficients taken from [106]. The hard scattering coefficients depend on the renormalization scale, while the PDFs and FFs depend on the factorization scale, both of which are taken to be $\mu_R = \mu_F = Q$ for the polarized SIDIS process. For the neutron functions $g_1^{n,h}$, the same proton helicity PDFs are used except with the switch $\Delta u \leftrightarrow \Delta d$ that follows from isospin symmetry. For the nuclear structure functions we use Eq. (4.11) to take nuclear effects into account. Note that the DIS observable provides sensitivity to Δq^+ (see Eq. (4.8)), while the SIDIS observable provides sensitivity to Δq and $\Delta\bar{q}$ individually. Thus the combination of DIS and SIDIS data can, in theory, provide information on the sea quarks, with pion production in SIDIS, in particular, providing sensitivity to $\Delta\bar{u}$ and $\Delta\bar{d}$ and kaon production to Δs and $\Delta\bar{s}$.

4.2 Parametrization

The parameterization of the unpolarized PDFs was discussed in Section 2.3. For the polarized PDFs and FFs at the input scale $\mu_0 = m_c$ we use the template Eq. (2.40),

except now with $\gamma = 0$ and $\mathcal{M} = \mathrm{B}[\alpha+1, \beta+1] + \eta \mathrm{B}[\alpha+2, \beta+1]$ normalizing the function to the first moment. This is in contrast to the spin-averaged PDFs being normalized to the second moment in Section 2.3 due to the momentum sum rule depending on the second moment and the first moments not existing for the spin-averaged PDFs. The first moment was chosen for the helicity PDFs in previous JAM analyses [112] due to potential correlations from fitting the weak neutron and hyperon decay constants (which are proportional to the first moment). Although we no longer fit these constants in the current analysis, we keep the convention of normalizing to the first moment. To characterize the nucleon valence region and discriminate it from the sea components, we parametrize the light-quark and strange PDFs according to

$$\begin{aligned} \Delta u &= \Delta u_v + \Delta \bar{u}, \qquad & \Delta d &= \Delta d_v + \Delta \bar{d}, \\ \Delta \bar{u} &= \Delta S + \Delta \bar{u}_0, \qquad & \Delta \bar{d} &= \Delta S + \Delta \bar{d}_0, \\ \Delta s &= \Delta S + \Delta s_0, \qquad & \Delta \bar{s} &= \Delta S + \Delta \bar{s}_0, \end{aligned} \tag{4.21}$$

where the dependence on x and the scale μ_0^2 has been suppressed for convenience. The input quark distributions Δu_v, Δd_v, $\Delta \bar{u}_0$, $\Delta \bar{d}_0$, Δs_0, and $\Delta \bar{s}_0$, as well as the gluon distribution Δg, are parametrized individually as in Eq. (2.40). For the sea quark PDFs, the additional function ΔS is also parametrized via Eq. (2.40), and is designed to allow a more singular small-x behavior compared to the valence distributions by allowing the corresponding α parameter to more negative values. The parametrization is such that all sea quarks have the same small-x behavior. We note that this parameterization is slightly less flexible than the one used for the spin-averaged PDFs in Section 2.3, with $\gamma = 0$ and all of the sea quarks sharing the same ΔS function. The former restriction can be afforded due to the fact that the polarized data are significantly less precise compared to the unpolarized data. The latter restriction does not affect the results of the analysis since there are no polarized data below $x \approx 0.01$.

The same template (Eq. (2.40)) is used for FFs (normalized to the second moment), but with x replaced by the momentum fraction z of the parton carried by the hadron, and with $\eta = \gamma = 0$ for the pion and kaon FFs. For all FFs we parameterize D_g^h with the template Eq. (2.40). For the π^+ FFs, we assume charge symmetry, $D_u^{\pi^+} = D_{\bar{d}}^{\pi^+}$, $D_d^{\pi^+} = D_{\bar{u}}^{\pi^+}$, as well as $D_q^{\pi^+} = D_{\bar{q}}^{\pi^+}$ for the heavier quarks $q = s, c, b$, while for the K^+ FFs we take $D_d^{K^+} = D_{\bar{u}}^{K^+} = D_{\bar{d}}^{K^+}$ and $D_q^{K^+} = D_{\bar{q}}^{K^+}$ for $q = c, b$, and allow $D_u^{K^+}$ and $D_{\bar{s}}^{K^+}$ FFs to differ. $D_s^{K^+}$ is parameterized independently. Although

the $D_u^{K^+}$ and $D_{\bar{s}}^{K^+}$ are the favored FFs for the kaon, we allow them to differ due to the fact that for u it is necessary to form a secondary $(s, \bar{s})$ pair before forming a $K^+ = (u, \bar{s})$. This is not true for $\bar{s}$, so we expect $D_u^{K^+}$ to be suppressed relative to $D_{\bar{s}}^{K^+}$ [266] (see Fig. 4.18). The FFs for negatively charged mesons are related by $D_q^{\pi^-/K^-} = D_{\bar{q}}^{\pi^+/K^+}$ for all flavors. We use two terms, each of the form Eq. (2.40), for $D_u^{\pi^+}$, $D_d^{\pi^+}$, $D_u^{K^+}$, and $D_d^{K^+}$, and one for all other pion and kaon FFs.

The parametrizations for the unidentified hadron FFs are identical to those in Ref. [116]. Namely, we parametrize the residual $D^{\delta h^+} \equiv D^{h^+} - D^{\pi^+} - D^{K^+}$; have γ and η free; and assume $D_u^{\delta h^+} = D_d^{\delta h^+} = D_s^{\delta h^+}$, $D_{\bar{u}}^{\delta h^+} = D_{\bar{d}}^{\delta h^+} = D_{\bar{s}}^{\delta h^+}$, $D_c^{\delta h^+} = D_{\bar{c}}^{\delta h^+}$, and $D_b^{\delta h^+} = D_{\bar{b}}^{\delta h^+}$. Further, we set the N, γ, and η parameters for the antiquarks to be equal to those of the quarks. For the residual FFs, one has $D_q^{\delta h^-} = D_{\bar{q}}^{\delta h^+}$. We tested that adding further flexibility to the FFs, such as $\eta \neq 0$ for the pion and kaon, does not affect the quality of the fit nor the extracted distributions. Overall, 35 leading-twist PDFs and FFs are fitted with a total of 146 parameters. Including parameters for higher twist and off-shell corrections to structure functions, plus data normalizations, brings the number of parameters to 227.

Recently the question of PDF positivity beyond leading order in α_s in the $\overline{\text{MS}}$ scheme has been disucssed [267, 268]. Such a constraint would require $|\Delta f(x, \mu_R^2)| \leq f(x, \mu_R^2)$ to hold for all flavors at all x and μ_R^2. To explore this question phenomenologically, we perform analyses with and without the positivity constraints. The baseline analysis, referred to in the following as "JAM", does not enforce positivity; however, when included, the positivity constraints are enforced approximately on each Monte Carlo replica by imposing a penalty on the χ^2 function when the bounds are violated [269].

In practice, we rewrite the constraints in terms of the helicity-basis PDFs $f^{\uparrow\downarrow}$

$$f^{\uparrow}(x, \mu_R^2) \equiv q(x, \mu_R^2) + \Delta q(x, \mu_R^2) \geq 0, \tag{4.22}$$

$$f^{\downarrow}(x, \mu_R^2) \equiv q(x, \mu_R^2) - \Delta q(x, \mu_R^2) \geq 0. \tag{4.23}$$

For each replica and at each step of the χ^2 minimization we first calculate the spin-averaged and polarized PDFs at the input scale $\mu_R^2 = \mu_0^2 = m_c^2$ at 100 evenly spaced points in the range $0.1 < x < 0.9$, noting that if positivity is enforced at the input scale it will automatically hold at larger scales [270, 271]. We repeat this process for all of the quarks, antiquarks, and the gluon. Any helicity-basis PDFs that are

negative contribute to the overall χ^2 through

$$\chi^2_{\text{positivity}} = N^2\Big(\sum_x \sum_q \sum_i \Theta\big[-f^i(x,\mu_0^2)\big]\ \big[f^i(x,\mu_0^2)\big]\Big)^2, \tag{4.24}$$

where $i = \uparrow, \downarrow$ and $\Theta(z)$ is the step function

$$\Theta(z) = 0 \text{ if } z < 0,\ \Theta(z) = 1 \text{ if } z > 0.$$

Through testing, it was found that a normalization $N = 150$ is sufficient to ensure that the positivity bounds are satisfied without $\chi^2_{\text{positivity}}$ dominating the entire χ^2 function and decreasing the quality of the fit to the data.

This method of approximately enforcing positivity constraints through a penalty on the χ^2 function is similar to the method of Ref. [269], but with two advantages that come from performing a simultaneous fit of spin-averaged and polarized PDFs. The first is that we can compare the spin-averaged and polarized PDFs for each replica individually, rather than comparing the polarized PDFs to the average of the spin-averaged PDFs as was done in Ref. [269]. The second is that both types of PDFs can be adjusted to ensure that the positivity constraints are satisfied, rather than just the polarized PDFs.

4.3 Data and Quality of Fit

In this analysis we include measurements of the DIS asymmetries $A_\parallel$ and A_1 from fixed-target experiments on proton, deuterium, and ^{3}He from EMC [222], SMC [223, 224], COMPASS [225–227], SLAC [228–233], HERMES [234, 235], and Jefferson Lab [272–276]. The cuts on the four-momentum transfer Q and the hadronic final state masses W for all DIS asymmetries are $Q^2 > m_c^2$ and $W^2 > 10$ GeV^2. For polarized SIDIS, we include measurements of A_1^h on proton, deuterium, and helium targets from HERMES [246,277], COMPASS [247,248], and SMC [245] with the cuts $Q^2 > m_c^2$ and $W^2 > 10$ GeV^2 and an additional cut on the fragmentation variable, $0.2 < z_h < 0.8$, to ensure the applicability of the leading power formalism and avoid large-z threshold corrections [116]. The included SIA data were discussed in Section 2.5.

Beyond DIS and SIA, we include polarized jet production data from STAR [143, 236–241] and PHENIX [242]. We include the single-spin asymmetries $A_L^{W^\pm}$ from

experiment	ref.	beam	target	obs.	$N_{\rm dat}$	$\chi^2_{\rm red}$	fitted norm.
EMC	[222]	μ	p	A_1	10	0.34	0.980(61)
SMC	[223]	μ	p	A_1	11	0.34	—
SMC	[223]	μ	D	A_1	11	1.81	—
SMC	[224]	μ	p	A_1	7	1.42	—
SMC	[224]	μ	D	A_1	7	0.67	—
COMPASS	[226]	μ	p	A_1	11	1.04	0.967(47)
COMPASS	[225]	μ	D	A_1	11	0.57	1.019(57)
COMPASS	[227]	μ	p	A_1	35	0.88	1.003(60)
SLAC E80/E130	[228]	e	p	$A_\parallel$	10	0.77	0.959(56)
SLAC E142	[229]	e	^{3}He	A_1	4	1.03	0.991(4)
SLAC E143	[231]	e	p	$A_\parallel$	39	0.83	1.030(32)
SLAC E143	[231]	e	D	$A_\parallel$	39	1.10	0.989(27)
SLAC E154	[230]	e	^{3}He	$A_\parallel$	15	0.42	0.973(46)
SLAC E155	[233]	e	p	$A_\parallel$	59	1.33	1.103(63)
SLAC E155	[232]	e	D	$A_\parallel$	59	0.96	0.976(36)
HERMES	[234]	e	"n"	A_1	5	0.15	0.995(7)
HERMES	[235]	e	p	$A_\parallel$	16	0.64	0.996(34)
HERMES	[235]	e	D	$A_\parallel$	16	1.07	0.961(33)
Total					365	**0.95**	

Table 4.1: χ^2 table: Polarized DIS A_1 and $A_\parallel$. Summary of the $\chi^2_{\rm red}$ values for the polarized DIS data on A_1 and $A_\parallel$ used in this analysis, as well as their fitted normalizations. Fitted normalizations are not provided in the cases where experimental papers did not provide a normalization uncertainty.

STAR [255], as well as the single-spin asymmetries $A_L^{W^\pm/Z}$ from PHENIX [256, 257]. For the data that is relevant only for spin-averaged PDFs, the included data are the same as those in Section 2.5. The kinematics (at LO in α_s) are shown for all experiments in Fig. 4.1.

The $\chi^2_{\rm red}$ values and fitted normalizations for the polarized DIS data are shown in Table 4.1, while Table 4.2 shows the values for the polarized hadron-hadron collision data. The $\chi^2_{\rm red}$ values and fitted normalizations for the polarized SIDIS data are shown in Table 4.3. Figs. (4.2)–(4.6) show the data and theory comparison for all of the inclusive data included in this analysis. The polarized DIS data are shown in Fig. 4.2 for proton, Fig. 4.3 for the deuteron, and Fig. 4.4 for the "neutron" and helium. The W/Z boson production data is shown in Fig. 4.5, while the polarized jet production process is shown in Fig. 4.6. The data and theory comparisons for polarized SIDIS are shown in Figs. (4.7)–(4.9).

type	experiment	ref.	process	diff.	obs.	N_{dat}	χ^2_{red}	fitted norm.
inclusive jets	STAR	[143]	$pp \to \text{jet}$	y_{jet}, p_T	A_{LL}^{jet}	4	0.22	—
	STAR	[236]	$pp \to \text{jet}$	y_{jet}, p_T	A_{LL}^{jet}	9	0.40	1.005(34)
	STAR	[237]	$pp \to \text{jet}$	y_{jet}, p_T	A_{LL}^{jet}	7	1.51	1.007(10)
	STAR	[238]	$pp \to \text{jet}$	y_{jet}, p_T	A_{LL}^{jet}	18	0.58	0.965(33)
	STAR	[239]	$pp \to \text{jet}$	y_{jet}, p_T	A_{LL}^{jet}	12	1.58	1.019(25)
	STAR	[241]	$pp \to \text{jet}$	y_{jet}, p_T	A_{LL}^{jet}	18	0.85	1.030(42)
	STAR	[240]	$pp \to \text{jet}$	y_{jet}, p_T	A_{LL}^{jet}	13	0.73	0.948(38)
	PHENIX	[242]	$pp \to \text{jet}$	y_{jet}, p_T	A_{LL}^{jet}	2	0.37	1.001(2)
W/Z production	STAR	[255]	$pp \to W \to e\nu_e$	η_e	A_L^W	12	0.65	0.972(25)
	PHENIX	[256]	$pp \to W \to e\nu_e$	η_e	$A_L^{W/Z}$	2	0.21	0.995(9)
	PHENIX	[257]	$pp \to W \to e\nu_e$	η_e	$A_L^{W/Z}$	4	0.64	1.003(2)
Total						101	**0.80**	

Table 4.2: χ^2 table: Polarized pp collisions. Summary of the χ^2_{red} values for the polarized hadron-hadron collision data used in this analysis, as well as their fitted normalizations. All processes are inclusive and the undetected part of the final system X has been suppressed in the "process" column.

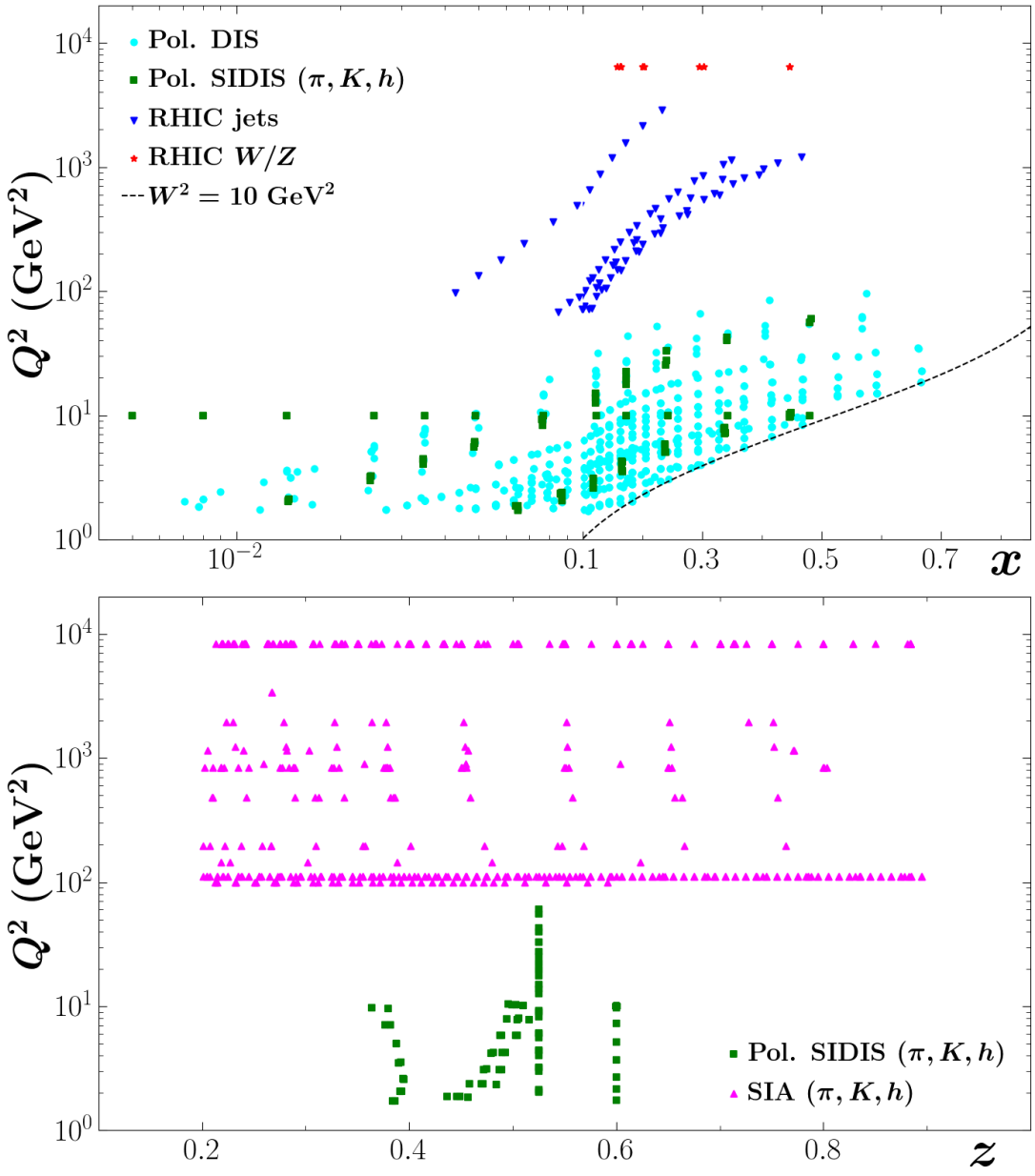

Figure 4.1: Kinematics of polarized datasets. Kinematic coverage of the datasets included in this analysis. The top panel shows the data as a function of x and Q^2. The variable x represents Bjorken-x for polarized DIS and SIDIS and Feynman-x for vector boson and jet production, while the scale Q^2 represents the four-momentum transfer squared for polarized DIS and SIDIS, the mass squared of the intermediate boson for vector boson production, and the transverse momentum squared for jet production. Also indicated is the DIS cut of $W^2 = 10$ GeV^2 (dashed black line). The bottom panel shows the semi-inclusive data as a function of z and Q^2. For polarized SIDIS, z is defined as in Eq. (2.33), while for SIA it is defined as in Eq. (2.36).

experiment	ref.	beam	target	hadron	N_{dat}	χ^2_{red}
HERMES	[246]	e	p	π^+	7	1.34
HERMES	[246]	e	p	π^-	7	0.39
HERMES	[246]	e	D	π^+	7	1.19
HERMES	[246]	e	D	π^-	7	2.57
COMPASS	[247]	μ	p	π^+	10	1.01
COMPASS	[247]	μ	p	π^-	10	1.19
COMPASS	[248]	μ	D	π^+	8	0.37
COMPASS	[248]	μ	D	π^-	8	0.45
HERMES	[246]	e	D	K^+	7	0.69
HERMES	[246]	e	D	K^-	7	0.62
HERMES	[246]	e	D	$K^\pm$	7	0.51
COMPASS	[247]	μ	p	K^+	10	0.18
COMPASS	[247]	μ	p	K^-	10	0.24
COMPASS	[248]	μ	D	K^+	8	0.33
COMPASS	[248]	μ	D	K^-	8	0.58
SMC	[245]	μ	p	h^+	12	1.71
SMC	[245]	μ	p	h^-	12	1.10
SMC	[245]	μ	D	h^+	12	0.61
SMC	[245]	μ	D	h^-	12	1.15
HERMES	[246]	e	p	h^+	7	0.70
HERMES	[246]	e	p	h^-	7	0.55
HERMES	[246]	e	D	h^+	7	1.60
HERMES	[246]	e	D	h^-	7	0.69
HERMES	[246]	e	^{3}He	h^+	7	0.62
HERMES	[246]	e	^{3}He	h^-	7	0.51
COMPASS	[248]	μ	D	h^+	10	0.69
COMPASS	[248]	μ	D	h^-	10	0.98
Total					231	**0.85**

Table 4.3: χ^2 table: Polarized SIDIS. Summary of the χ^2_{red} values for the polarized SIDIS data on A_1^h used in this analysis.

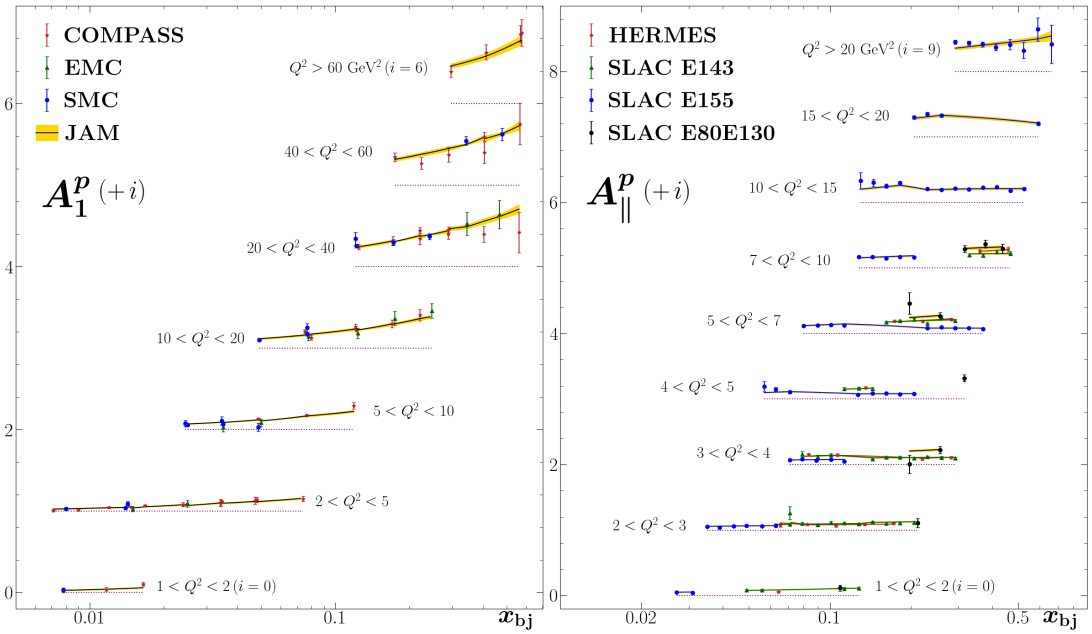

Figure 4.2: Data vs. Theory: Polarized DIS (proton A_1 and $A_\parallel$). All proton polarized inclusive DIS data on A_1 and $A_\parallel$ included in this analysis, plotted as a function of $x_{\rm bj}$ with different ranges of Q^2 against the mean JAM result (colored lines) with 1σ uncertainty bands in gold. The left panel shows A_1 data from COMPASS [226, 227], EMC [222], and SMC [223, 224]. The right panel shows $A_\parallel$ data from HERMES [235] and SLAC [228, 231, 233]. Data in different Q^2 bins are increased by i for clarity, with the dashed purple lines showing the different values of i.

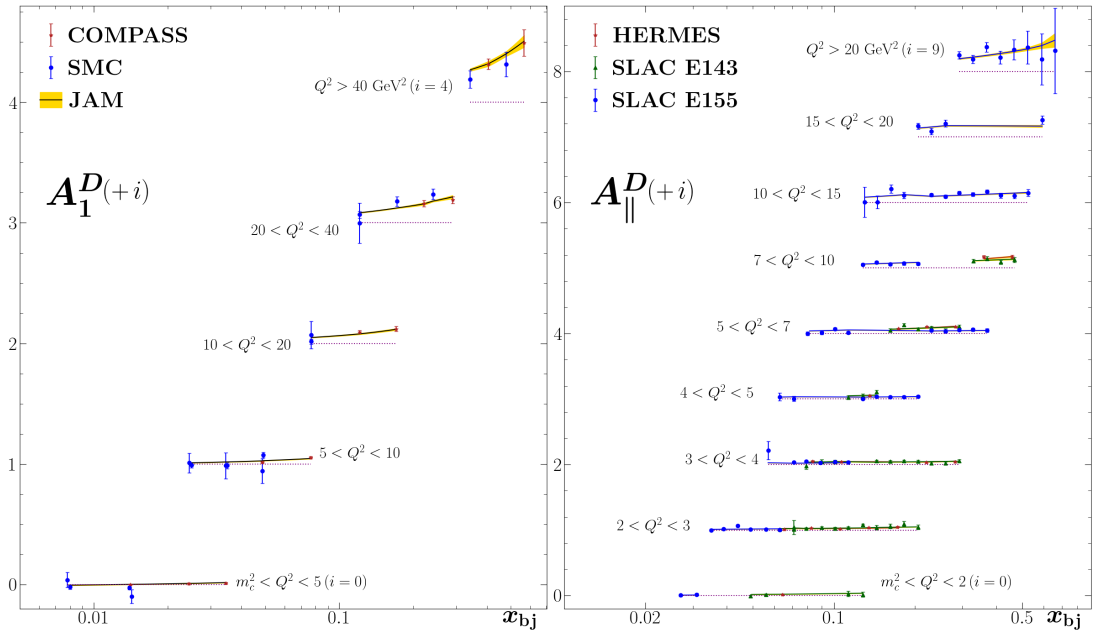

Figure 4.3: Data vs. Theory: Polarized DIS (deuteron A_1 and $A_\parallel$). All deuteron polarized inclusive DIS data on A_1 and $A_\parallel$ included in this analysis, plotted as a function of $x_{\rm bj}$ with different ranges of Q^2 against the mean JAM result (colored lines) with 1σ uncertainty bands in gold. The left panel shows A_1 data from COMPASS [225] and SMC [223, 224]. The right panel shows $A_\parallel$ data from HERMES [235] and SLAC [231, 232]. Data in different Q^2 bins are increased by i for clarity, with the dashed purple lines indicating the different values of i.

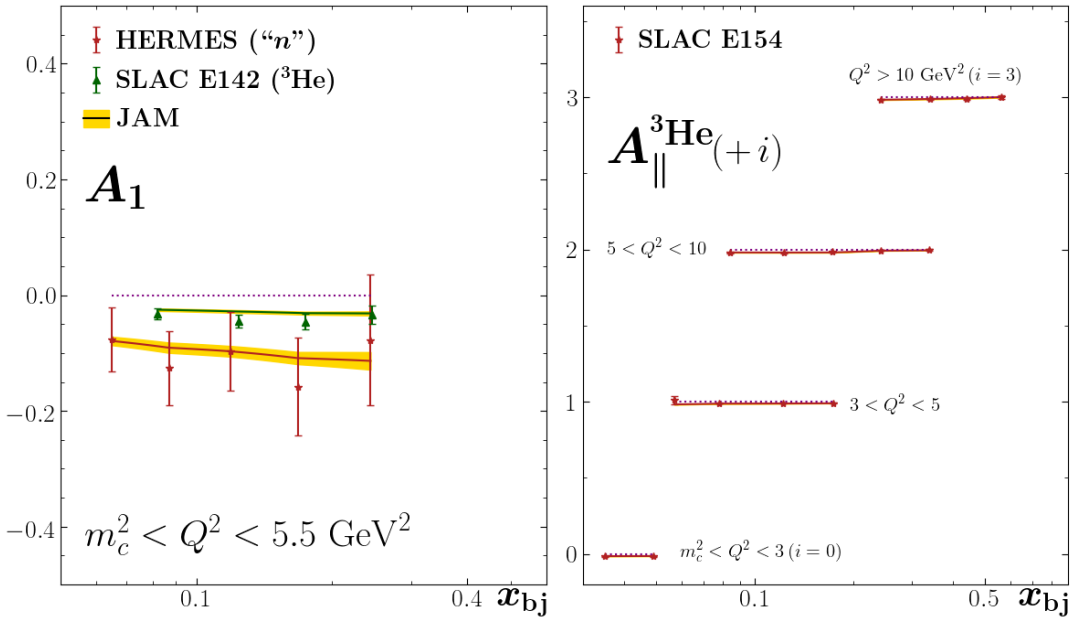

Figure 4.4: Data vs. Theory: Polarized DIS (neutron and helium A_1 and $A_\parallel$). All "neutron" and helium polarized DIS data on A_1 and $A_\parallel$ included in this analysis, plotted as a function of x_{bj} with different ranges of Q^2 against the mean JAM result (colored lines) with 1σ uncertainty bands in gold. The left panel shows A_1 "neutron" data from HERMES [234] and helium data from SLAC [229]. The right panel shows $A_\parallel$ helium data from SLAC [230]. Data in different Q^2 bins are increased by i for clarity, with the dashed purple lines indicating the different values of i. The "neutron" data from HERMES was extracted using a measurement on a ^{3}He target, corrected for nuclear effects.

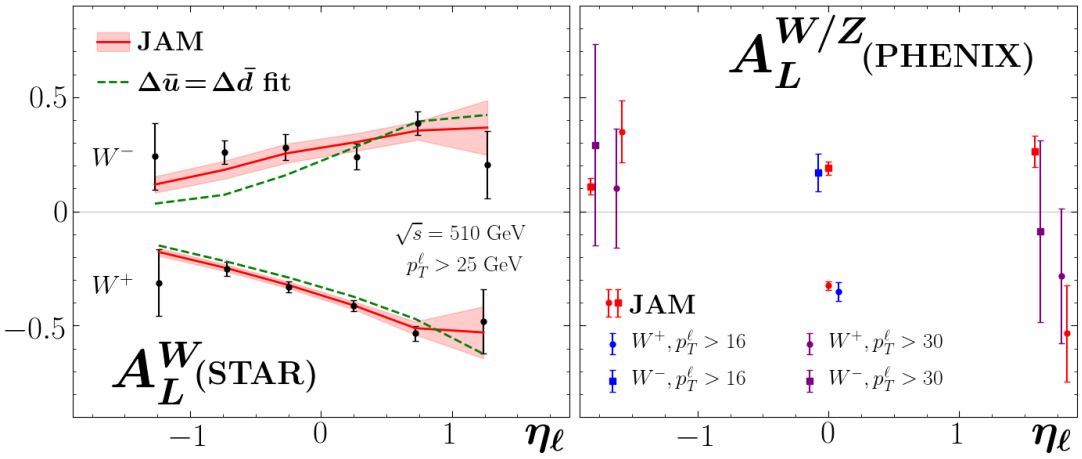

Figure 4.5: Data vs. Theory: Polarized W/Z-lepton Production. Single-spin asymmetries A_L^W and $A_L^{W/Z}$ versus lepton pseudorapidity η_ℓ. Left panel: Asymmetries from STAR [255] (black circles) at $\sqrt{s} = 510$ GeV and integrated over $p_T^\ell > 25$ GeV, compared with the full JAM fit (red solid lines and 1σ uncertainty bands) and with a fit where $\Delta\bar{u}$ is set equal to $\Delta\bar{d}$ (green dashed lines). Right panel: Asymmetries from PHENIX [256, 257] at $\sqrt{s} = 510$ GeV and integrated over $p_T^\ell > 16$ GeV (blue points) or $p_T^\ell > 30$ GeV (purple points), compared with the full JAM fit (red points). The W^+/Z asymmetries are shown with circles, while the W^-/Z asymmetries are shown with squares (we show points instead of bands due to the wide binning of the PHENIX data).

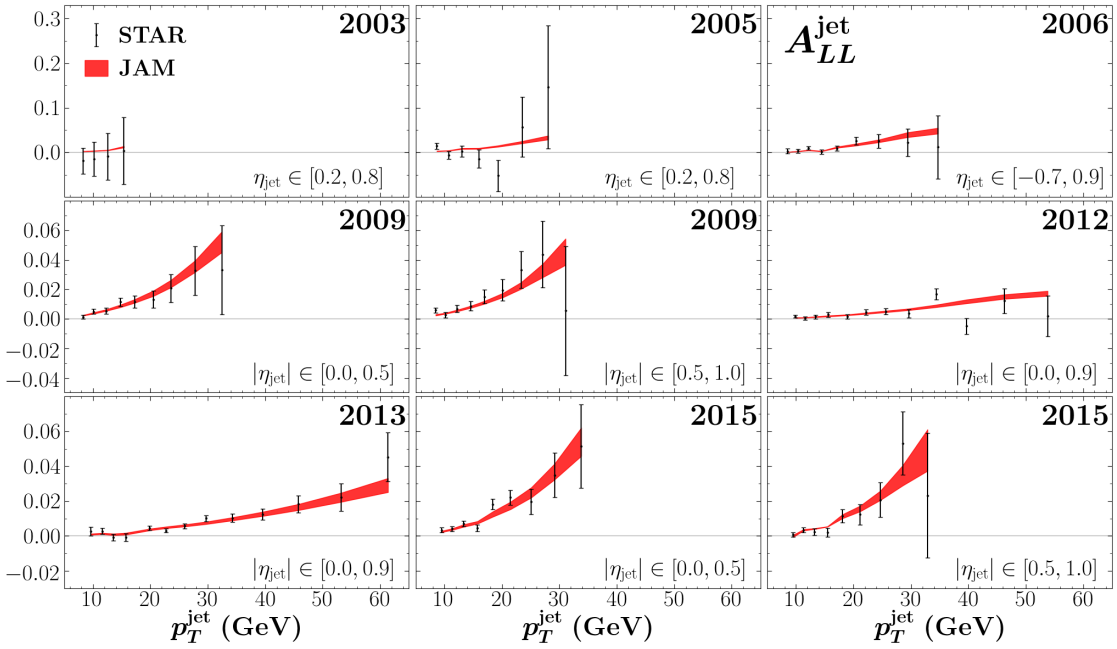

Figure 4.6: Data vs. Theory: Polarized jets. Polarized double longitudinal jet asymmetries A_{LL}^{jet} from STAR [143, 236–241] (black circles) plotted as a function of p_T^{jet} and compared to the JAM fit (red 1σ uncertainty bands). Each subplot shows the year when the data was taken and the pseudorapidity bins η_{jet} or $|\eta_{\text{jet}}|$.

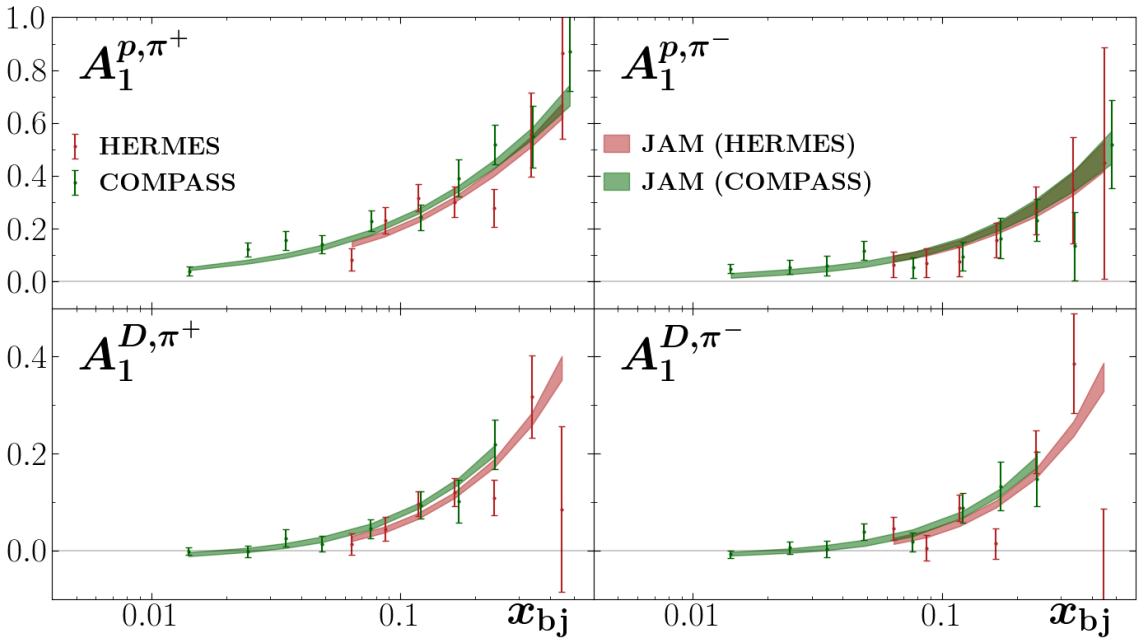

Figure 4.7: Data vs. Theory: Polarized SIDIS (pion). Polarized SIDIS asymmetries $A_1^{\pi^\pm}$ from HERMES [246] (red points) and COMPASS [247, 248] (green points) plotted as a function of $x_{\rm bj}$ and compared to the JAM fit (1σ uncertainty bands). The top left panel shows the results for the proton π^+ asymmetry, the top right for the proton π^- asymmetry, the bottom left for the deuteron π^+ asymmetry, and the bottom right for the deuteron π^- asymmetry. z_H is integrated from 0.2 to 1 for SMC and HERMES, and from 0.2 to 0.85 for COMPASS.

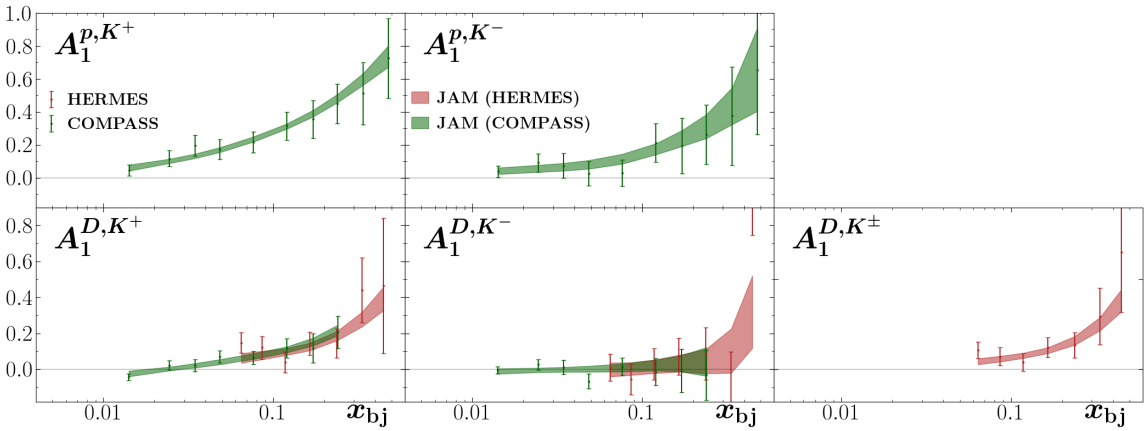

Figure 4.8: Data vs. Theory: Polarized SIDIS (kaon). Polarized SIDIS asymmetries $A_1^{K^\pm}$ from HERMES [246] (red points) and COMPASS [247,248] (green points) plotted as a function of x_{bj} and compared to the JAM fit (1σ uncertainty bands). The top left panel shows the results for the proton K^+ asymmetry, the top middle for the proton K^- asymmetry, the bottom left for the deuteron K^+ asymmetry, the bottom middle for the deuteron K^- asymmetry, and the bottom right for the deuteron $K^\pm$ asymmetry. z_H is integrated from 0.2 to 1 for SMC and HERMES, and from 0.2 to 0.85 for COMPASS.

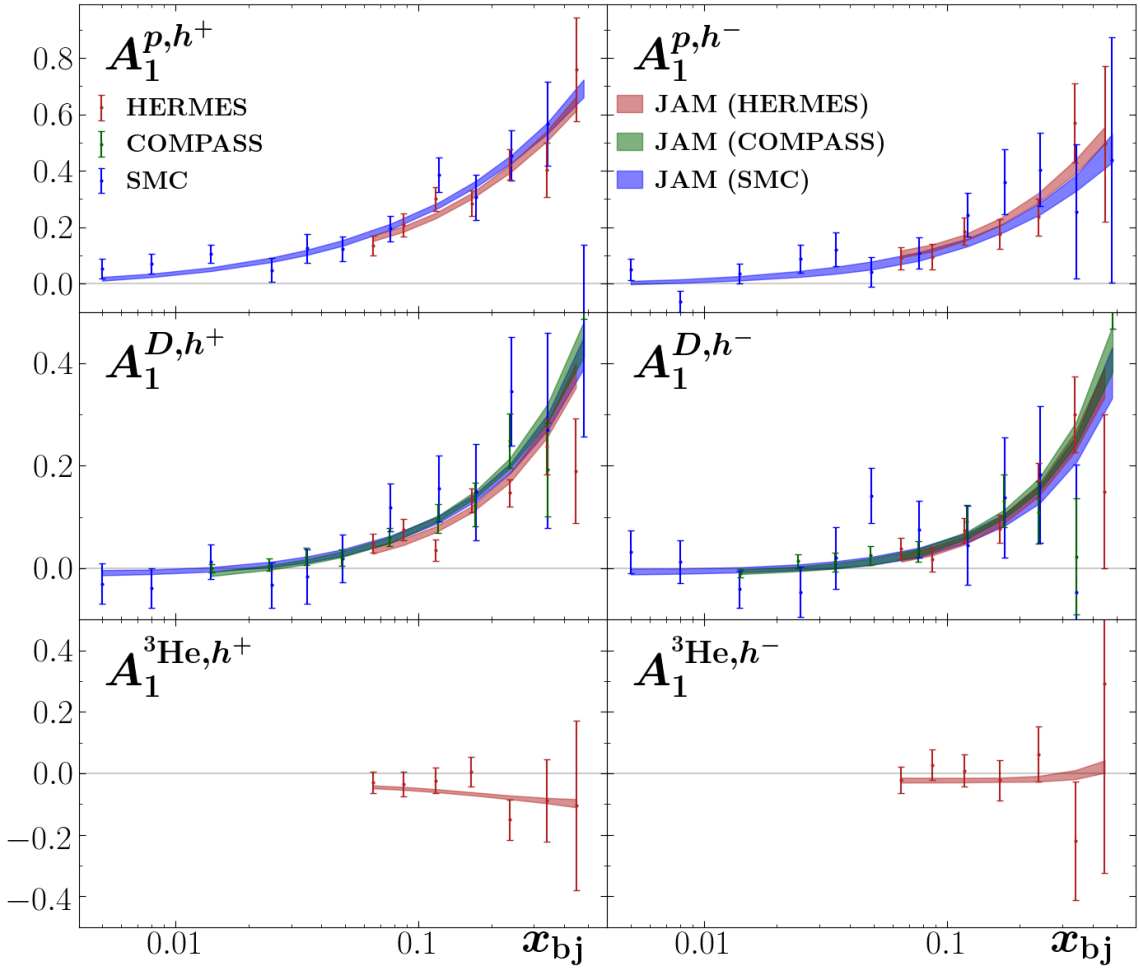

Figure 4.9: Data vs. Theory: Polarized SIDIS (hadron). Polarized SIDIS asymmetries $A_1^{h^\pm}$ from HERMES [246] (red points), COMPASS [248] (green points), and SMC [245] (blue points) plotted as a function of x_{bj} and compared to the JAM fit (1σ uncertainty bands). The top left panel shows the results for the proton h^+ asymmetry, the top right for the proton h^- asymmetry, the middle left for the deuteron h^+ asymmetry, the middle right for the deuteron h^- asymmetry, the bottom left for the helium h^+ asymmetry, and the bottom right for the helium h^- asymmetry. z_H is integrated from 0.2 to 1 for SMC and HERMES, and from 0.2 to 0.85 for COMPASS.

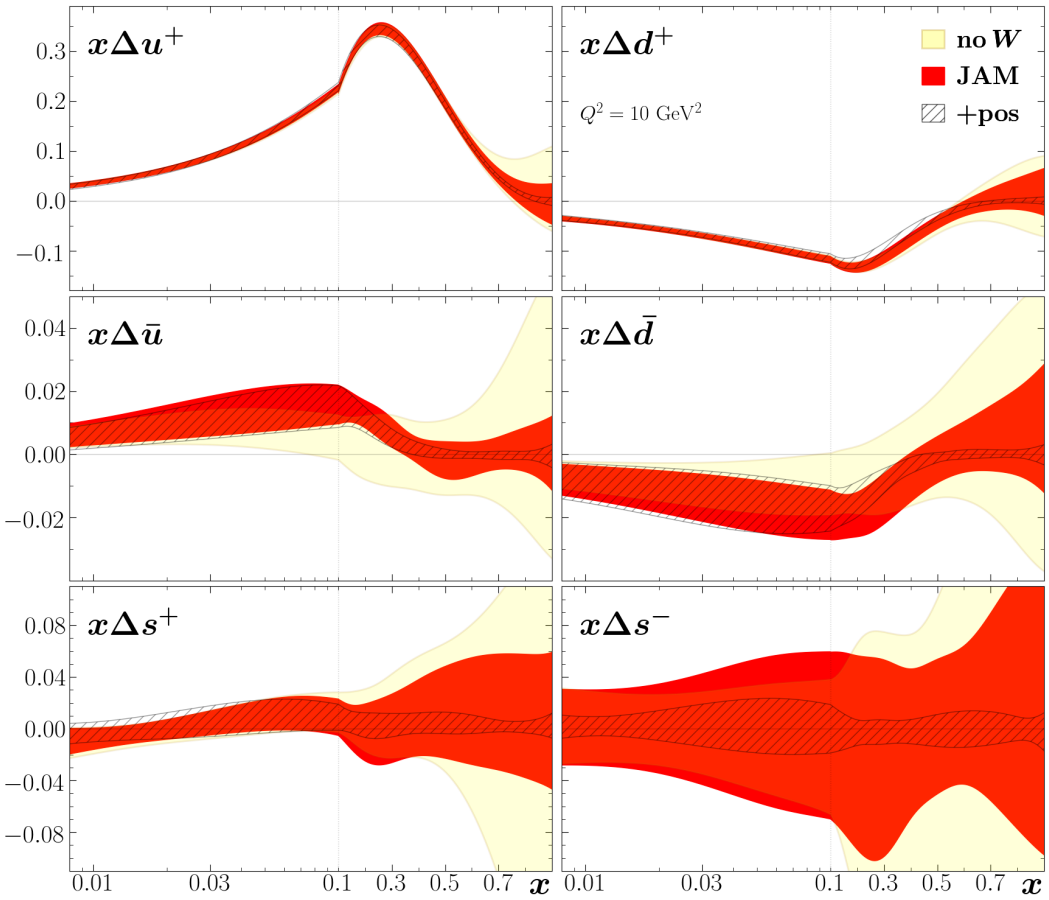

Figure 4.10: Extracted helicity PDFs. Helicity PDFs from JAM (red 1σ bands) at $Q^2 = 10$ GeV^2 compared with a fit without RHIC W/Z data (yellow band) and the result with positivity constraints (hatched band). The different panels show different flavors, including $x\Delta u^+$, $x\Delta d^+$, $x\Delta\bar{u}$, $x\Delta\bar{d}$, $x\Delta s^+$, and $x\Delta s^-$.

4.4 Extracted Helicity PDFs and Hadron FFs

This analysis [58] is based on more than 900 Monte Carlo samples, which we use to ensure the statistical convergence of the PDFs and FFs, from which the means and expectation values are then computed using Eq. (2.4). Although the spin-averaged PDFs are included in this analysis, the resulting distributions do not change significantly from what was shown in Chapter 2 and they will not be shown again here. This lack of change is due to the high precision of the unpolarized data when compared to the polarized. The resulting helicity parton densities are displayed in Fig. 4.10 at the scale $Q^2 = 10$ GeV^2, compared with the fit including positivity constraints and

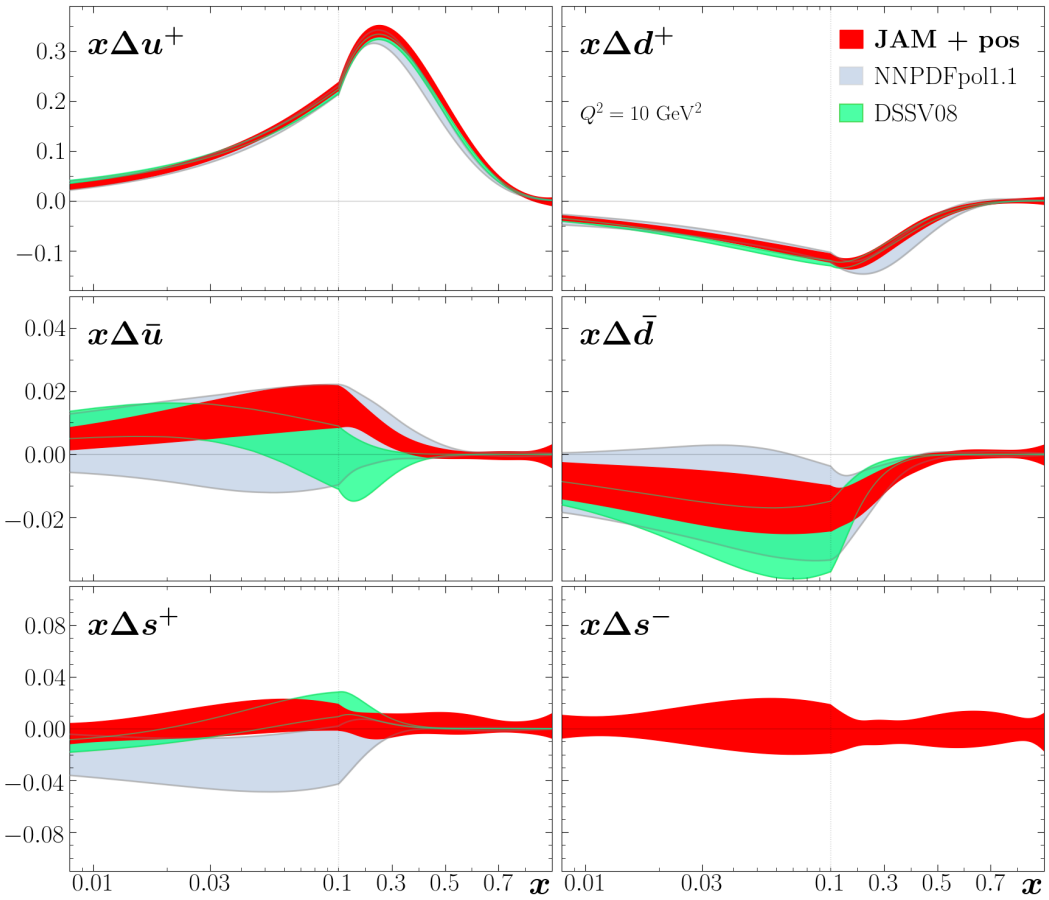

Figure 4.11: Extracted helicity PDFs compared to other groups. Helicity PDFs from JAM with positivity constraints (red 1σ bands) at $Q^2 = 10$ GeV^2 compared with the NNPDFpol1.1 [260] and DSSV08 [258] analyses. The different panels show different flavors, including $x\Delta u^+$, $x\Delta d^+$, $x\Delta\bar{u}$, $x\Delta\bar{d}$, $x\Delta s^+$, and $x\Delta s^-$. Note that $\Delta s^- = 0$ for NNPDFpol1.1 and DSSV08.

the fit without the RHIC W/Z data. For the Δu^+ and Δd^+ distributions, we find that the inclusion of the RHIC W/Z data significantly decreases the uncertainty at high $x > 0.5$, primarily due to the constraints on the antiquark distributions. The inclusion of positivity constraints decreases the uncertainties even further at high x, but does not drastically change the result for the Δu^+ and Δd^+ distributions. For the strange distributions we again see a reduction of uncertainties at large x when including the RHIC W/Z data. Positivity constraints provide a massive reduction in uncertainties due to the strange helicity being bound by the spin-averaged strange distribution. Both the Δs^+ and Δs^- distributions are consistent with zero through-

out the entire range of x, as was found in the JAM17 analysis [114]. As discussed in Section 4.2, we do not assume $\Delta s = \Delta\bar{s}$ and thus a nonzero Δs^- is possible. A nonzero Δs^- is expected from chiral symmetry breaking [278–280] from interactions with mesons (such as Λ) that accompany the Goldstone boson, and the kaon SIDIS data potentially provide sensitivity to the asymmetry. However, these data are not precise enough to distinguish the asymmetry from zero.

In Fig. 4.11, the JAM results with positivity constraints are compared to the NNPDF [260] and DSSV [258] analyses, which both also include positivity constraints, at the scale $Q^2 = 10$ GeV^2. For the Δu^+ and Δd^+ distributions we see complete agreement between the three groups. This is to be expected as these distributions are primarily constrained by the polarized DIS data which is included in all three analyses. For the Δs^+ distribution we find agreement within errors as well, although the DSSV result suggests a slightly positive result around $x \approx 0.1$ and the NNPDF result suggests a negative result below $x \approx 0.1$. Both DSSV and NNPDF assume $\Delta s = \Delta\bar{s}$ and thus $\Delta s^- = 0$. The light sea quarks will be discussed in detail in Section 4.4.1.

Given the phenomenological interest in the behavior of $\Delta q/q$ as $x \to 1$ [281–283], our simultaneous extraction of unpolarized and helicity PDFs including the W-lepton data provides the most reliable determination of the ratios to date. The results for the light quark polarization ratios $\Delta q/q$ are shown in Fig. 4.12. As is well known, the polarization is positive for u quarks and negative for d quarks. Without positivity constraints, a nonzero ratio can be extracted for u up to $x \approx 0.8$ and for d up to $x \approx 0.6$. With positivity constraints this is extended further up to $x \approx 0.85$ and $x \approx 0.7$ for u and d, respectively. The inclusion of the latest W data also provides unambiguous signs for $\Delta\bar{u}$ and $\Delta\bar{d}$, leading to a positive $\Delta\bar{u}/\bar{u}$ and a negative $\Delta\bar{d}/\bar{d}$, matching their quark counterparts. Without (with) positivity constraints, $\Delta\bar{u}/\bar{u}$ can be distinguished from zero up to values of $x \approx 0.35$ ($x \approx 0.40$), while for $\Delta\bar{d}/\bar{d}$ it can be distinguished from zero up to $x \approx 0.35$ ($x \approx 0.45$). The inclusion of positivity constraints makes little difference below $x = 0.1$ for both the quarks and antiquarks but reduces the uncertainties at larger x.

In Fig. 4.13 we show the truncated integral $\int_{0.01}^{1} \mathrm{d}x\, \Delta q(x)$ at $Q^2 = 4$ GeV^2 for the light quarks and antiquarks before and after including the RHIC W data. The lower limit of integration is chosen to roughly match the lower x limit of the data (see Fig. 4.1). We see an improvement in the uncertainties for Δu^+ and Δd^+ of

roughly 30%, while $\Delta\bar{u}$ sees an improvement of roughly 40% and $\Delta\bar{d}$ an improvement of roughly 20%. While prior to the inclusion of the RHIC W data the sign of the $\Delta\bar{u}$ contribution to the proton spin was consistent with zero, after including these data we find that $\Delta\bar{u}$ provides a small but unambiguously positive contribution to the proton spin. Prior to the inclusion of the RHIC data, the result for $\Delta\bar{u}$ depends heavily on the inclusion of positivity constraints. When the RHIC data are included, however, this dependence is significantly reduced, allowing for an extraction that is far less dependent on theoretical assumptions.

Our truncated moments for Δu^+ and Δd^+, with values 0.771(25) and $-0.363(23)$, respectively, are only slightly smaller in magnitude than the corresponding full moments from lattice QCD calculations, which find 0.864(16) for Δu^+ and $-0.426(16)$ for Δd^+ [284]. This comparison suggests that the contributions to the light quark moments below $x = 0.01$ are small. We find nonzero truncated moments for $\Delta\bar{d}$ and, for the first time, $\Delta\bar{u}$, which was found to be consistent with zero in both the NNPDF and DSSV analyses. Interestingly, the contributions from $\Delta\bar{u}$ [+0.044(17)] and $\Delta\bar{d}$ [$-0.056(24)$] approximately cancel in the sum.

4.4.1 Helicity Sea Asymmetry

The polarized antiquark asymmetry, shown in Fig. 4.14, is clearly nonzero for $0.01 < x < 0.3$. The inclusion of positivity constraints significantly reduces the uncertainties at $x \gtrsim 0.2$, since the polarized sea quarks are restricted by the size of the unpolarized sea quarks. In contrast to the final result, the results without the RHIC W data are consistent with zero for $x \gtrsim 0.07$, illustrating the importance of the STAR W data for the extraction of the polarized antiquark asymmetry in the intermediate-x region.

We furthermore compare our results to the asymmetries from the DSSV [258] and NNPDF [260] groups. The DSSV fit [258] is qualitatively similar to our result without the RHIC W data, as expected, with significantly smaller errors at high x due to the inclusion of positivity constraints. The differences in the shape at $x \lesssim 0.1$ may have two causes. The first is the propagation of FF uncertainties, which are not propagated in the DSSV analysis due to the lack of a simultaneous fit. The second is the parameterization flexibility of the helicity PDFs. While DSSV has 20 free parameters for the helicity PDFs, this analysis has 32, which may impact the results. These factors may also explain why the two analyses have similar error bands (except at high x where positivity constraints explain the difference), despite the DSSV analysis not including

any W-lepton production data. The NNPDF result [260], on the other hand, shows only a slight deviation from zero at high values of x. This is consistent with this fit taking the DSSV result [258] as the prior for $\Delta\bar{u}$ and $\Delta\bar{d}$, but with 4σ uncertainty, and including the older STAR W data [254] in their reweighting analysis. These factors also explain why the NNPDF analysis has larger errors throughout most of the range of x. At high x the differences are again explained by positivity constraints. Our analysis was thus the first extraction of a nonzero polarized antiquark asymmetry in the intermediate-x region, where model calculations generally indicate the largest effects [250–252].

In Fig. 4.15, we compare our result directly to several models. The meson cloud result is taken directly from [249] at $Q^2 = 2.5$ GeV^2 and the statistical model result is taken directly from [252] at $Q^2 = 10$ GeV^2. For the chiral soliton model, we use the relation $\Delta\bar{d} - \Delta\bar{u} = -2x^{0.12}(\bar{d} - \bar{u})$ between the helicity and spin-averaged asymmetries at $Q^2 = 4$ GeV^2 derived in [251] and use the spin-averaged asymmetry from the simultaneous extraction. While the results are plotted at slightly different scales, we do not expect evolution to change the qualitative findings here. We see complete agreement with both the statistical and chiral soliton models, which both also peak around $x \approx 0.1$. On the other hand, the meson cloud model predicts a much smaller asymmetry with the opposite sign. This may indicate that the meson cloud alone is not sufficient to explain the observed asymmetry.

4.4.2 Gluon Helicity

The extracted gluon helicity is shown in Fig. 4.16, both with and without positivity constraints. As was found in Ref. [244], without positivity constraints there are two distinct gluon solutions with opposite signs. When positivity constraints are enforced, the negative solution is eliminated due to the fact that $|\Delta g| > g$ at high x for the negative solution. Table 4.4 shows the χ^2_{red} for the polarized jet data, which is most sensitive to the gluon helicity, for both the positive ($+\Delta g$) and negative ($-\Delta g$) solutions. We see that the χ^2_{red} is similar for both solutions, and there is no basis to ignore the negative solution based on the data alone.

As discussed in Ref. [244], the positive solution results in positive contributions from the gg and qg channels to the asymmetry. The negative solution results in a large positive gg contribution and a large negative qg contribution, which mostly cancel and end up giving a positive asymmetry similar to the positive solution. Solutions with

experiment	ref.	$N_{\rm dat}$	$\chi^2_{\rm red}(+\Delta g)$	$\chi^2_{\rm red}(-\Delta g)$
STAR	[143]	4	0.22	0.22
STAR	[236]	9	0.40	0.51
STAR	[237]	7	1.51	1.49
STAR	[238]	18	0.58	0.51
STAR	[239]	12	1.60	1.42
STAR	[241]	18	0.84	0.90
STAR	[240]	13	0.72	1.09
PHENIX	[242]	2	0.37	0.38
Total		83	**0.84**	**0.88**

Table 4.4: χ^2 table: Polarized jet data with positive and negative Δg. Summary of the $\chi^2_{\rm red}$ values for the polarized jet production data for both the positive and negative Δg solutions.

Δg near zero would lead to an asymmetry that is too small, thus leading to the two distinct solutions. When positivity is enforced, the resulting single solution is similar to the positive solution. All of these findings are consistent with those in Ref. [244]. Recently, it was found that charged pion production with large transverse momentum in polarized SIDIS measured at Jefferson Lab could potentially distinguish between the two solutions [285]. Future measurements at the EIC may also help [286], as well as lattice QCD measurements [287–289].

In Fig. 4.16 the result with positivity constraints is compared to the NNPDF [260] and DSSV [259] results, which also include positivity constraints. All three analyses agree within errors. We note that the DSSV19 analysis also includes dijet production data from STAR [290, 291] and single-hadron production data from PHENIX [292–294], which provide extra constraints in the $x \approx 0.1$ region. These datasets will be included in future analyses. Further constraints on the gluon helicity at low x will be provided by the EIC [38, 39].

4.4.3 Fragmentation Functions

In Fig. 4.17, Fig. 4.18, and Fig. 4.19, we compare our extracted fragmentation functions for pions, kaons, and unidentified hadrons to those of the JAM20-SIDIS [116] and MAPFF1.0 [295] analyses. The MAPFF1.0 analysis did not include unidentified hadrons, and so is not shown in Fig. 4.19. Generally we see agreement between the two JAM analyses within errors, and significantly smaller errors on the JAM20-SIDIS

Analysis	$\boldsymbol{\pi^+}$	$\boldsymbol{K^+}$
JAM	$u = \bar{d}, d = \bar{u}, s = \bar{s}$	$u, \bar{s}, s, d = \bar{d} = \bar{u}$
JAM20-SIDIS [116]	$u = \bar{d}, d = \bar{u} = s = \bar{s}$	$u, \bar{s}, s = d = \bar{d} = \bar{u}$
MAPFF1.0 [295]	$u, \bar{d}, d = \bar{u}, s = \bar{s}$	$u, \bar{s}, s = \bar{u}, d = \bar{d}$

Table 4.5: Fragmentation function comparison. Assumptions made about the pion and kaon FFs in this analysis, the JAM20-SIDIS analysis [116], and the MAPFF1.0 [295] analysis. All three analyses fit the $c = \bar{c}$, $b = \bar{b}$, and gluon FFs in addition to the FFs listed in the table.

analysis, which will be discussed below. We note that the two analyses include the same SIA and unpolarized SIDIS data, while this analysis includes polarized SIDIS data in addition. The polarized SIDIS data does little to constrain the FFs, however, due to its large errors compared to the unpolarized data. Comparing to the MAPFF1.0 analysis, we see some significant differences, such as for $D_u^{\pi^+}$ and $D_{\bar{d}}^{\pi^+}$, with the errors of MAPFF1.0 generally being larger than those in the JAM20-SIDIS analysis.

One major difference between the MAPFF1.0 and JAM analyses is that MAPFF1.0 places a cut of $W^2 > 25$ GeV^2 on the COMPASS SIDIS data, significantly reducing the number of points from 498 to 314 for the pion and 494 to 312 for the kaon. But the most likely cause of the differences are the different parametrization choices, summarized in Table 4.5 (the parametrization for this analysis is also summarized in Section 4.2). For the pion, MAPFF1.0 has the most flexible parametrization, choosing to have $D_u^{\pi^+} \neq D_{\bar{d}}^{\pi^+}$ and leaving four free functions for the three lightest (anti)quarks, compared to three from this analysis, which sets $D_u^{\pi^+} = D_{\bar{d}}^{\pi^+}$, and two from the JAM20-SIDIS analysis, which further sets $D_s^{\pi^+} = D_d^{\pi^+}$. For the kaon, both this analysis and MAPFF1.0 have four free functions for the three lightest (anti)quarks, but assume different relations between the four unfavored light quarks, with this analysis assuming $D_d^{K^+} = D_{\bar{d}}^{K^+} = D_{\bar{u}}^{K^+}$ and MAPFF1.0 instead assuming $D_s^{K^+} = D_{\bar{u}}^{K^+}$ and $D_d^{K^+} = D_{\bar{d}}^{K^+}$. The JAM20-SIDIS analysis reduces the number of free light quark functions to three by further assuming that $D_d^{K^+} = D_{\bar{d}}^{K^+} = D_{\bar{u}}^{K^+} = D_s^{K^+}$. While the parametrizations for $D^{\delta h^+}$ are the same between the two JAM analyses, the unidentified hadron FFs $D^{h^+} = D^{\pi^+} + D^{K^+} + D^{\delta h^+}$ depends upon the pion and kaon FFs and thus are also affected by these parametrization choices. This comparison shows the importance of parametrization choices when it comes to extracting FFs.

The MAPFF1.0 analysis also performs the analysis at NNLO accuracy, while the two JAM analyses are at NLO accuracy, but this does not appear to make a large difference above $z = 0.2$ (see Figures 1 and 2 in [295]).

In Fig. 4.20 we compare directly the different FFs, including the "residual" $\delta h^+ \equiv h^+ - \pi^+ - K^+$, for all flavors. The conclusions generally match those found in the JAM20-SIDIS analysis [116]. Namely, we find that the pion FF dominates for most channels, as expected due to the fact that the pion is the lightest hadron in the QCD spectrum. The most notable exception to this is $\bar{s} \to K^+$. We also find that δh^+ can be sizeable and comparable to K^+, indicating that the production of hadrons heavier than the kaon can be sizeable. This is particularly true in the d channel, which can be explained by proton production. The s channel is also large due to the assumption that $D_s^{\delta h^+} = D_d^{\delta h^+}$. Finally, we find that $D_b^{\delta h^+}$ is larger than $D_c^{\delta h^+}$, which can be explained by the fact that the bottom quark has a larger phase space to transition into heavier hadrons when compared to the lighter charm quark.

4.5 Summary

In this chapter we performed the first simultaneous extraction of helicity PDFs, spin-averaged PDFs, and pion, kaon, and unidentified hadron FFs. This analysis included the latest polarized W-lepton production data from the STAR collaboration, and led to a data-driven extraction of a nonzero polarized sea asymmetry at intermediate x. This extracted asymmetry is consistent with predictions from the statistical [252] and chiral soliton models [251], but disagrees with the meson cloud model [249]. This analysis also provided a self-consistent extraction of the light quark polarizations $\Delta q/q$, and found $\Delta\bar{u}/u$ ($\Delta\bar{d}/d$) is positive (negative) like its quark counterpart $\Delta u/u$ ($\Delta d/d$). From the extracted helicity PDFs, the truncated contributions to the proton's spin were calculated and compared to those of other extractions and lattice QCD. This analysis found a nonzero contribution to the proton's spin from $\Delta\bar{u}$, and that the contributions from the quarks below $x \approx 0.01$ must be small in order to match lattice QCD calculations [284].

It was also found that the extraction of the gluon helicity is heavily dependent upon the imposition of positivity constraints, whose application beyond leading order has been debated recently in the literature [267, 268]. Without positivity constraints, the data permit both a positive and negative solution. When the constraints are

enforced, only the positive solution remains. These findings are in agreement with those of another recent JAM analysis [244]. In the future, SIDIS data from Jefferson Lab [285], jet production data from the EIC [286], and information from lattice QCD [287–289] could help resolve this puzzle. Finally the extracted pion, kaon, and unidentified hadron FFs were compared to those of other recent extractions. The comparison demonstrated the importance of parametrization choices particularly when it comes to FFs.

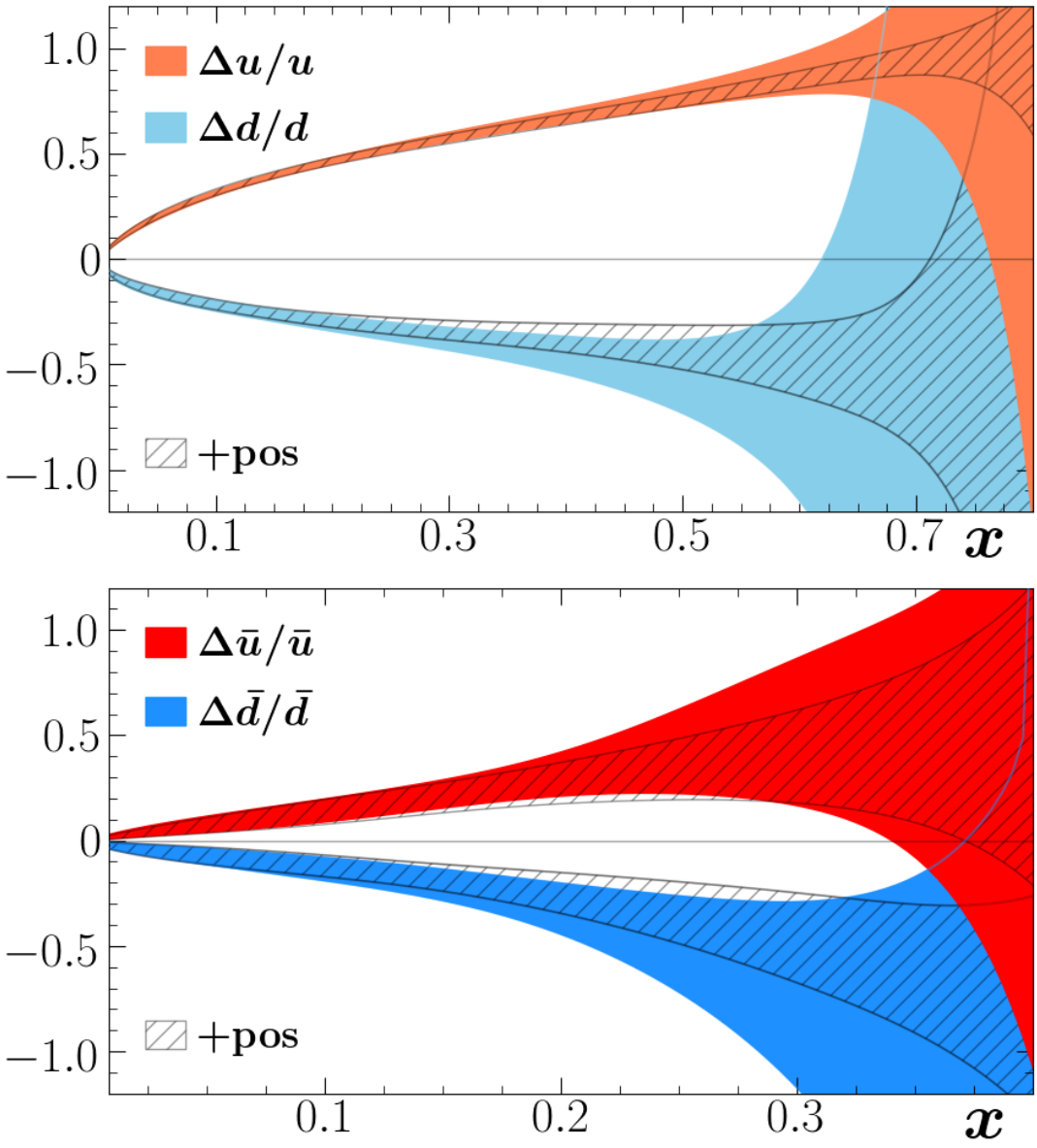

Figure 4.12: Light quark polarizations. The ratios $\Delta q/q$ are shown at $Q^2 = 10$ GeV^2. The top panel shows u and d (coral and skyblue 1σ bands) while the bottom panel shows $\bar{u}$ and $\bar{d}$ (red and blue 1σ bands). The results are compared to those with positivity constraints (hatched bands).

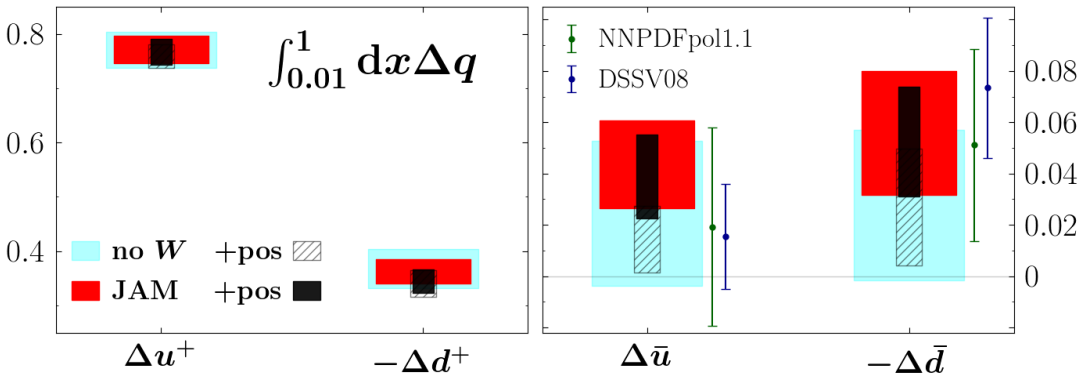

Figure 4.13: Truncated moments of helicity distributions. Truncated integrals $\int_{0.01}^{1} \mathrm{d}x\, \Delta q(x)$ at $Q^2 = 4$ GeV^2 for Δu^+, $-\Delta d^+$, $\Delta \bar{u}$ and $-\Delta \bar{d}$ from this analysis (red rectangles) compared to the fit without the RHIC W/Z data (cyan) and with positivity constraints (small hatched squares without RHIC and black squares with RHIC). For the antiquarks, NNPDFpol1.1 (green points) and DSSV08 (blue points) are also shown. The vertical height of the bands represents 1σ uncertainty.

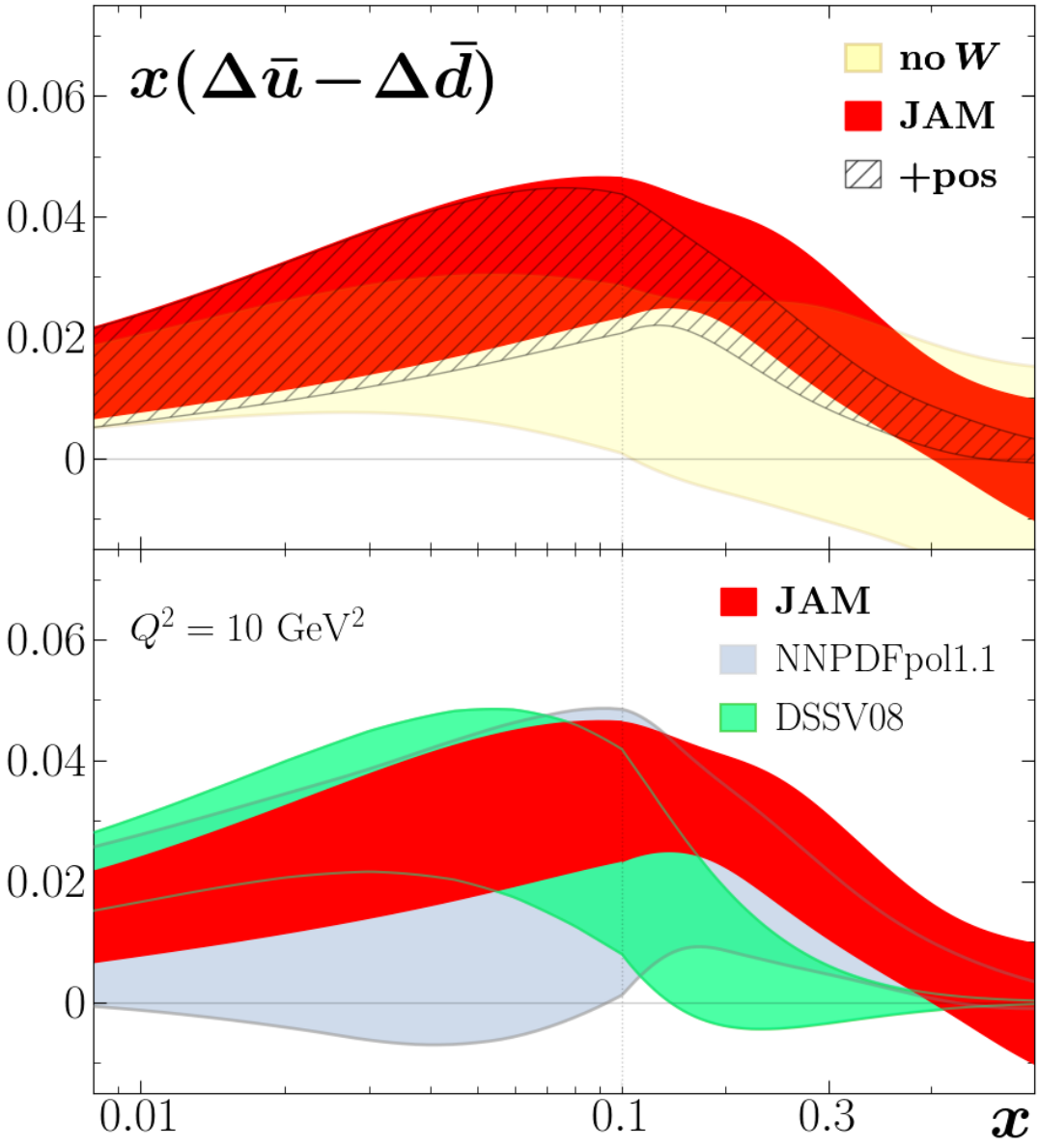

Figure 4.14: Extracted helicity sea asymmetry. Helicity sea quark asymmetry $x(\Delta\bar{u}-\Delta\bar{d})$ from JAM (red 1σ bands) at $Q^2 = 10$ GeV^2. In the top panel it is compared with the fit without RHIC W/Z data (yellow band) and the result with positivity constraints (hatched band). In the bottom panel it is compared to the NNPDFpol1.1 [260] and DSSV08 [258] analyses.

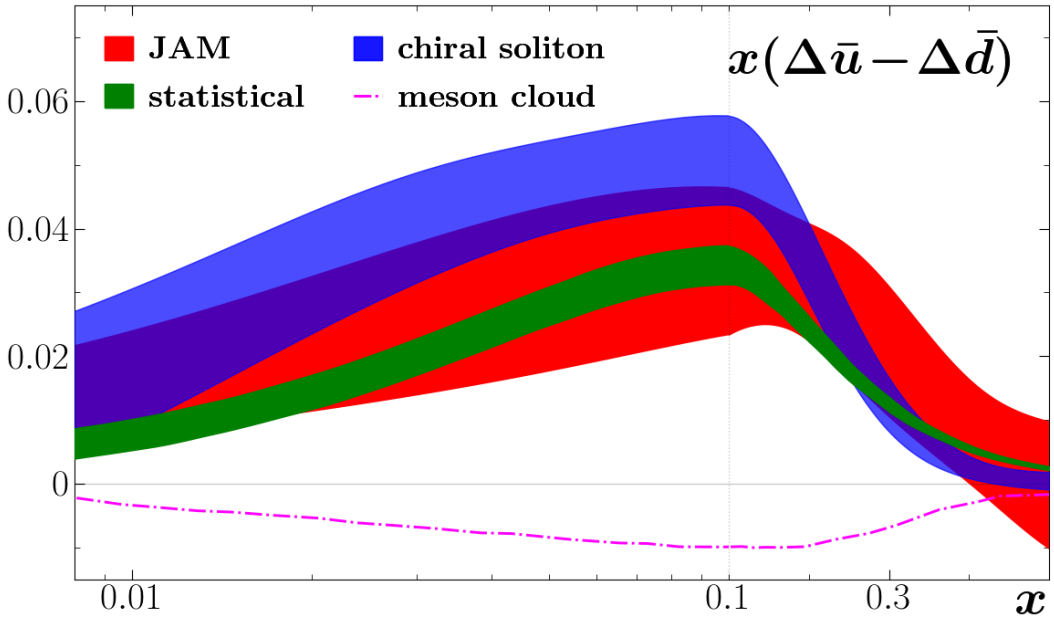

Figure 4.15: Extracted helicity sea asymmetry. Polarized sea quark asymmetry $x(\Delta\bar{u} - \Delta\bar{d})$ from JAM (red 1σ band) at $Q^2 = 10$ GeV^2 compared with the statistical model [252] at $Q^2 = 10$ GeV^2 (green 1σ band), the chiral soliton model [251] at $Q^2 = 4$ GeV^2 (blue 1σ band), and the meson cloud model [249] at $Q^2 = 2.5$ GeV^2 (dot-dash magenta line).

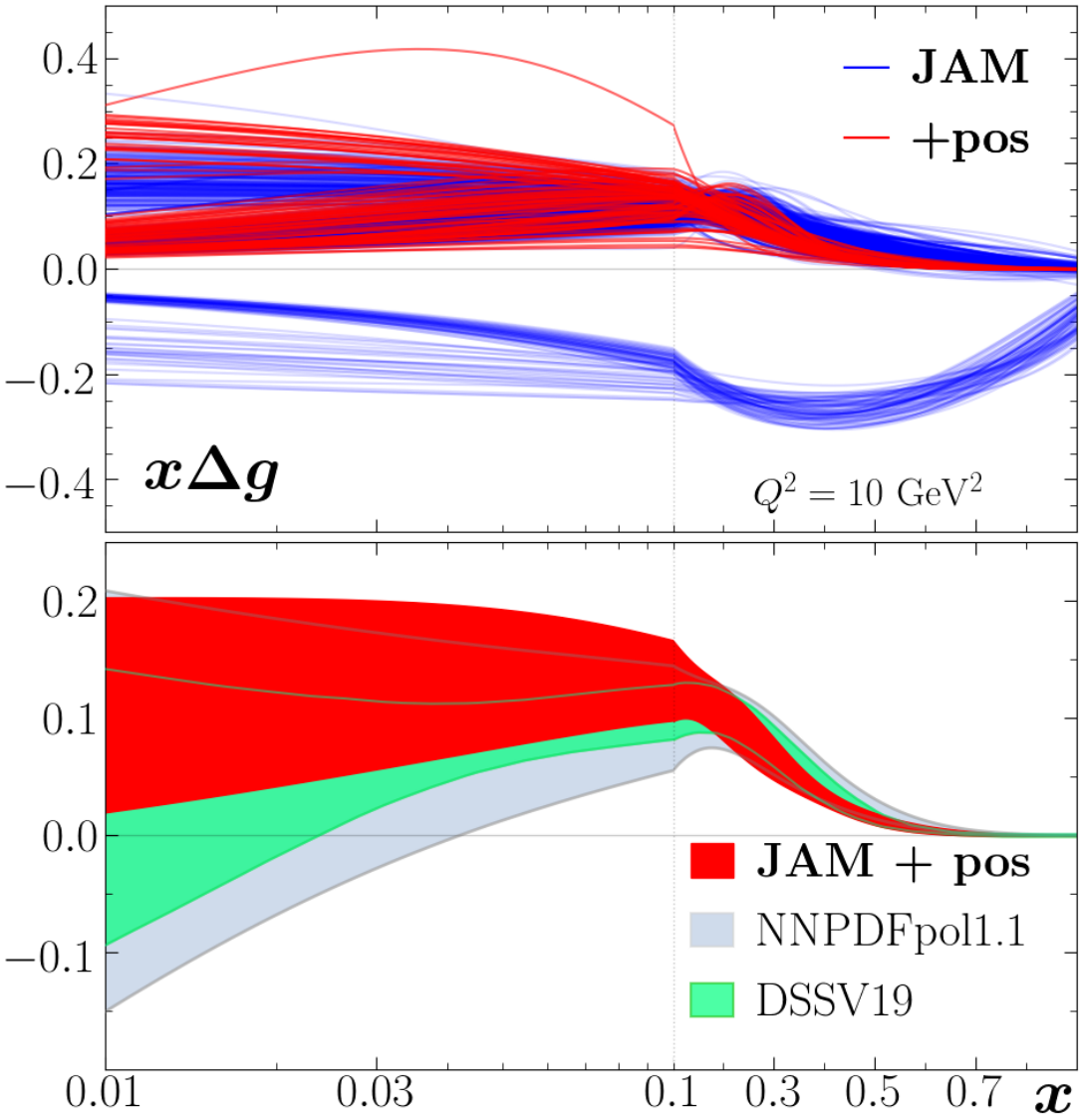

Figure 4.16: Extracted gluon helicity. Helicity gluon distribution $x\Delta g$ from JAM at $Q^2 = 10$ GeV^2. The top panel shows the replicas from the base fit (blue lines) and the replicas from the fit with positivity constraints (red lines). The bottom panel compares the result with positivity constraints (red 1σ band) with those of NNPDFpol1.1 [260] and DSSV19 [259].

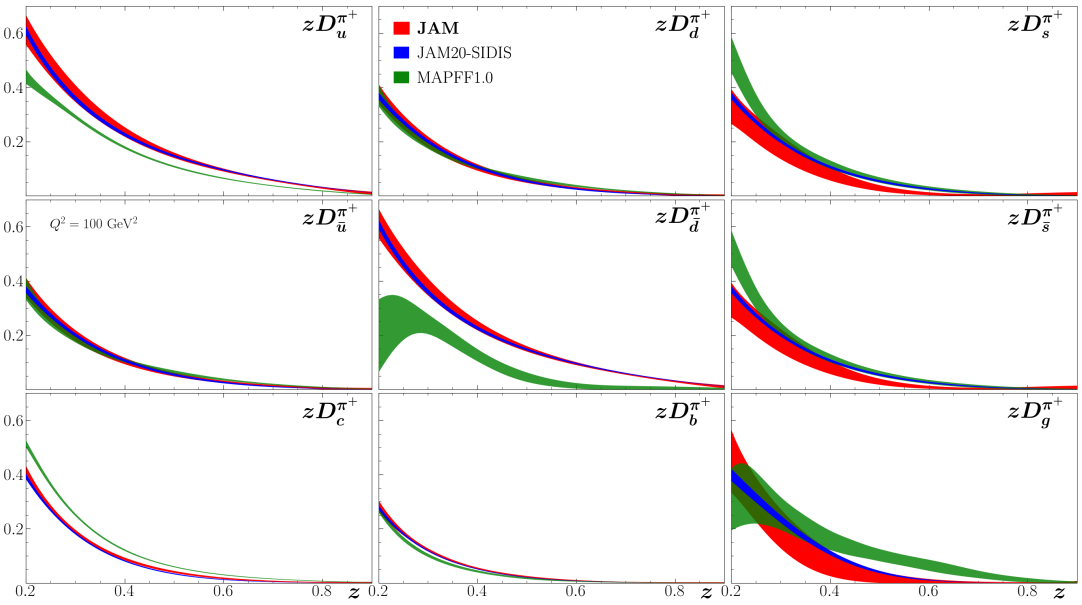

Figure 4.17: Extracted pion FFs. Pion FFs from this analysis (red 1σ bands) at $Q^2 = 100$ GeV^2 compared with the JAM20-SIDIS [116] (blue bands) and MAPFF1.0 [295] (green bands) analyses. The different panels show different flavors, and the FFs are multiplied by z.

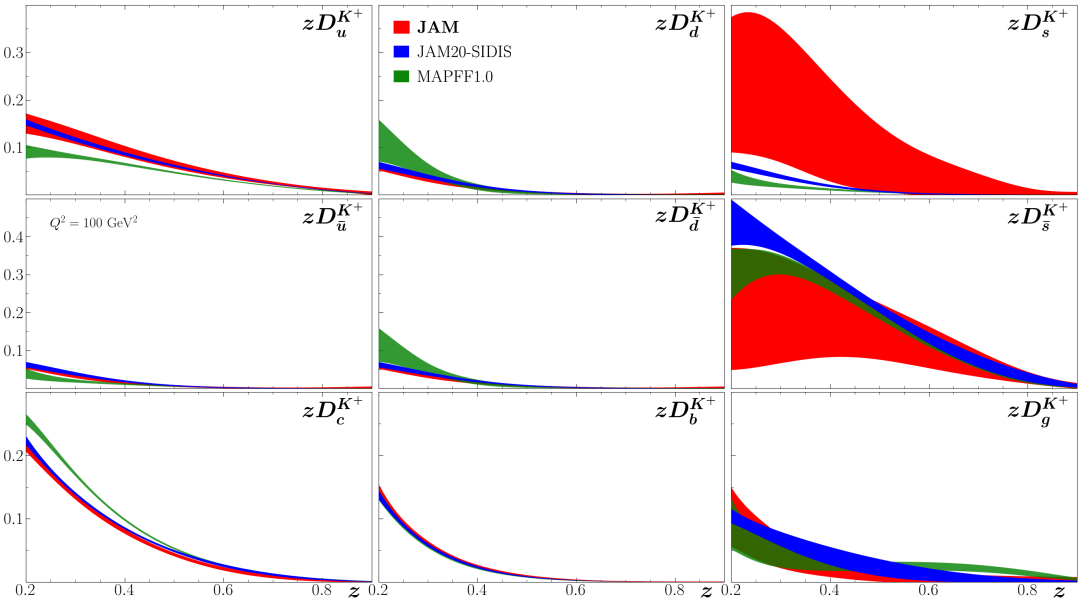

Figure 4.18: Extracted kaon FFs. Kaon FFs from this analysis (red 1σ bands) at $Q^2 = 100$ GeV^2 compared with the JAM20-SIDIS [116] (blue bands) and MAPFF1.0 [295] (green bands) analyses. The different panels show different flavors, and the FFs are multiplied by z.

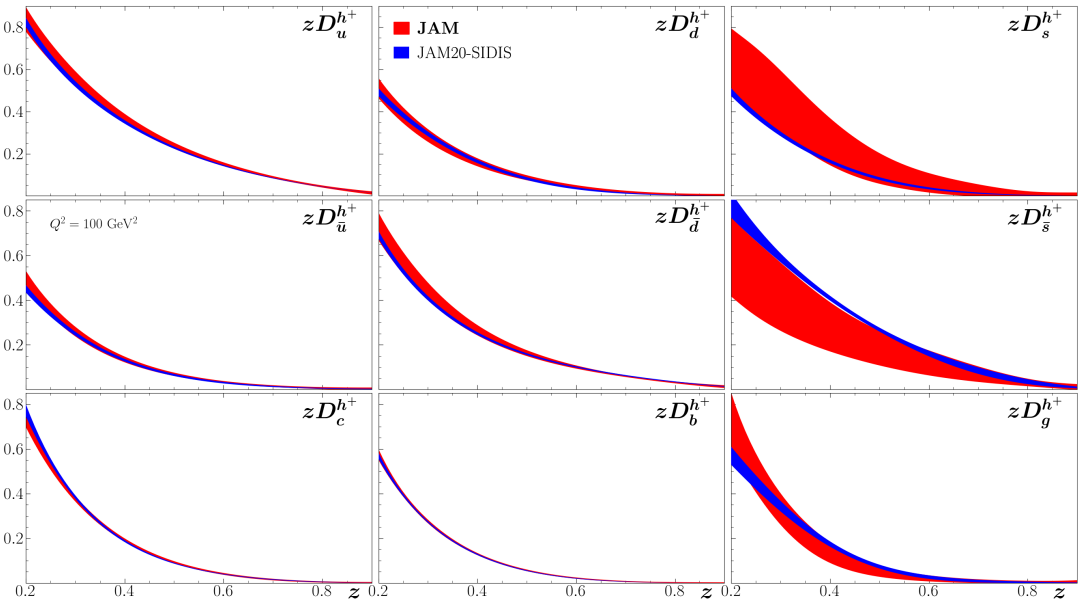

Figure 4.19: Extracted hadron FFs. Hadron FFs from this analysis (red 1σ bands) at $Q^2 = 100$ GeV^2 compared with the JAM20-SIDIS [116] (blue bands) analysis. The different panels show different flavors, and the FFs are multiplied by z.

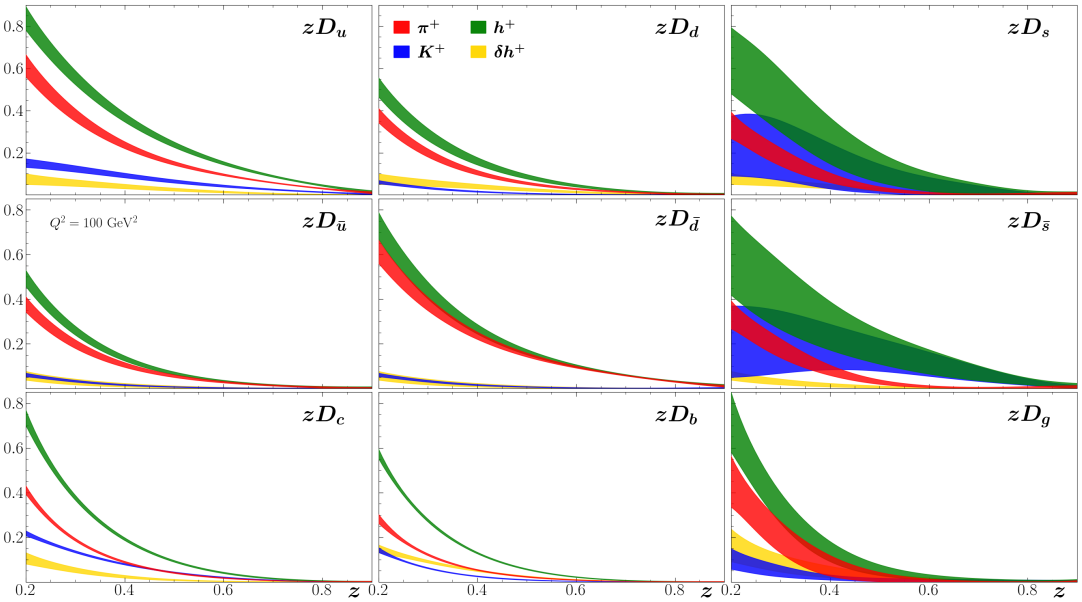

Figure 4.20: All extracted FFs. FFs from this analysis at $Q^2 = 100$ GeV^2 for the pion (red bands), kaon (blue bands), hadron (green bands), and the residual defined as $\delta h^+ \equiv h^+ - \pi^+ - K^+$ (gold bands) . The different panels show different flavors, and the FFs are multiplied by z.

CHAPTER 5

DI-HADRON PRODUCTION AND TRANSVERSITY PARTON DISTRIBUTION FUNCTIONS

In this chapter we present the details of a simultaneous global QCD analysis of $\pi^+\pi^-$ DiFFs and transversity PDFs. The general details of such an analysis were discussed in Section 2.1. We discuss two types of DiFFs: the "unpolarized DiFF"[1] D_1 , and the Interference Fragmentation Function (IFF) $H_1^\sphericalangle$. We emphasize that the IFF is also a DiFF, so that the term DiFF can refer to both D_1 and $H_1^\sphericalangle$, but we will use the term IFF to distinguish $H_1^\sphericalangle$ from D_1. The IFF was first suggested in Ref. [296] and is sensitive to the transverse polarization of the fragmenting quark. As the transversity PDF is a chiral-odd function, in equations for cross sections it must always appear alongside another chiral-odd function. In the case of di-hadron production in SIDIS and pp, that chiral-odd function is the IFF. This allows an extraction of transversity within a collinear framework, in contrast to the TMD framework used in the "JAM3D" analyses [33, 34]. Such a collinear extraction has been performed previously by Radici, Bacchetta and coworkers from the Pavia group [31, 32][2], and results from lattice QCD also provide information on the proton's transversity [35–37, 297]. However, the results from these three approaches currently cannot be reconciled.

[1]Since we only consider $\pi^+\pi^-$ DiFFs in this analysis, the notations D_1 and $H_1^\sphericalangle$ will always refer to the $\pi^+\pi^-$ DiFFs and IFFs, respectively.

[2]As we will be frequently referring to the 2018 analysis by Radici, Bacchetta et al., we will abbreviate it as RB18 from now on.

In Ref. [59] we will propose a new definition for the unintegrated DiFFs that is compatible with the probability interpretation of collinear DiFFs. We will also derive the LO evolution equations with this new definition and compare our findings with previous results in the literature. In Ref. [60] we will use this new definition and perform the first simultaneous global QCD analysis of $\pi^+\pi^-$ DiFFs, IFFs, and transversity PDFs. We improve upon previous collinear DiFF analyses [31, 32] by including new data from Belle on the SIA cross section, new data from STAR on pp collisions, and additional binnings of the SIA and SIDIS asymmetry data. We also investigate in more detail how to use event generators (in this case PYTHIA) to constrain D_1. We will show results for the DiFFs, IFF, and transversity PDFs h_1^q. Finally, we will show results for the tensor charges

$$\delta u \equiv \int_0^1 \mathrm{d}x (h_1^u - h_1^{\bar{u}}), \tag{5.1a}$$

$$\delta d \equiv \int_0^1 \mathrm{d}x (h_1^d - h_1^{\bar{d}}), \tag{5.1b}$$

and the isovector tensor charge

$$g_T \equiv \delta u - \delta d, \tag{5.2}$$

and compare to previous DiFF extractions, extractions in the TMD framework, as well as lattice QCD. The tensor charge is a fundamental property of the proton similar to the vector, axial, and scalar charges. It is relevant for QCD phenomenology [31–33, 298–302], ab initio studies in lattice QCD [35, 36, 297, 303, 304], and beyond the Standard Model physics [305–308].

5.1 Di-Hadron Fragmentation Functions

In this section we start with a new definition of the field-theoretic correlator for the fragmentation into a di-hadron pair and derive the resulting DiFFs and evolution equations. In this section we generalize to arbitrary hadron pairs h_1h_2 with masses M_1 and M_2.

5.1.1 Definition

We begin by defining the kinematics of hadron pair production. Consider a parton with momentum k that fragments into two detected hadrons h_1 and h_2 with momenta P_1 and P_2, respectively, with $P_1^2 = M_1^2$ and $P_2^2 = M_2^2$. The total momentum of the di-hadron pair is $P_h = P_1 + P_2$ and its relative momentum is $R = (P_1 - P_2)/2$. The invariant mass of the di-hadron pair is $M_h^2 = P_h^2$. The hadrons carry fractions $z_1 = P_1^-/k^-$ and $z_2 = P_2^-/k^-$ of the quark's minus-lightcone momentum. We will find it convenient to use the variables $z \equiv z_1 + z_2$ and $\zeta \equiv (z_1 - z_2)/z$. Working in a reference frame where P_h has no transverse component ($\vec{P}_{h\perp} = 0$), the momenta P_1 and P_2 can be written as

$$P_1 = \left(\frac{M_1^2 + \vec{R}_T^2}{(1+\zeta)P_h^-}, \frac{1+\zeta}{2} P_h^-, \vec{R}_T \right), \qquad P_2 = \left(\frac{M_2^2 + \vec{R}_T^2}{(1-\zeta)P_h^-}, \frac{1-\zeta}{2} P_h^-, -\vec{R}_T \right).$$

Note that one readily finds

$$\vec{R}_T^2 = \frac{1-\zeta^2}{4} M_h^2 - \frac{1-\zeta}{2} M_1^2 - \frac{1+\zeta}{2} M_2^2, \tag{5.3}$$

which implies the lower bound on M_h,

$$M_h^2 \geq \frac{2}{1+\zeta} M_1^2 + \frac{2}{1-\zeta} M_2^2.$$

We define the terms fully unintegrated DiFFs (uDiFFs) as those that depend upon $(z, \zeta, \vec{R}_T, \vec{k}_T)$ which, more precisely, means the scalars $(z, \zeta, \vec{R}_T^2, \vec{k}_T \cdot \vec{R}_T, \vec{k}_T^2)$. Upon integrating over $\vec{k}_T$, we have "extended" DiFFs (extDiFFs) that depend upon $(z, \zeta, \vec{R}_T^2)$, or alternatively (z, ζ, M_h^2). Subsequent integration over $\vec{R}_T$ leads to collinear DiFFs (cDiFFs) that depend upon (z, ζ), or alternatively (z_1, z_2). These functions give the probability density for a parton i to fragment into, say, hadron h_1 carrying the fraction z_1 of the parton's longitudinal momentum if hadron h_2 carries the momentum fraction z_2. Integrating the extDiFF over ζ leads to DiFFs that depend upon (z, M_h^2). These are the objects that have been frequently used in the transversity-related phenomenology of di-hadron production, and they will be referred to simply as DiFFs.

We define the correlator for fragmentation into a di-hadron pair as

$$\begin{aligned}\Delta_{ij}^{h_1h_2/q}(z,\vec{k}_T;P_1,P_2) = \frac{1}{16\pi^3 z^2(1-\zeta^2)}\frac{1}{N_c}\sum\!\!\!\!\!\!\int_X \frac{\mathrm{d}\xi^+\mathrm{d}^2\vec{\xi}_T}{(2\pi)^3}\, e^{i(k^-\xi^+-\vec{k}_T\cdot\vec{\xi}_T)} \\ \times\ \langle 0|\mathcal{W}_1(\infty,\xi)\psi_{q,i}(\xi^+,0^-,\vec{\xi}_T)|P_1,P_2;X\rangle \\ \times\ \langle P_1,P_2;X|\bar{\psi}_{q,j}(0^+,0^-,\vec{0}_T)\mathcal{W}_2(0,\infty)|0\rangle\,, \end{aligned} \tag{5.4}$$

where $\sum\!\!\!\!\!\!\int_X \equiv \sum_X \int \mathrm{d}^3\vec{P}_X/((2\pi)^3 2P_X^0)$, ψ_q are quark fields, and $\mathcal{W}_1, \mathcal{W}_2$ are the Wilson lines that ensure SU(3) color gauge invariance. They explicitly take the form

$$\mathcal{W}_1(\infty,\xi) = [\infty^+,0^-,\vec{\infty}_T;\infty^+,0^-,\vec{\xi}_T]\times[\infty^+,0^-,\vec{\xi}_T;\xi^+,0^-,\vec{\xi}_T]\,, \tag{5.5}$$

$$\mathcal{W}_2(0,\infty) = [0^+,0^-,\vec{0}_T;\infty^+,0^-,\vec{0}_T]\times[\infty^+,0^-,\vec{0}_T;\infty^+,0^-,\vec{\infty}_T]\,, \tag{5.6}$$

where $[a^+,a^-,\vec{a}_T;b^+,b^-,\vec{b}_T] \equiv \mathcal{P}\exp\Big[-ig\int_a^b ds_\mu A^\mu(s)\Big]$, with A^μ a gluon field. The two twist-2 Dirac traces that are relavent for this study can be parameterized in terms of uDiFFs [309]:

$$\mathrm{Tr}\left[\Delta^{h_1h_2/q}(z,\vec{k}_T;P_1,P_2)\gamma^-\right] = D_1^{h_1h_2/q}(z,\zeta,\vec{R}_T^2,\vec{k}_T\cdot\vec{R}_T,\vec{k}_T^2)\,, \tag{5.7}$$

$$\begin{aligned}\mathrm{Tr}\left[\Delta^{h_1h_2/q}(z,\vec{k}_T;P_1,P_2)i\sigma^{i-}\gamma_5\right] = -\frac{\epsilon_T^{ij}R_T^j}{M_h}H_1^{\sphericalangle' h_1h_2/q}(z,\zeta,\vec{R}_T^2,\vec{k}_T\cdot\vec{R}_T,\vec{k}_T^2) \\ -\frac{\epsilon_T^{ij}k_T^j}{M_h}H_1^{\perp\, h_1h_2/q}(z,\zeta,\vec{R}_T^2,\vec{k}_T\cdot\vec{R}_T,\vec{k}_T^2)\,. \end{aligned} \tag{5.8}$$

The tensor $\epsilon_T^{ij} \equiv \epsilon^{-+ij}$ is the "transverse" Levi-Civita tensor, and we use the convention $\epsilon_T^{12} = +1$. The extDiFFs are found by carrying out the $z^2\int \mathrm{d}^2\vec{k}_T$ integral on both sides of Eqs. (5.7)–(5.8), which leaves two structures:

$$z^2\int \mathrm{d}^2\vec{k}_T\mathrm{Tr}\left[\Delta^{h_1h_2/q}(z,\vec{k}_T;P_1,P_2)\gamma^-\right] = D_1^{h_1h_2/q}(z,\zeta,M_h^2)\,, \tag{5.9}$$

$$z^2\int \mathrm{d}^2\vec{k}_T\mathrm{Tr}\left[\Delta^{h_1h_2/q}(z,\vec{k}_T;P_1,P_2)i\sigma^{i-}\gamma_5\right] = -\frac{\epsilon_T^{ij}R_T^j}{M_h}H_1^{\sphericalangle\, h_1h_2/q}(z,\zeta,M_h^2)\,, \tag{5.10}$$

where, due to Eq. (5.3), we have chosen to write the extDiFFs as depending on M_h^2

instead of $\vec{R}_T^2$. Further integration over $\vec{R}_T$ leads to a single unpolarized cDiFF:

$$\begin{aligned} D_1^{h_1h_2/q}(z_1,z_2) = \frac{1}{16\pi^3(1-\zeta^2)}\frac{1}{N_c}\sum\!\!\!\!\!\!\int_X \int \mathrm{d}^2\vec{R}_T \int \frac{\mathrm{d}\xi^+}{2\pi}\, e^{ik^-\xi^+} \\ \times \mathrm{Tr}\Big[\langle 0|\mathcal{W}(\infty^+,\xi^+)\psi_q(\xi^+,0^-,\vec{0}_T)|P_1,P_2;X\rangle \\ \times\ \langle P_1,P_2;X|\bar{\psi}_q(0^+,0^-,\vec{0}_T)\mathcal{W}(0^+,\infty^+)|0\rangle\gamma^-\Big]\,, \end{aligned} \tag{5.11}$$

where $\mathcal{W}(a^+,b^+) = [a^+,0^-,\vec{0}_T;b^+,0^-,\vec{0}_T]$. We again emphasize that $D_1^{h_1h_2/q}(z_1,z_2)$ has the aforementioned definition of a (conditional) probability density. An analogous derivation for the gluon leads to

$$\begin{aligned} D_1^{h_1h_2/g}(z_1,z_2) = \frac{z}{4\pi^3(1-\zeta^2)}\frac{1}{2P_h^-}\frac{1}{N_c^2-1}\sum\!\!\!\!\!\!\int_X \int \mathrm{d}^2\vec{R}_T \int \frac{\mathrm{d}\xi^+}{2\pi}\, e^{ik^-\xi^+} \\ \times \langle 0|\mathcal{W}_{ba}(\infty^+,\xi^+)F_a^{-i}(\xi^+,0^-,\vec{0}_T)|P_1,P_2;X\rangle \\ \times\ \langle P_1,P_2;X|F_c^{-i}(0^+,0^-,\vec{0}_T)\mathcal{W}_{cb}(0^+,\infty^+)|0\rangle\,, \end{aligned} \tag{5.12}$$

where the gluons are represented by components of the field strength tensor $F_a^{\mu\nu}$, and the Wilson lines are the same as Eqs. (5.5)–(5.6) but in the adjoint representation.

In order to define the DiFFs $D_1^{h_1h_2/q}(z,M_h^2)$, we consider the expression for the cross section $\frac{\mathrm{d}\sigma}{\mathrm{d}z\,\mathrm{d}M_h}$ for the process $e^+e^- \to (h_1h_2)X$, given by

$$\frac{\mathrm{d}\sigma}{\mathrm{d}z\,\mathrm{d}M_h} = \frac{4\pi N_c\alpha_{\mathrm{em}}^2}{3Q^2}\sum_q e_q^2 \int_{-1}^1 \mathrm{d}\zeta\,\Big[\frac{\pi}{4}z\,M_h\,(1-\zeta^2)\,D_1^{h_1h_2/q}(z,\zeta,M_h^2)\Big]\,, \tag{5.13}$$

where the sum q runs over all quark flavors and $Q=\sqrt{s}$, the COM energy of the collision. We then define

$$D_1^{h_1h_2/q}(z,M_h^2) \equiv \frac{\pi}{4}z\,M_h\int_{-1}^1 \mathrm{d}\zeta\,(1-\zeta^2)\,D_1^{h_1h_2/q}(z,\zeta,M_h^2)\,, \tag{5.14}$$

so that

$$\frac{\mathrm{d}\sigma}{\mathrm{d}z\,\mathrm{d}M_h} = \frac{4\pi\alpha_{\mathrm{em}}^2 N_c}{3Q^2}\sum_q e_q^2\,D_1^{h_1h_2/q}(z,M_h^2)\,. \tag{5.15}$$

Analogously, we define

$$H_1^{\sphericalangle,h_1h_2/q}(z,M_h^2) \equiv \frac{\pi}{4} z\, M_h \int_{-1}^{1} d\zeta\, \frac{|\vec{R}_T|}{M_h}\,(1-\zeta^2)\, H_1^{\sphericalangle,h_1h_2/q}(z,\zeta,M_h^2), \tag{5.16}$$

with the extra factor $\frac{|\vec{R}_T|}{M_h}$ coming from the $\frac{\epsilon_T^{ij} R_T^j}{M_h}$ in Eq. (5.10).

5.1.2 Evolution

In this section we discuss evolution equations for the extDiFFs. Perturbative corrections to the quark DiFF correlator (Eq. (5.4)) generate two pieces: a "homogeneous term" involving only DiFFs and a "mixing term" involving single-hadron FFs. The same is true for the gluon DiFF correlator. In Ref. [310] it was argued that, for extDiFFs, the mixing term is not divergent in the ultraviolet region and therefore does not contribute to the evolution of these functions. We have confirmed this point [59]. The evolution of $D_1^{h_1h_2/q}(z,\zeta,M_h^2)$ can be computed at the level of the correlator, with the calculation proceeding exactly as that for the single-hadron FF $D_1^{h/q}(z)$ except for the prefactor in Eq. (5.4). The final result reads

$$\begin{aligned}\frac{\partial D_1^{h_1h_2/q}(z,\zeta,M_h^2;\mu_R)}{\partial \ln \mu_R^2} = \int_z^1 \frac{dw}{w^2}\Bigg[& D_1^{h_1h_2/q}\left(\frac{z}{w},\zeta,M_h^2;\mu_R\right) P_{q\to q}(w) \\ & + D_1^{h_1h_2/g}\left(\frac{z}{w},\zeta,M_h^2;\mu_R\right) P_{q\to g}(w)\Bigg],\end{aligned} \tag{5.17}$$

where $P_{q\to q}(w)$ and $P_{q\to g}$ are the standard unpolarized time-like splitting functions. An analogous expression exists for $H_1^{\sphericalangle,h_1h_2/q}(z,\zeta,M_h^2)$, with the splitting functions adjusted to those of the transversity distribution. We call the reader's attention to the $1/w^2$ factor that enters the integrand, whereas for the single-hadron case this factor is just $1/w$. The $1/w^2$ for the di-hadron case can be traced back to the prefactor in the definition Eq. (5.4), as it eventually leads to the first factor in brackets in the first line of Eq. (5.17). That factor is simply 1, whereas for the $D_1^{h/q}(z)$ case it is $z/(z/w) = w$.

For the phenomenology we need the evolution equation for $D_1^{h_1h_2/q}(z,M_h)$. Switching out the variables z_1 and z_2 for z and ζ, and defining $D_1^{h_1h_2/q}(z,M_h^2)$ as in Eq. (5.14),

we find

$$\frac{\partial D_1^{h_1h_2/q}(z,M_h^2;\mu_R)}{\partial \ln \mu_R^2} = \int_z^1 \frac{dw}{w}\left[D_1^{h_1h_2/q}\left(\frac{z}{w},M_h^2;\mu_R\right)P_{q\to q}(w) + D_1^{h_1h_2/g}\left(\frac{z}{w},M_h^2;\mu_R\right)P_{q\to g}(w)\right], \tag{5.18}$$

with an analogous expression for $H_1^{\sphericalangle,h_1h_2/q}(z,M_h^2)$. We also mention that after integrating Eq. (5.17) over $\vec{R}_T$, our result for the evolution of the homogeneous term of $D_1^{h_1h_2/q}(z_1,z_2)$ fully agrees with previous results derived in the literature [311–313]. An analogous evolution equation holds for the gluon extDiFF $D_1^{h_1h_2/g}(z,M_h^2;\mu_R)$.

5.2 Processes

In this analysis we include di-hadron production in SIA, SIDIS, and proton-proton collisions. In the subsections below we summarize each of these processes and the relevant equations.

5.2.1 Di-Hadron Production in SIA

The basics of the SIA process were summarized in Subsection 2.2.5, with the LO diagram shown in Fig. 2.6. We consider two observables, the first where a single di-hadron pair, with invariant mass M_h and combined fractional energy z, is detected and the observable is sensitive to $D_1(z,M_h^2)$. At LO in α_s, the cross section for this process is given by Eq. (5.15). The factorization scale of the process is $Q=\sqrt{s}$. Due to the symmetry of $\pi^+\pi^-$ pairs (see 5.3.1) there are 5 free DiFFs (D_1^u, D_1^s, D_1^c, D_1^b, and D_1^g) with this observable only capable of constraining one of them (which we choose to be D_1^u). Thus we have supplemented this data with data gathered from an event generator such as PYTHIA, which will be discussed in Subsection 5.4.1.

The second observable involves the detection of two di-hadron pairs, with the second having invariant mass $\overline{M}_h$ and combined fractional energy $\bar{z}$. By measuring an azimuthal correlation of two hadron pairs detected in opposite hemispheres, one obtains an observable that is sensitive to the IFF $H_1^{\sphericalangle}(z,M_h^2)$. At LO in α_s, this

modulation (known as the Artru-Collins asymmetry [314]) is given by [315]:

$$a_{12R}(z, M_h, \bar{z}, \overline{M}_h) = \frac{\sin^2\theta \sum_q e_q^2 \, H_1^{\sphericalangle,q}(z, M_h^2) H_1^{\sphericalangle,\bar{q}}(\bar{z}, \overline{M}_h^2)}{(1+\cos^2\theta)\sum_q e_q^2 \, D_1^q(z, M_h^2) D_1^{\bar{q}}(\bar{z}, \overline{M}_h^2)}\,, \tag{5.19}$$

where the sum q runs over all quark flavors and θ is defined as the polar angle between the beam axis and the reference axis in the COM system. As discussed in the introduction to this chapter, there is only one free IFF, which we choose to be $H_1^{\sphericalangle,u}$. Thus only one observable is needed for the IFFs. However, one can see from Eq. (5.36) that a_{12R} is proportional to $[H_1^{\sphericalangle,u}]^2$. Thus the asymmetry cannot uniquely determine the sign of $H_1^{\sphericalangle,u}$, and one must be applied by hand. Knowing the sign of the SIDIS data (see the following section) and assuming that the transversity up quark distribution must be positive (as is found in all phenomenological, model, and lattice QCD studies) leads to the conclusion that $H_1^{\sphericalangle,u}$ must be negative. This is also supported by model calculations for $H_1^{\sphericalangle}$ [316], and so in the analysis we choose $H_1^{\sphericalangle,u}$ to be negative.

5.2.2 Di-Hadron Production in SIDIS

The basics of the SIDIS process were summarized in Section 2.2.4, with the LO diagram shown in Fig. 2.5. In this case, the target is transversely polarized and the outgoing hadrons are a $\pi^+\pi^-$ pair with invariant mass M_h. We denote the incoming lepton momentum, virtual-photon momentum, and relative hadron momentum by $\mathbf{l}$, $\mathbf{q}$, and $\mathbf{R}$, respectively. With these variables, experimentalists define the azimuthal angle ϕ_R by [317]:

$$\phi_R = \frac{(\mathbf{q}\times\mathbf{l})\cdot\mathbf{R}}{|(\mathbf{q}\times\mathbf{l})\cdot\mathbf{R}|}\arccos\frac{(\mathbf{q}\times\mathbf{l})\cdot(\mathbf{q}\times\mathbf{R})}{|(\mathbf{q}\times\mathbf{l})||(\mathbf{q}\times\mathbf{R})|}. \tag{5.20}$$

They also define ϕ_S, the azimuthal angle of the initial nucleon spin. COMPASS [318] defines the angle ϕ_{RS} as $\phi_{RS} \equiv \phi_R + \phi_S - \pi$, while HERMES [319] defines it as $\phi_{RS} \equiv \phi_R + \phi_S$.

At leading twist and LO in α_s, the asymmetry can be written as [320, 321]:

$$A_{UT}^{\text{COMPASS}} = \frac{\sum_q e_q^2 \, h_1^q(x) \, H_1^{\sphericalangle,q}(z, M_h^2)}{\sum_q e_q^2 \, f_1^q(x) \, D_1^q(z, M_h^2)}\,, \tag{5.21}$$

for COMPASS and

$$A_{UT}^{\text{HERMES}} = -D_{nn}(y)\frac{\sum_q e_q^2\, h_1^q(x)\, H_1^{\sphericalangle,q}(z, M_h^2)}{\sum_q e_q^2\, f_1^q(x)\, D_1^q(z, M_h^2)}, \tag{5.22}$$

for HERMES. The DiFFs and transversity PDFs depend on the factorization scale, which is taken to be $\mu_F = Q$ for the SIDIS process. For HERMES, the transverse-spin-transfer coefficient $D_{nn}(y)$ is given by

$$D_{nn}(y) = \frac{1-y}{1-y+\frac{y^2}{2}}. \tag{5.23}$$

The opposite signs between the asymmetries for COMPASS and HERMES come from the different definitions for the angle ϕ_{RS}, which differ by a phase π as discussed above. The extra factor D_{nn} for HERMES comes from the fact that HERMES calculated the asymmetries in the lepton-nucleon system, while COMPASS calculated them in the photon-nucleon system. Ultimately these different definitions do not impede the possibility of fitting both measurements simultaneously and simply require the extra factors for the HERMES asymmetry.

Both COMPASS and HERMES provide the asymmetry for proton targets, while COMPASS also provides it for a deuteron target. The neutron PDFs are calculated using the proton PDFs with the switch $u \leftrightarrow d$ that is derived from isospin symmetry. Nuclear effects are neglected, and the deuterium PDFs are simply taken as the sum of the proton and neutron PDFs. Due to the symmetry of the $\pi^+\pi^-$ IFFs (see Eq. (5.36)), one has in the numerator for a proton target

$$\sum_q e_q^2\, h_1^q\, H_1^{\sphericalangle,q} = \frac{1}{9}H_1^{\sphericalangle,u}[4(h_1^u - h_1^{\bar{u}}) - (h_1^d - h_1^{\bar{d}})] = \frac{1}{9}H_1^{\sphericalangle,u}[4h_1^{u_v} - h_1^{d_v}], \tag{5.24}$$

while for the deuteron one has

$$\sum_q e_q^2\, h_1^{q,D}\, H_1^{\sphericalangle,q} = \frac{1}{3}H_1^{\sphericalangle,u}[h_1^u - h_1^{\bar{u}} + (h_1^d - h_1^{\bar{d}})] = \frac{1}{3}H_1^{\sphericalangle,u}[h_1^{u_v} + h_1^{d_v}], \tag{5.25}$$

where $h_1^{q,D}$ are the deuteron transversity PDFs. From this one immediately sees that the SIDIS data is sensitive only to the valence distributions, with the proton asymmetry more sensitive to u_v than d_v and the deuteron asymmetry equally sensitive to both.

5.2.3 Di-Hadron Production in Proton-Proton Collisions

For di-hadron production in proton-proton collisions, the process is similar to the jet production process shown in Fig. 2.4, with the outgoing jet replaced by a di-hadron pair with invariant mass M_h. The cross sections are written as differential in η, the pseudo-rapidity of the di-hadron pair, P_{hT}, the transverse momentum of the di-hadron pair, and M_h. The asymmetry is defined as

$$A_{UT}^{\sin(\phi_R-\phi_S)} \equiv \frac{\frac{\mathrm{d}\sigma_{UT}}{\mathrm{d}\eta \mathrm{d}M_h \mathrm{d}P_{hT}}}{\frac{\mathrm{d}\sigma_{UU}}{\mathrm{d}\eta \mathrm{d}M_h \mathrm{d}P_{hT}}}\,, \tag{5.26}$$

where σ_{UT} is the cross section with one proton transversely polarized and σ_{UU} the cross section with both protons unpolarized. They are given by [322]

$$\begin{aligned}\frac{\mathrm{d}\sigma_{UT}}{\mathrm{d}\eta \mathrm{d}M_h \mathrm{d}P_{hT}} &= 2P_{hT} \sum_{abcd} \int_{x_a^{\min}}^1 \mathrm{d}x_a \int_{x_b^{\min}}^1 \frac{\mathrm{d}x_b}{z} \qquad (5.27)\\ &\times f_1^a(x_a,\mu_F)\, h_1^b(x_b,\mu_F) \frac{\mathrm{d}\Delta\hat{\sigma}_{ab^\uparrow\to c^\uparrow d}}{\mathrm{d}\hat{t}}(x_a,x_b,s,\eta,P_{hT},\mu_R,\mu_F)\, H_1^{\sphericalangle,c}(z,M_h^2,\mu_F)\,,\\ \frac{\mathrm{d}\sigma_{UU}}{\mathrm{d}\eta \mathrm{d}M_h \mathrm{d}P_{hT}} &= 2P_{hT} \sum_{abcd} \int_{x_a^{\min}}^1 \mathrm{d}x_a \int_{x_b^{\min}}^1 \frac{\mathrm{d}x_b}{z} \qquad (5.28)\\ &\times f_1^a(x_a,\mu_F)\, f_1^b(x_b,\mu_F) \frac{\mathrm{d}\hat{\sigma}_{ab\to cd}}{\mathrm{d}\hat{t}}(x_a,x_b,s,\eta,P_{hT},\mu_R,\mu_F)\, D_1^c(z,M_h^2,\mu_F)\,,\end{aligned}$$

where x_a (x_b) is the momentum fraction of the parton coming from h_A (h_B) with momentum P_A (P_B). The sum over the parton flavors $abcd$ runs over all partonic channels that can contribute to the process. The limits on the integration are given by

$$x_a^{\min} = \frac{P_{hT}e^{\eta}}{\sqrt{s} - P_{hT}e^{-\eta}}\,, \qquad x_b^{\min} = \frac{x_a P_{hT}e^{-\eta}}{x_a\sqrt{s} - P_{hT}e^{\eta}}\,, \tag{5.29}$$

and one also has the relation

$$z = \frac{P_{hT}}{\sqrt{s}}\left(\frac{x_a e^{-\eta} + x_b e^{\eta}}{x_a x_b}\right). \tag{5.30}$$

The hard (perturbative) partonic cross sections $\frac{\mathrm{d}\Delta\hat{\sigma}_{ab^\uparrow\to c^\uparrow d}}{\mathrm{d}\hat{t}}$ and $\frac{\mathrm{d}\hat{\sigma}_{ab\to cd}}{\mathrm{d}\hat{t}}$ are written as differential in $\hat{t}$, where $\hat{t} = tx_a/z$ and t is given by $t = (P_A - P_B)^2$. They are taken from the appendix of Ref. [322]. (We note that there is an error there on all of

the quark channels, which should have the opposite sign.) The renormalization and factorization scales μ_R and μ_F are taken to be equal to P_{hT}. The STAR collaboration defines h_A to be the transversely polarized proton rather than h_B as we do here. This difference in the definition of the asymmetry is taken care of by taking $\eta \to -\eta$ when evaluating the asymmetries provided by STAR. For later use, we define the x value that the transversity PDFs are approximately evaluated at as

$$x \approx \frac{P_{hT}}{\sqrt{s}} e^{\eta}. \tag{5.31}$$

5.3 Parametrizations

In this section we discuss the parameterization of the DiFFs, IFFs, and transversity PDFs. All functions are evolved using the DGLAP evolution equation (see Subsection 5.1.2) with the appropriate splitting functions, as detailed in Section 2.3.

5.3.1 DiFF Parametrization

The $\pi^+\pi^-$ DiFFs have the advantage of being highly symmetric and thus reducing the number of free functions that need to be parameterized. For D_1, one has the symmetry relations [323]

$$\begin{aligned} D_1^u = D_1^d = D_1^{\bar{u}} = D_1^{\bar{d}}, \\ D_1^s = D_1^{\bar{s}}, \quad D_1^c = D_1^{\bar{c}}, \quad D_1^b = D_1^{\bar{b}}. \end{aligned} \tag{5.32}$$

Thus, including also the gluon, there are 5 free functions to be fitted. D_1^u, D_1^s, and D_1^g are parametrized at the input scale $\mu_0 = 1$ GeV, while D_1^c and D_1^b are parametrized at $\mu_R = m_c$ and $\mu_R = m_b$, respectively. It is computationally efficient to choose a parametrization for D_1^q that can be converted into Mellin space. As such, we choose to parametrize the z dependence with a functional form on a grid of M_h, choosing for the up quark:

$$\mathbf{M}_h^u = [2m_\pi, 0.40, 0.50, 0.70, 0.75, 0.80, 0.90, 1.00, 1.20, 1.30, 1.40, 1.60, 1.80, 2.00] \text{ GeV}.$$

We note that the grid is not uniform and is instead chosen in a way to best describe the detailed structure of the e^+e^- cross section. For s, c, b and g we choose grids that

are less dense:

$$\begin{aligned}\mathbf{M}_h^s = \mathbf{M}_h^c = \mathbf{M}_h^b &= [2m_\pi, 0.50, 0.75, 1.00, 1.20, 1.60, 2.00] \text{ GeV}, \\ \mathbf{M}_h^g &= [2m_\pi, 0.70, 1.00, 1.40, 2.00] \text{ GeV}.\end{aligned}$$

We find that these grids are sufficient to describe the Belle and PYTHIA data while keeping the number of parameters to a minimum. At each value of M_h on the grid, denoted by $\mathbf{M}_h^{q,i}$, the z dependence is parametrized as:

$$D_1^q(z, \mathbf{M}_h^{q,i}) = \sum_{j=1,2,3} \frac{N_{ij}^q}{\mathcal{M}_{ij}^q} z^{\alpha_{ij}^q} (1-z)^{\beta_{ij}^q}, \tag{5.33}$$

where $\boldsymbol{a} = \{N_{ij}^q, \alpha_{ij}^q, \beta_{ij}^q\}$ is the set of parameters to be inferred, and $\mathcal{M}_{ij}^q = \mathrm{B}[\alpha_{ij}^q + 1, \beta_{ij}^q + 1]$ normalizes the function to the first moment to decorrelate the normalization and shape parameters. For the up quark, it is necessary to include the $j = 2$ and $j = 3$ parameters as it will be constrained by the Belle cross section data. The other functions will be constrained by PYTHIA-generated data (see Subsection 5.4.1) with larger errors, and so less flexibility is needed and we set all of the $j = 2, 3$ parameters to zero. This leads to $14 \times 9 = 126$ parameters for D_1^u, $7 \times 3 = 21$ parameters each for D_1^s, D_1^c, and D_1^b, and $5 \times 3 = 15$ parameters for D_1^g, for a total of 204 free parameters. This functional form is easily converted into Mellin space at each $\mathbf{M}_h^{q,i}$, and interpolation is used to retrieve the Mellin space representation at any value of M_h.

In this analysis, we also enforce the positivity bound [320]

$$D_1^q(z, M_h; \mu_R^2) > 0, \tag{5.34}$$

with $q = u, s, c, b, g$. As was done with the positivity constraints in Chapter 4, we enforce this bound approximately on each Monte Carlo replica by imposing a penalty on the χ^2 function when the bounds are violated [269]. For each replica and at each step of the χ^2 minimization we first calculate D_1^q at the input scale $\mu_0^2 = 1$ GeV^2 at 300 points in the z, M_h plane, with $0.2 < z < 1$ and $2m_\pi < M_h < 2.0$, noting that if positivity is enforced at the input scale it will automatically hold at larger scales for any function that evolves through the DGLAP equation [270, 271]. We repeat this process for u, s, c, b, and g. Any DiFFs that are negative contribute to the overall χ^2

through

$$\chi^2_{D_1<0} = N^2\Big(\sum_{z,M_h}\sum_q \Theta\big[-D_1^q(z,M_h;\mu_0^2)\big]\ |D_1^q(z,M_h;\mu_0^2)|\Big)^2. \tag{5.35}$$

The normalization N is chosen to be 3 so that the initial contribution to the χ^2 is generally $\mathcal{O}(1000)$, although the size of the initial contribution can vary significantly depending on the starting parameters. This ensures that the contribution is non-negligible but that it also does not dominate the entire χ^2 function which is otherwise $\mathcal{O}(1000)$.

5.3.2 IFF Parametrization

For $H_1^\sphericalangle$, the symmetry relations for $\pi^+\pi^-$ are [323]

$$\begin{aligned} H_1^{\sphericalangle,u} &= -H_1^{\sphericalangle,d} = -H_1^{\sphericalangle,\bar{u}} = H_1^{\sphericalangle,\bar{d}}, \\ H_1^{\sphericalangle,s} &= -H_1^{\sphericalangle,\bar{s}} = H_1^{\sphericalangle,c} = -H_1^{\sphericalangle,\bar{c}} = 0. \end{aligned} \tag{5.36}$$

The (quark) transversity cannot couple to a FF for gluons. We therefore need just a single free IFF, which we choose to be $H_1^{\sphericalangle,u}$. The IFF $H_1^{\sphericalangle,u}$ is parameterized similarly to D_1^q. Since the relevant data for $H_1^{\sphericalangle,u}$ is comparatively sparse and has larger errors, far fewer parameters are needed. Thus we are able to choose a less dense M_h grid:

$$\mathbf{M}_h^u \ = \ [2m_\pi, 0.50, 0.70, 0.85, 1.00, 1.60, 2.00]\ \text{GeV},$$

and at each value of $\mathbf{M}_h^{u,i}$, the z dependence:

$$H_1^{\sphericalangle,u}(z,\mathbf{M}_h^{u,i}) = \sum_{j=1,2}\frac{N_{ij}^q}{\mathcal{M}_{ij}^q} z^{\alpha_{ij}^q}(1-z)^{\beta_{ij}^q}, \tag{5.37}$$

at the input scale $\mu_0 = 1$ GeV. This leads to a total of $8\times 6 = 48$ free parameters for the IFF. As with D_1^q, this form is easily converted into Mellin space.

We enforce the bound [320]:

$$|H_1^q(z,M_h;\mu_R^2)| < D_1^q(z,M_h;\mu_R^2), \tag{5.38}$$

which, in practice, only needs to be applied to $q = u$. This is enforced similarly to

the positivity constraint on D_1 in Eq. (5.34), with a χ^2 penalty equal to:

$$\chi^2_{|H_1|>D_1} = N^2\Big(\sum_{\pm}\sum_{z,M_h}\Theta\big[H_1^u(z,M_h;\mu_0^2)\pm D_1^u(z,M_h;\mu_0^2)\big] \times\ \big[|H_1^u(z,M_h;\mu_0^2)\pm D_1^u(z,M_h;\mu_0^2)|\big]\Big)^2, \tag{5.39}$$

with N chosen to be 3.

5.3.3 Transversity PDF Parametrization

Our parameterization for the transversity PDFs follows that of the spin-averaged PDFs in Section 2.3 and the helicity PDFs in Section 4.2. The transversity PDFs are parameterized at the input scale $\mu_0 = 1$ GeV using the template Eq. (2.40) normalized to the first moment $\mathcal{M} = \mathrm{B}[\alpha+1,\beta+1] + \gamma\mathrm{B}[\alpha+\frac{3}{2},\beta+1] + \eta\mathrm{B}[\alpha+2,\beta+1]$. We choose to parameterize the valence distributions $h_1^{u_v}$ and $h_1^{d_v}$, as well as the antiquark distributions $h_1^{\bar{u}} = -h_1^{\bar{d}}$. Since we only have three unique observables to constrain the PDFs (proton SIDIS, deuteron SIDIS, and pp collisions), we choose the relation between the antiquarks based on predictions from the large-N_C limit [324]. We note that previous phenomenological analyses found small or negligible contributions from antiquarks [34, 298, 325–328]. We have a total of $3 \times 5 = 15$ parameters for the transversity PDFs. Combined with the 204 parameters for D_1, the 48 parameters for $H_1^{\sphericalangle}$, and 7 normalization parameters, we end up with a total of 274 fitted parameters.

As the first moment of the transversity PDFs is of great interest (see Section 5.6), we also place a limit on our parameterization to constrain the small-x behavior where no experimental data is available. The behavior of the PDFs at small-x is governed by the α parameter in Eq. (2.40). Recent theoretical works using an operator-based approach [329] have placed limits on this parameter as $x \to 0$ [330]. They find:

$$\alpha \xrightarrow[x\to 0]{} 1 - 2\sqrt{\frac{\alpha_s N_c}{2\pi}}. \tag{5.40}$$

This limit applies for both the valence quarks and antiquarks, and there is a roughly 50% uncertainty on this value from $1/N_c$ corrections and NLO corrections [331]. At the input scale, we calculate $\alpha \to 0.17$, and so limit the α parameter to the range $0.17 \pm 50\%$ for all quark flavors. Technically, this limit only applies as $x \to 0$, while our approach places a limit on the entire range of x. We find, however, that limiting

α as such has no impact on the resulting PDFs in the measured region or on our ability to describe the experimental data. Thus this simplified approach is sufficient to capture the $x \to 0$ behavior while not affecting results at moderate or high-x.

We also enforce the Soffer Bound [332] on the transversity PDFs:

$$\left|h_1^q(x,Q^2)\right| \leq \frac{1}{2}\left[f_1^q(x,\mu_R^2) + g_1^q(x,\mu_R^2)\right], \tag{5.41}$$

where $f_1^q(x,\mu_R^2)$ and $g_1^q(x,\mu_R^2)$ are the spin-averaged and helicity PDFs, respectively, taken from the analysis of Chapter 4 [58]. This is again enforced through a χ^2 penalty similar to Eq. (5.35) and Eq. (5.39)

$$\chi^2_{\rm SB} = N^2\Big(\sum_{\pm}\sum_{x}\sum_{q}\Theta\big[-F^{\pm}(x,\mu_0^2)\big]\ |F^{\pm}(x,\mu_0^2)|\Big)^2, \tag{5.42}$$

where $q = u, d, \bar{u}, \bar{d}$ and $F^{\pm}(x,\mu_0^2) \equiv \frac{1}{2}\big[f_1^q(x,\mu_0^2) + g_1^q(x,\mu_0^2)\big] \pm h_1^q(x,\mu_0^2)$. The normalization N is chosen to be 10.

5.4 Data and Quality of Fit

In this section we discuss the experimental and lattice data that enters the analysis as well as the PYTHIA generated data. This analysis must be supplemented by the PYTHIA data due to the fact that there are 5 free D_1 functions (see Subsection 5.3.1), but only one experimental observable currently available to constrain them (although D_1 appears in the denominators of the SIDIS and pp asymmetries, those observables must also constrain the transversity PDFs and so do not have the ability to significantly constrain any D_1 functions). Thus we will discuss how we generate the PYTHIA data and how it is used to constrain these functions.

5.4.1 PYTHIA Data

Our goal with the PYTHIA generated data is to be able to provide reasonable constraints on the D_1 functions for the strange, charm, bottom, and gluon in the absence of experimental data. To do this, we first generate data at the Belle energy $\sqrt{s} = 10.58$ GeV for the ratio $\sigma^q/\sigma^{\rm tot}$ with $q = s, c, b$, where $\sigma^{\rm tot} \equiv \sum_q \sigma^q$ with the sum over all quark flavors. This generated data is capable of constraining the strange, charm,

and bottom distributions. In order to get some constraint on the gluon distribution, we repeat this process at five energy scales, evenly spaced between Belle and LEP energies,

$$\sqrt{s} \;=\; [10.58, 30.73, 50.88, 71.04, 91.19] \text{ GeV}.$$

This strategy allows us to get some estimate of the gluon DiFF through evolution. We do not include the bottom quark contribution at the lowest Belle energy $\sqrt{s} = 10.58$ GeV but include it at all energies above this, due to the fact that our massless quark formalism would be highly inadequate at such low energies comparable to $2m_b$.

This leads to a total of $5 \times 3 = 15$ generated datasets from PYTHIA, with each one covering the same (z, M_h) region as the Belle data. The same cuts used on the real Belle data are used on the PYTHIA data. The cut in Eq. (5.43) is used with $\sqrt{s} = 10.58$ GeV regardless of the value of $\sqrt{s}$, as there is no benefit within our analysis of going beyond the kinematic region of the Belle data.

The PYTHIA data can be generated with arbitrarily high statistics, so we neglect the tiny statistical error. In order to have a quantifiable, non-negligible error, we generate each $(\sqrt{s}, q)$ dataset using four different tunes: "PYTHIA 6 def.", "PYTHIA 6 Aleph", "PYTHIA 6 LEP/Tev.", and "PYTHIA 8." The tunes are summarized in Table 5.1. The data that we fit are the mean values of the four tunes, and we take the variance of the tunes as an uncorrelated systematic error. We do not apply any correlated or normalization errors. We choose to fit the ratio $\sigma^q/\sigma^{\text{tot}}$ as it leads to smaller variance between the different tunes compared to taking the absolute cross sections σ^q. Taking such a ratio should also lead to cancellations of NLO thrust effects. The quality of fit to the PYTHIA data is shown in Table 5.2. The comparison between the PYTHIA generated data and the theory is shown in Appendix A.

Par	PYTHIA 6 def.	PYTHIA 6 Aleph	PYTHIA 6 LEP/Tev.
PARJ(1)	0.1	0.106	0.073
PARJ(2)	0.3	0.285	0.2
PARJ(3)	0.4	0.71	0.94
PARJ(4)	0.05	0.05	0.032
PARJ(11)	0.5	0.55	0.31
PARJ(12)	0.6	0.47	0.4
PARJ(13)	0.75	0.65	0.54
PARJ(14)	0	0.02	
PARJ(15)	0	0.04	
PARJ(16)	0	0.02	
PARJ(17)	0	0.2	
PARJ(19)	1	0.57	
PARJ(21)	0.36	0.37	0.325
PARJ(25)	1		0.63
PARJ(26)	0.4	0.27	0.12
PARJ(33)	0.8	0.8	0.8
PARJ(41)	0.3	0.4	0.5
PARJ(42)	0.58	0.796	0.6
PARJ(45)	0.5		
PARJ(46)	1		
PARJ(47)	1		
PARJ(54)	-0.05	-0.04	-0.05
PARJ(55)	-0.005	-0.0035	-0.005
PARJ(81)	0.29	0.292	0.29
PARJ(82)	1	1.57	1.65
MSTJ(11)	4	3	5
MSTJ(12)	2	3	
MSTJ(26)	2	2	2
MSTJ(45)	5		
MSTJ(107)	0	0	0

Table 5.1: Summary of PYTHIA tunes. Summary of the different PYTHIA 6 tunes used to generate the PYTHIA data. The "PYTHIA 6 def." tune uses the corresponding default parameters shown above. The "PYTHIA 6 Aleph" and "PYTHIA 6 LEP/Tev." tunes modify those parameters as shown in their respective columns (if the row is blank then the default value is used). For the "PYTHIA 8" tune we use all default parameters.

$\sqrt{s}$	q	N_{dat}	χ^2_{red}
10.58	s	206	4.43
10.58	c	204	1.41
30.73	s	235	0.32
30.73	c	231	0.62
30.73	b	205	0.89
50.88	s	229	0.39
50.88	c	228	0.65
50.88	b	209	0.74
71.04	s	229	0.46
71.04	c	223	0.44
71.04	b	207	0.66
91.19	s	231	0.56
91.19	c	215	0.65
91.19	b	210	0.93
Total		3062	0.91

Table 5.2: χ^2 table: Di-hadron PYTHIA generated data. Summary of χ^2 values per number of points N_{dat} for the PYTHIA datasets where the observable is $\sigma^q/\sigma^{\mathrm{tot}}$.

5.4.2 Experimental and Lattice Data

To constrain D_1, we use the $\pi^+\pi^-$ production in SIA cross section data from the Belle collaboration [333] at $\sqrt{s} = 10.58$ GeV. Since we can only extract information on the IFF up to $M_h = 2.0$ GeV (see below), we place a cut of $M_h < 2.0$ GeV on the cross section data. The data are provided in bins of M_h, with width $\Delta M_h = 0.02$ GeV, and bins of z, with width $\Delta z = 0.05$. We average over each z bin, and evaluate the cross-section at $\langle M_h \rangle$. In order to avoid the sharp kaon and D0 resonances, we also cut out all data points with $\langle M_h \rangle = 0.49$ GeV and $\langle M_h \rangle = 1.87$ GeV. The kinematic limit of the Belle data is given by $M_h < \frac{\sqrt{s}}{2} z$, and this limit is reached in the bins of z below $z = 0.45$. We place a cut of

$$M_h < \alpha_{\text{cut}} \frac{\sqrt{s}}{2} z, \tag{5.43}$$

with $\alpha_{\text{cut}} = 0.7$, in order to avoid the region near the kinematic limit. We also cut out the lowest z bin $0.20 < z < 0.25$ where the applicability of the leading power formalism is questionable [116]. In total, with these cuts we include 1,121 of the 1,468 data points provided by Belle, and the kinematic ranges are $0.25 < z < 1.0$ and $0.3 < M_h < 2.0$ GeV.

To constrain $H_1^\sphericalangle$, we include the SIA asymmetry data provided by Belle [334]. The same cuts on M_h and z discussed above are applied to both hadron pairs here. The data is available in three binnings: (z, M_h), $(M_h, \overline{M}_h)$, and $(z, \bar{z})$, with the non-binned variables integrated over $2m_\pi < M_h, \overline{M}_h < 2$ GeV and $0.2 < z, \bar{z} < 1.0$. We include all three binnings in our analysis. When integrating over M_h or $\overline{M}_h$, we apply the kinematic constraints $M_h < \frac{\sqrt{s}}{2} z$ and $\overline{M}_h < \frac{\sqrt{s}}{2} \bar{z}$. For each binning we also average over the two variables that are not integrated over. We note that in all Belle measurements in this analysis, a cut is placed on the thrust T [335] of $T > 0.8$. In our LO formalism, there is no way to take this thrust cut into account. We expect that the errors from this omission cancel out in the ratio of $H_1^\sphericalangle / D_1$ which appears in all of the asymmetries, thus reducing the impact on the extraction of the transversity PDFs.

The transversity PDFs are constrained by SIDIS and pp data, with the SIDIS data coming from HERMES [319] and COMPASS [318]. This data is again available in three binnings (x_{bj}, z, and M_h) and we use all three. For HERMES, the non-binned variables are averaged over $0.023 < x_{\text{bj}} < 0.4$, $0.2 < z < 1.0$ and $0.5 < M_h < 1.0$

GeV. COMPASS has the same range in z, but a larger range in x and M_h with $0.003 < x_{\rm bj} < 1$ and $0.29 < M_h < 1.29$ GeV. We integrate over all four variables: $x_{\rm bj}$, y, z, and M_h. We denote the integration ranges on $x_{\rm bj}$ and y provided by the experiment as $x^{\rm exp}_{\rm min} < x < x^{\rm exp}_{\rm max}$ and $y^{\rm exp}_{\rm min} < y < y^{\rm exp}_{\rm max}$. But in order to respect the additional phase space cuts on $Q^2 > Q^2_{\rm min}$ and $W^2 > W^2_{\rm min}$, these limits must be adjusted appropriately. We choose to integrate over $x_{\rm bj}$ before y, and adjust the lower and upper limits through [336]

$$x_{\rm min}(y) = {\rm MAX}\Big(x^{\rm exp}_{\rm min}, \frac{Q^2_{\rm min}}{(s-M^2)y}\Big), \tag{5.44}$$

$$x_{\rm max}(y) = {\rm MIN}\Big(x^{\rm exp}_{\rm max}, \frac{y(s-M^2) - W^2_{\rm min} + M^2}{y(s-M^2)}\Big). \tag{5.45}$$

Then the lower limit on y is adjusted through

$$y_{\rm min} = {\rm MAX}\Big(y^{\rm exp}_{\rm min}, \frac{Q^2_{\rm min}}{x^{\rm exp}_{\rm max}(s-M^2)}, \frac{W^2_{\rm min} - M^2}{(1-x^{\rm exp}_{\rm min})(s-M^2)}\Big). \tag{5.46}$$

While integrating over z and M_h, one must also respect the limit

$$M_h < \sqrt{zys - M^2}, \tag{5.47}$$

which is required so that $u \equiv (P_h - P)^2 < 0$. We choose to integrate over M_h before y and z, and adjust the upper limit as

$$M^{\rm max}_h(y,z) = {\rm MIN}\Big(M^{\rm exp}_{h,{\rm max}}, \sqrt{zys - M^2}\Big), \tag{5.48}$$

where $M^{\rm exp}_{h,{\rm max}}$ is the upper limit provided by the experiment.

The pp collision data is provided by STAR, at both $\sqrt{s} = 200$ GeV [337,338] (the latter being preliminary) and $\sqrt{s} = 500$ GeV [339]. The published $\sqrt{s} = 200$ GeV data is provided with three different upper cuts (0.2, 0.3, 0.4) on the opening angle R of the pion pair, with 0.3 treated as the default. This cut is used to filter out pion pairs that do not originate from a single parton. We use the data corresponding to $R < 0.3$ and have tested that this choice does not significantly impact our results. For the preliminary $\sqrt{s} = 200$ GeV data we also use the data corresponding to the cut $R < 0.3$. The $\sqrt{s} = 500$ GeV data is provided with an opening angle of $R < 0.7$. A larger opening angle cut is acceptable here as the increased energy means that a

experiment	ref.	observable	reaction	non-perturbative function(s)	$\sqrt{s}/E_{\rm lab}$ (GeV)
Belle	[333]	$\frac{d^2\sigma}{dzdM_h}$	$e^+e^- \to \pi^+\pi^-$	D_1	10.58
Belle	[334]	a_{12R}	$e^+e^- \to (\pi^+\pi^-)(\pi^+\pi^-)$	$D_1, H_1^\sphericalangle$	10.58
HERMES	[319]	$A_{UT}^{\rm HERMES}$	$ep^\uparrow \to e'\pi^+\pi^-$	$D_1, H_1^\sphericalangle, f_1, h_1$	52
COMPASS	[318]	$A_{UT}^{\rm COMPASS}$	$ep^\uparrow \to e'\pi^+\pi^-$	$D_1, H_1^\sphericalangle, f_1, h_1$	301
STAR	[337–339]	A_{UT}^{pp}	$pp^\uparrow \to \pi^+\pi^-$	$D_1, H_1^\sphericalangle, f_1, h_1$	200, 500
Total					

Table 5.3: Summary of di-hadron production data. Available data for the DiFFs D_1 and $H_1^\sphericalangle$ and the transversity PDF h_1.

pion pair produced from a single initial parton is more likely to be spread out. The data is provided binned in P_{hT}, M_h, and η, with the results (often) provided for both $\eta > 0$ and $\eta < 0$ when binned in P_{hT} or M_h. We include the data binned in all three variables. In contrast to the SIDIS data, we do not integrate over the non-binned variables here. This is due to the fact that the phase space is dependent upon the choice of the opening angle, which cannot be written in terms of P_{hT}, M_h, and η alone. Instead of integrating over the non-binned variables, we take the provided central values which depend upon the opening angle. All of the data is summarized in Table 5.3, which details the reactions and the relevant non-perturbative functions, and the kinematic coverage of the data is shown in Fig. 5.1.

We also consider the inclusion of lattice data in our analysis, in particular from the ETMC collaboration [37] which provides results for the tensor charges δu and δd as well as the isovector tensor charge g_T. One can consider including g_T alone or both δu and δd. We choose the latter, as fitting g_T alone leads to large correlations between δu and δd, which can result in values for δu and δd that are significantly different than the ETMC results.

We discuss now the analysis of the transversity PDFs, where we consider five different scenarios. The "Baseline" fit uses all of the experimental data, including the preliminary STAR data, but no lattice data. The "Lattice" fit includes, in addition, the ETMC results for δu and δd. The "Published" fit excludes the preliminary STAR data. The "RB18-like" fit attempts to replicate the analysis of RB18 [32] as closely as possible within our framework. The "SIDIS only" excludes all of the STAR pp collision data. The details of these fits are summarized in Table 5.4. Unless otherwise specified, results shown are for the "Baseline" fit, which will also be referred to as JAM.

The χ^2 for the "Baseline" fit is summarized in Table 5.5. We are able to describe the Belle cross section, SIDIS, and STAR $\sqrt{s} = 500$ GeV data well. We note that

fit name	SIA asymmetry data	SIDIS data	pp data ($\sqrt{s} = 200$ GeV)	pp data ($\sqrt{s} = 500$ GeV)	lattice data	antiquarks	$H_1^\sphericalangle$ bound	D_1^q
Lattice	All 3 binnings	All 3 binnings	Published + Prelim.	Published	Yes	$\bar{u} = -\bar{d}$	Yes	Fitted
Baseline	All 3 binnings	All 3 binnings	Published + Prelim.	Published	No	$\bar{u} = -\bar{d}$	Yes	Fitted
Published only	All 3 binnings	All 3 binnings	Published	Published	No	$\bar{u} = -\bar{d}$	Yes	Fitted
RB18-like	(z, M_h) binning only	x binning only	Published	—	No	—	No	Zero
SIDIS only	All 3 binnings	All 3 binnings	—	—	No	—	Yes	Fitted

Table 5.4: Summary of transversity PDF fits. The fits differ based on their inclusion of datasets, antiquark transversity PDFs, and the bound on $H_1^\sphericalangle$ (Eq. (5.38)). All fits include the same SIA cross section data from Belle and the same treatment for D_1. All fits include the Soffer Bound. For the RB18-like fit, we do not perform a simultaneous fit of DiFFs and transversity PDFs, rather leaving the DiFFs fixed from the SIA data alone.

there is some difficulty in describing the Belle a_{12R} data, particularly the $M_h, \overline{M}_h$ binning, with a total χ^2_{red} of 1.78. This could be attributed to our LO analysis and inability to take the thrust cut into account. For the $\sqrt{s} = 200$ GeV data, we struggle the most in describing the data binned in η, for both the published and preliminary data. This may be due to the fact that this binning is integrated over the entire range of both M_h and P_{hT}, while in our analysis we are only taking the central values due to our inability to take the opening angle into account. Removing this data from the fit does not improve the description of the other data nor significantly change the extracted PDFs, so we keep it in the fit.

A simplified χ^2 table is shown for all of the fits in Table 5.6. We start with discussing the "Lattice" fit, where we find it easy to describe δd but more difficult to describe δu. From comparison to the "Baseline" fit, it is clear to see that HERMES is the primary source of tension with the lattice data. Upon its removal, the χ^2_{red} for δu decreases to 2.56. Regardless, this relatively minor tension between the experimental and lattice data can be attributed to (presently unknown) systematic uncertainties in both our LO analysis and the lattice results for the tensor charges.

Comparison of the "Baseline" and "Published" fits shows that the preliminary data is largely compatible with the rest of the data. Furthermore, comparisons to the "SIDIS only" fit show that the SIDIS data are compatible with the pp data. Finally, we note that our χ^2_{red} from the "RB18-like" analysis is significantly lower than the value of 2.08 ± 0.09 that was found in Ref. [32]. This difference could be due to the differences in our extractions for the DiFFs and IFFs.

The data vs. theory for the Belle cross section data is shown in Fig. 5.2, while the asymmetry data is shown in Figs. (5.3)–(5.5). The SIDIS data is shown in Fig. 5.6.

experiment	observable	binning	$N_{\rm dat}$	$\chi^2_{\rm red}$	fitted norm.
Belle [333]	$\frac{d^2\sigma}{\mathrm{dz}\mathrm{d}M_h}$	z, M_h	1121	1.24	0.992(20)
Belle [334]	a_{12R}	z, M_h	55	0.53	—
		$M_h, \overline{M}_h$	64	3.43	
		$z, \bar{z}$	64	1.54	
HERMES [319]	$A_{UT}^{\rm HERMES}$	$x_{\rm bj}$	4	1.84	1.101(43)
		M_h	4	1.27	
		z	4	1.74	
COMPASS (p) [318]	$A_{UT}^{\rm COMPASS}$	$x_{\rm bj}$	9	0.88	0.994(4)
		M_h	10	1.12	
		z	7	1.58	
COMPASS (D) [318]	$A_{UT}^{\rm COMPASS}$	$x_{\rm bj}$	9	1.20	1.002(5)
		M_h	10	0.39	
		z	7	0.47	
STAR [337] $\sqrt{s} = 200$ GeV $R < 0.3$	A_{UT}^{pp}	$M_h, \eta < 0$	5	2.54	0.982(17)
		$M_h, \eta > 0$	5	1.52	
		$P_{hT}, \eta < 0$	5	0.92	
		$P_{hT}, \eta > 0$	5	1.05	
		η	4	1.72	
STAR [339] $\sqrt{s} = 500$ GeV $R < 0.7$	A_{UT}^{pp}	$M_h, \eta < 0$	32	0.78	1.078(27)
		$M_h, \eta > 0$	32	1.16	
		$P_{hT}, \eta > 0$	35	1.09	
		η	7	1.57	
STAR [338] $\sqrt{s} = 200$ GeV $R < 0.3$ PRELIMINARY	A_{UT}^{pp}	$M_h, \eta < 0$	31	0.94	0.955(16)
		$M_h, \eta > 0$	31	1.25	
		$P_{hT}, \eta < 0$	29	0.85	
		$P_{hT}, \eta > 0$	29	1.05	
		η	9	2.06	
Total			1627	**1.29**	

Table 5.5: χ^2 table: Di-hadron production data. Summary of $\chi^2_{\rm red}$ values for the various datasets used in the "Baseline" fit. The number of points corresponds to the cuts discussed in Section 5.4.

experiment	N_{dat}	χ^2_{red} Lattice	Baseline	Published only	RB18-like	SIDIS only
HERMES [319]	12	1.92	1.62	1.54	1.97	1.46
COMPASS (p) [318]	26	1.28	1.16	1.13	0.56	1.08
COMPASS (D) [318]	26	0.71	0.69	0.74	1.46	0.71
STAR (2015) [337]	24	1.62	1.54	1.92	1.21	—
STAR (2018) [339]	106	1.09	1.05	1.04	—	—
STAR (PRELIM) [338]	129	1.09	1.10	—	—	—
ETMC δu [37]	1	4.04	—	—	—	—
ETMC δd [37]	1	0.15	—	—	—	—
Total	325	**1.15**	**1.11**	**1.15**	**1.20**	**1.00**

Table 5.6: χ^2 table: Di-hadron production data with different fits. Summary of χ^2_{red} values for the datasets involving the transversity PDFs for the different fit configurations defined in Table 5.4. Note that N_{dat} for SIDIS is smaller for the RB18-like fit, as it does not include all binnings.

The published $\sqrt{s} = 200$ GeV STAR data is shown in Fig. 5.7, while the $\sqrt{s} = 500$ GeV data is shown in Figs. (5.8)–(5.10). As discussed previously, we struggle the most to describe the Belle a_{12R} data binned in $M_h, \overline{M}_h$ and the STAR $\sqrt{s} = 200$ GeV data binned in η. From the plots, we see that we are generally underestimating the a_{12R} data. We also underestimate the STAR data at $\sqrt{s} = 200$ GeV data binned in η, which may be due to our inability to properly integrate over M_h and P_{hT}.

Finally, we mention that performing a simultaneous fit of the DiFFs and PDFs (as opposed to keeping the DiFFs fixed from the SIA data alone) significantly improves the description of the STAR data, particularly at $\sqrt{s} = 200$ GeV. The χ^2_{red} for the published $\sqrt{s} = 200$ GeV data improves from 1.86 to 1.54, while it improves from 1.49 to 1.10 for the preliminary data. For the $\sqrt{s} = 500$ GeV data there is a small improvement from 1.13 to 1.05. For all of the STAR data, the χ^2_{red} improves from 1.38 to 1.12. The improvements are the largest for the data that is binned in η. These changes emphasize the potential importance of performing simultaneous analyses of the different functions that enter a global QCD analysis.

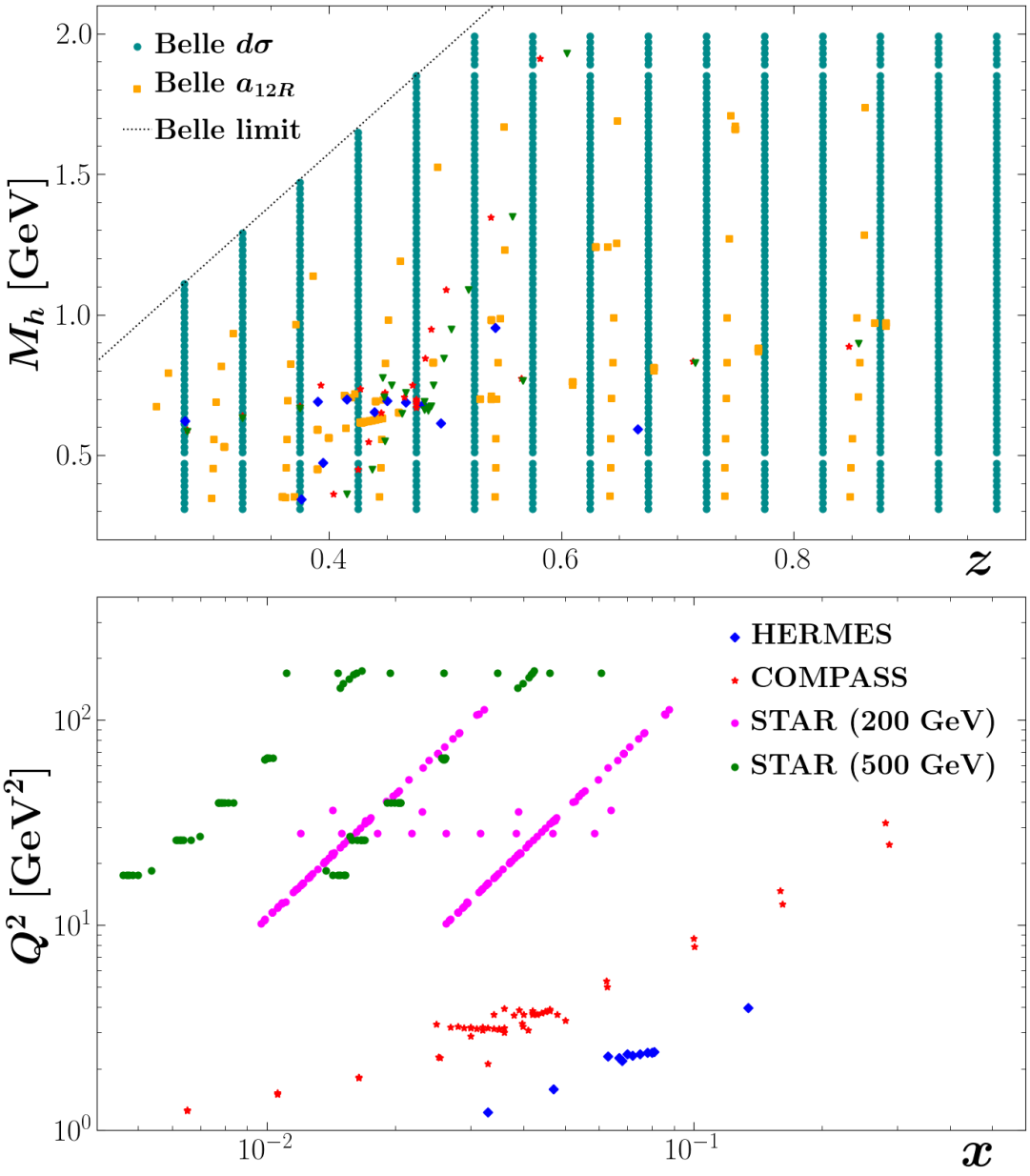

Figure 5.1: Kinematics of di-hadron production datasets. Kinematic coverage of the datasets included in this analysis. The top panel shows the data as a function of z and M_h. The kinematic limit of the Belle data (Eq. (5.43)) is also shown. The bottom panel shows the data as a function of x and Q^2. The variable x represents $x_{\rm bj}$ (Eq. (2.7)) for SIDIS and Eq. (5.31) for pp collisions, while the scale Q^2 represents the four-momentum transfer squared for SIDIS and the transverse momentum squared of the di-hadron pair for pp collisions.

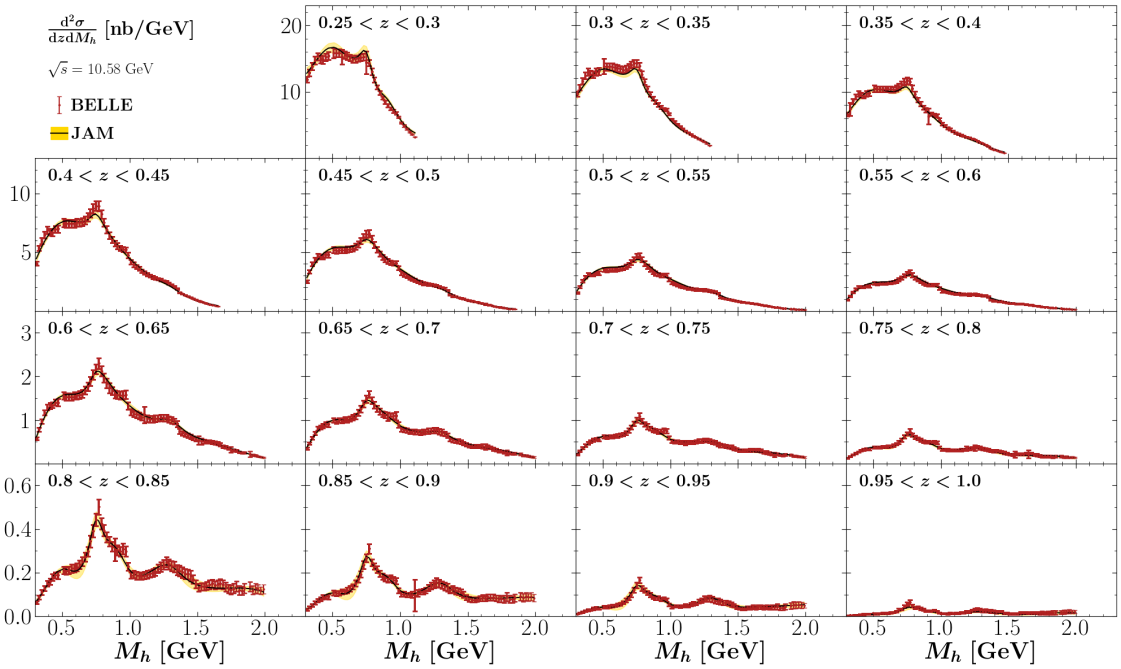

Figure 5.2: Data vs. Theory: Di-hadron SIA cross section. The Belle cross section data [333] at $\sqrt{s} = 10.58$ GeV (red points) plotted as a function of M_h against the mean JAM result (black line) with 1σ uncertainty bands in gold. The different panels show different bins of z.

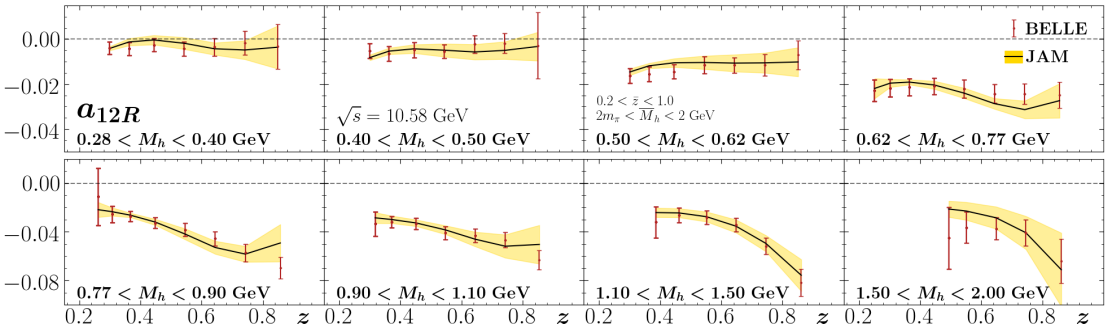

Figure 5.3: Data vs. Theory: Di-hadron SIA asymmetry binned in (z, M_h). The Belle a_{12R} data [334] binned in (z, M_h) at $\sqrt{s} = 10.58$ GeV (red points) plotted as a function of z against the mean JAM result (black line) with 1σ uncertainty bands in gold. The different panels show different bins of M_h, while $\bar{z}$ is integrated over $0.2 < \bar{z} < 1.0$ and $\overline{M}_h$ over $2m_\pi < \overline{M}_h < 2$ GeV.

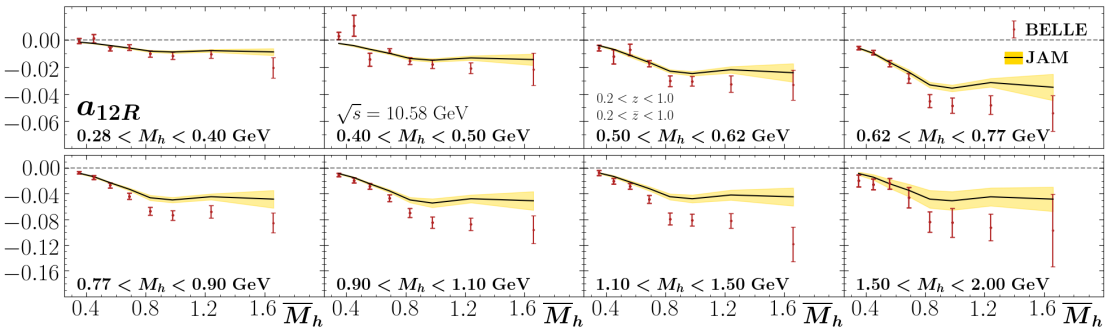

Figure 5.4: Data vs. Theory: Di-hadron SIA asymmetry binned in $(M_h, \overline{M}_h)$. The Belle a_{12R} data [334] binned in $(M_h, \overline{M}_h)$ at $\sqrt{s} = 10.58$ GeV (red points) plotted as a function of $\overline{M}_h$ against the mean JAM result (black line) with 1σ uncertainty bands in gold. The different panels show different bins of M_h, while z and $\bar{z}$ are integrated over $0.2 < z, \bar{z} < 1.0$.

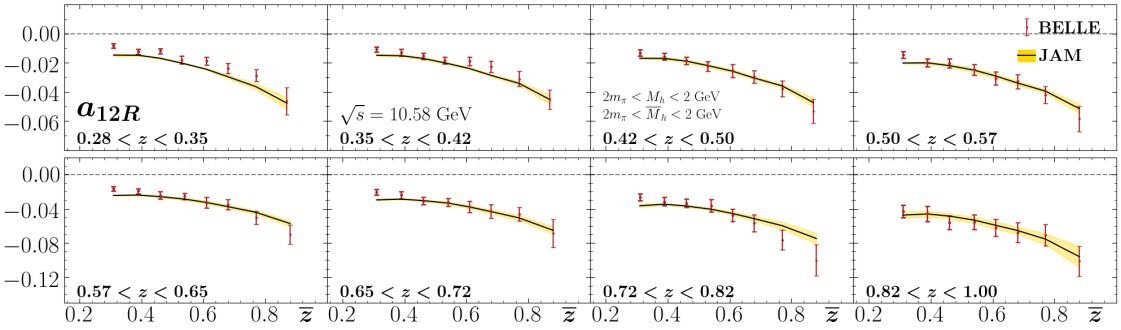

Figure 5.5: Data vs. Theory: Di-hadron SIA asymmetry binned in $(z, \bar{z})$. The Belle a_{12R} data [334] binned in $(z, \bar{z})$ at $\sqrt{s} = 10.58$ GeV (red points) plotted as a function of $\bar{z}$ against the mean JAM result (black line) with 1σ uncertainty bands in gold. The different panels show different bins of z, while M_h and $\overline{M}_h$ are integrated over $2m_\pi < M_h, \overline{M}_h < 2$ GeV.

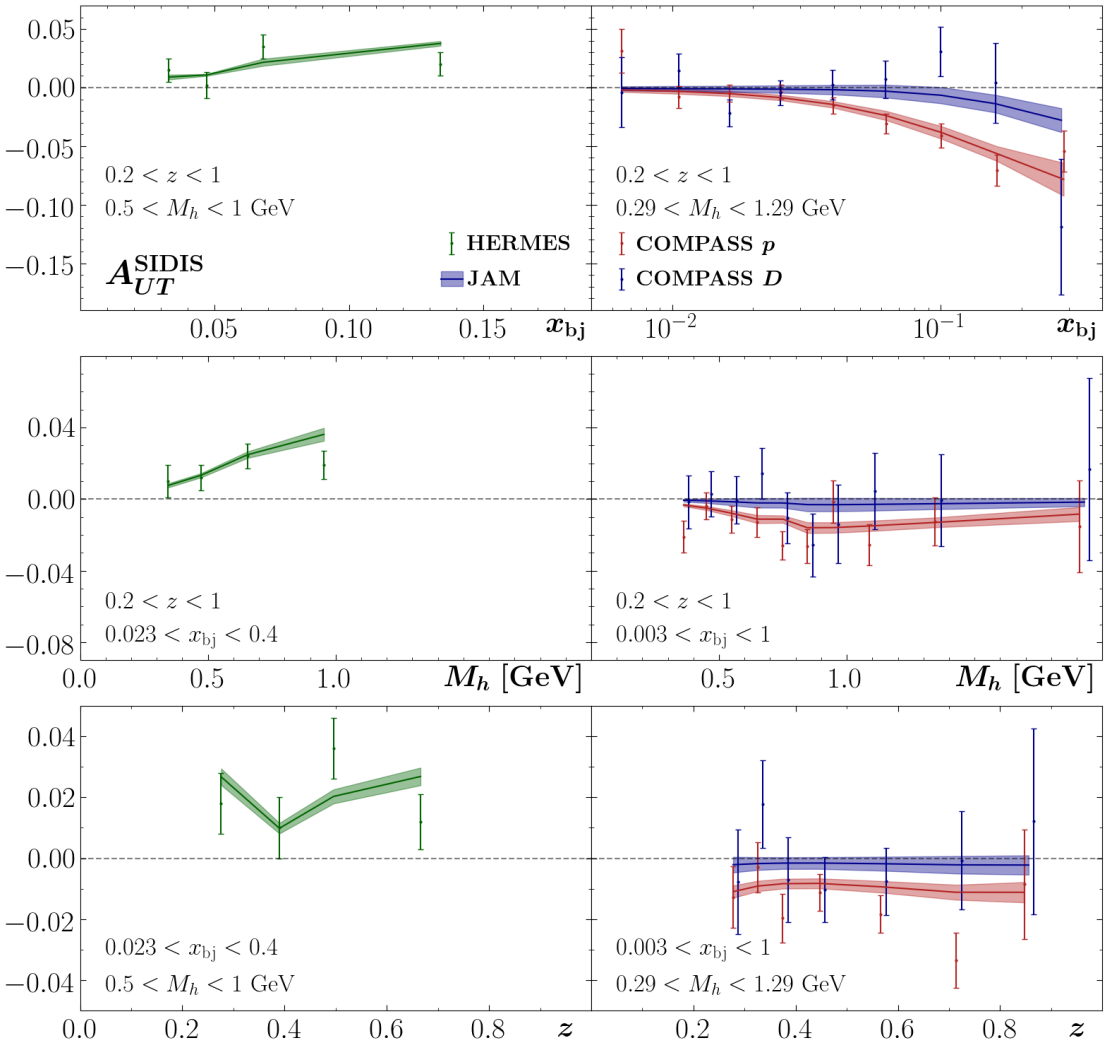

Figure 5.6: Data vs. Theory: Di-hadron SIDIS. The SIDIS A_{UT} data from HERMES [319] (green points) and COMPASS [318] (red points for proton, blue points for deuteron) plotted against the mean JAM result (colored lines and 1σ uncertainty bands). Results are shown for the data binned in $x_{\rm bj}$ in the top row, M_h in the middle row, and z in the bottom row. The non-binned variables are integrated over $0.023 < x_{\rm bj} < 0.4$, $0.5 < M_h < 1$ GeV, and $0.2 < z < 1.0$ for HERMES and $0.003 < x_{\rm bj} < 1$, $0.29 < M_h < 1.29$ GeV, and $0.2 < z < 1.0$ for COMPASS.

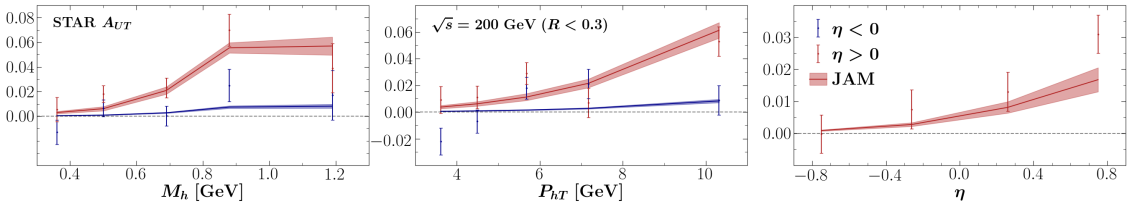

Figure 5.7: Data vs. Theory: Di-hadron production from pp at $\sqrt{s} = 200$ GeV. The STAR data [337] with opening angle cut $R < 0.3$ (colored points) are plotted against the mean JAM result (colored lines and 1σ uncertainty bands). The left panel shows the results as a function of M_h, the middle panel as a function of P_{hT}, and the right panel as a function of η. In the left and middle panel, the results are shown in red for $0 < \eta < 1$ and blue for $-1 < \eta < 0$.

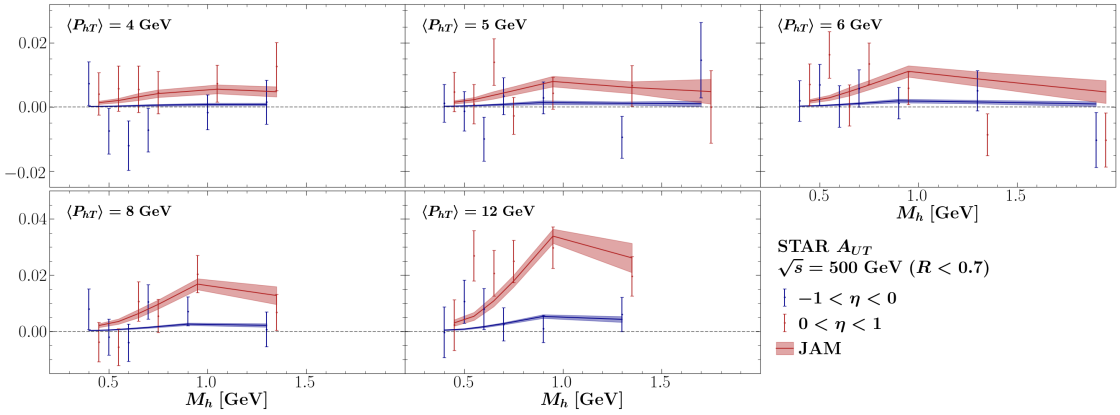

Figure 5.8: Data vs. Theory: Di-hadron production from pp at $\sqrt{s} = 500$ GeV binned in M_h. The STAR data [339] with opening angle cut $R < 0.7$ (red points for $0 < \eta < 1$, blue points for $-1 < \eta < 0$) are plotted as a function of M_h against the mean JAM result (colored lines and 1σ uncertainty bands). The different panels show different bins of P_{hT}.

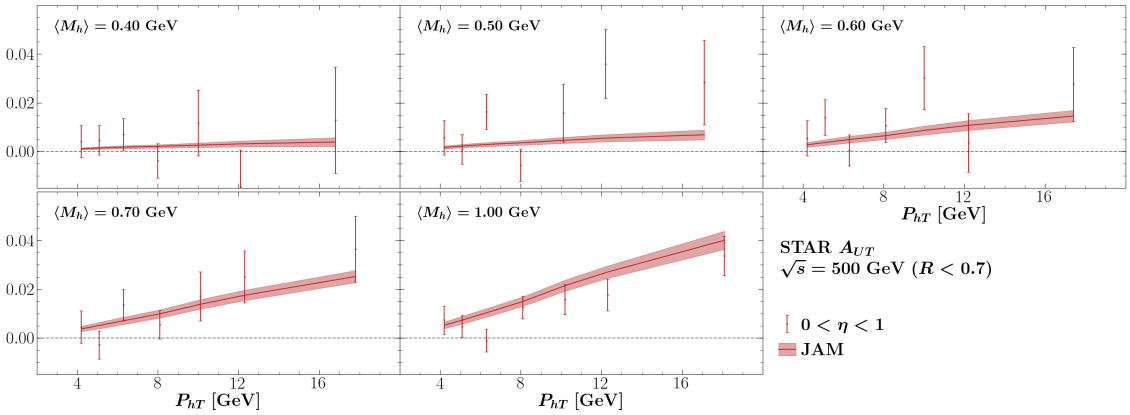

Figure 5.9: Data vs. Theory: Di-hadron production from pp at $\sqrt{s} = 500$ GeV binned in P_{hT}. The STAR data [339] with opening angle cut $R < 0.7$ (red points for $0 < \eta < 1$) are plotted as a function of P_{hT} against the mean JAM result (colored lines and 1σ uncertainty bands). The different panels show different bins of M_h.

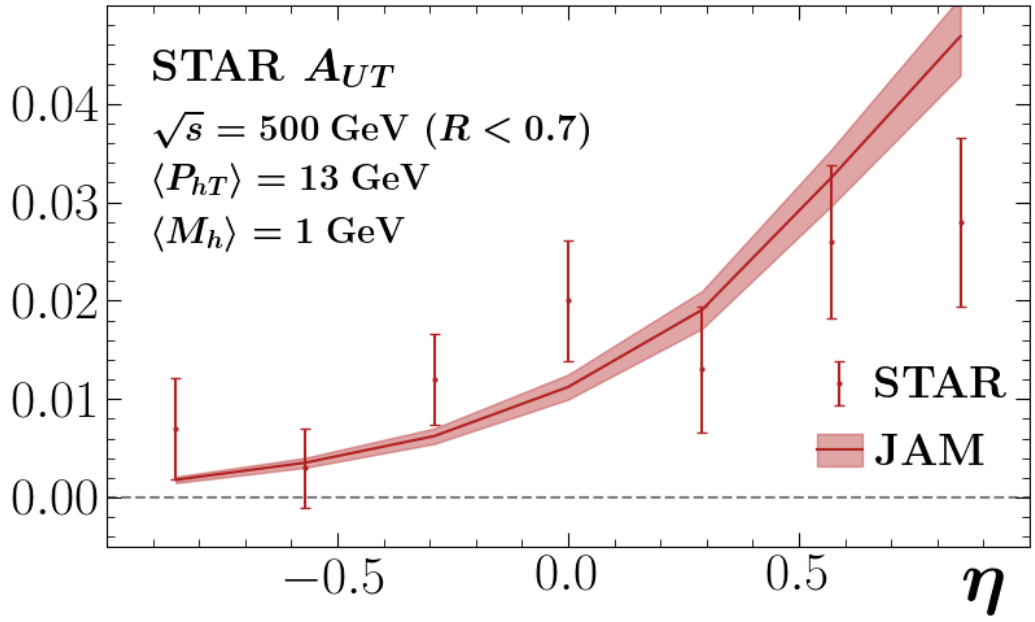

Figure 5.10: Data vs. Theory: Di-hadron production from pp at $\sqrt{s} = 500$ GeV binned in η. The STAR data [339] with opening angle cut $R < 0.7$ (red points) are plotted as a function of η against the mean JAM result (red line and 1σ uncertainty band). The results are plotted at the mean values $\langle P_{hT} \rangle = 13$ GeV and $\langle M_h \rangle = 1$ GeV.

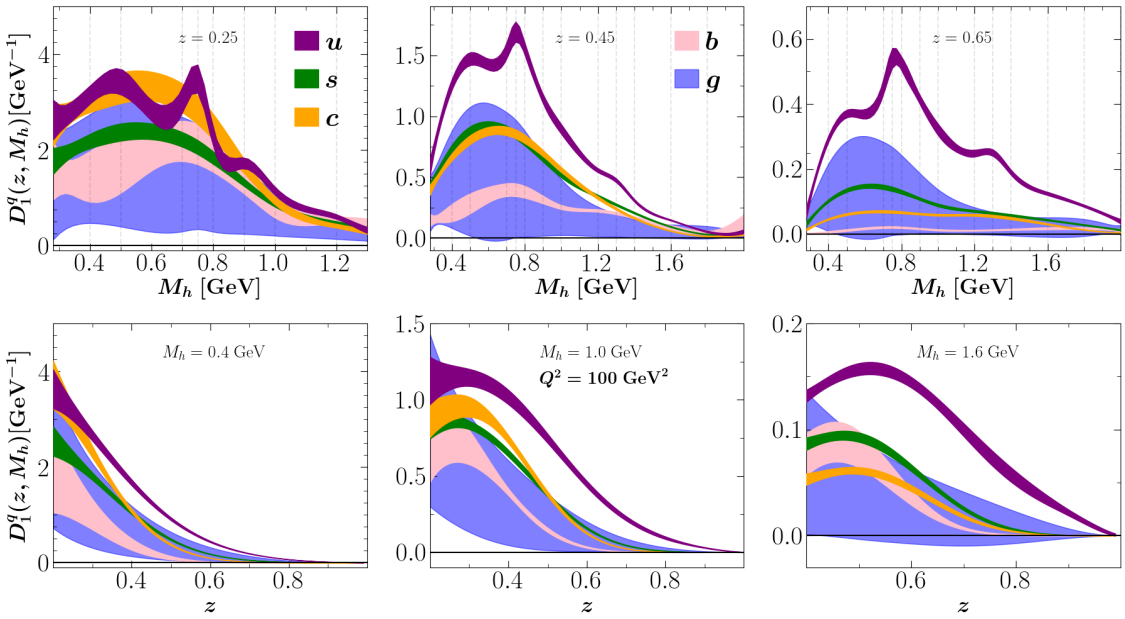

Figure 5.11: Extracted di-Hadron fragmentation functions. DiFFs D_1^q plotted as a function of M_h with $z = 0.25, 0.45, 0.65$ (top row) and as a function of z with $M_h = 0.4, 1.0, 1.6$ (bottom row) at the scale $Q^2 = 100$ GeV^2. The up, strange, charm, bottom, and gluon are shown in purple, green, orange, pink, and blue respectively. The vertical dashed lines on the top row show where the DiFFs are parametrized (see Section 5.3.1).

5.5 Extracted Di-Hadron Fragmentation Functions

As noted in Subsection 5.4.2, the cut on the thrust used in the Belle measurements cannot be accounted for in our LO formalism and may have a significant impact on the extractions of D_1 and $H_1^{\sphericalangle}$ that is expected to partially cancel in the ratio $H_1^{\sphericalangle}/D_1$. Nevertheless, we show here our results for both DiFFs with the caveat that there may be significant changes in an NLO extraction that is able to take the thrust cut into account.

In Fig. 5.11 we show all of the unpolarized DiFFs at the scale $Q^2 = 100$ GeV^2 (approximately the scale of the Belle experiment). We generally find strong constraints on all of the quark DiFFs, but very weak constraints on the gluon DiFFs. This is to be expected, as we have no observables sensitive to the gluon distribution at LO and instead depend entirely on evolution to constrain it (as discussed previously, the pp asymmetry cannot help significantly here). Measurements on the unpolarized pp cross section will provide stronger constraints on D_1^g in the future. In the RB18 anal-

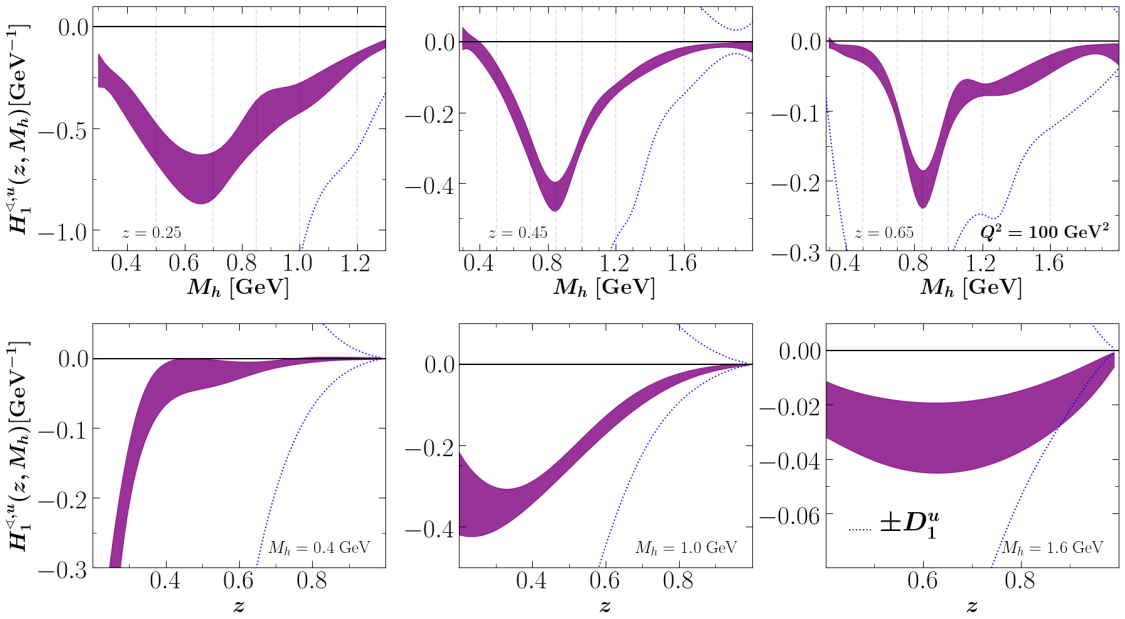

Figure 5.12: Extracted interference fragmentation functions. IFF $H_1^{\sphericalangle,u}$ plotted as a function of M_h with $z = 0.25, 0.45, 0.65$ (top row) and as a function of z with $M_h = 0.4, 1.0, 1.6$ (bottom row) at the scale $Q^2 = 100$ GeV^2. The vertical dashed lines on the top row show where the DiFFs are parametrized (see Section 5.3.2). The dashed blue line represents the $|H_1^{\sphericalangle,u}| < D_1^u$ bound.

ysis [32], they considered three scenarios for the gluon distribution at the input scale: $D_1^g = 0$, $D_1^g = D_1^u/4$, and $D_1^g = D_1^u$. The continuous distribution of gluon solutions shown here should provide a more realistic estimate of the errors on the extracted transversity PDFs.

In Fig. 5.12 we show the IFF at the scale $Q^2 = 100$ GeV^2. As mentioned in Subsection 5.2.1, the sign of $H_1^{\sphericalangle,u}$ cannot be fixed by the experimental data alone, and we have thus forced $H_1^{\sphericalangle,u}$ to be negative. We generally find that it does not struggle to satisfy the bound $|H_1^{\sphericalangle,u}| < D_1^u$, except at larger M_h and z where it begins to be limited by it.

5.6 Extracted Transversity PDFs and Tensor Charges

As discussed in Subsection 5.4.2, we perform five different analyses for the transversity PDFs, summarized in Table 5.4, to assess the impact of experimental datasets and the inclusion of lattice QCD in the fit. In the following we will frequently compare our results to the analyses of JAM3D [34] and RB18 [32]. The former extracts the transversity distributions alongside TMDs using the Collins and Sivers effects, while the latter uses the collinear DiFF approach as in this analysis. The two approaches have led to results for the transversity PDFs and tensor charges that seemingly disagree rather strongly, especially for the up quark. The JAM3D analysis included lattice QCD results into their fit and found good agreement between lattice and experiment. On the other hand, the RB18 results seem to be in strong disagreement with lattice. This situation has led to a "transverse spin puzzle" in trying to reconcile the results from the DiFF approach, TMD approach, and lattice QCD. Below we will show our own results for the transversity PDFs and tensor charges and provide evidence that the three approaches are in fact reconcilable.

In this context it is imperative to first discuss the fact that lattice QCD calculations are for the full tensor charge, while experimental data is only available down to $x \approx 0.005$ in both the DiFF and TMD approaches. Thus, unless there are restrictions on the behavior of the PDFs at small-x (either explicitly or through an inflexible parameterization), the full moments extracted from experimental data alone are subject to extrapolation errors which are entirely dependent upon the choice of parameterization. In this scenario, our results cannot give a realistic estimate of the errors. Furthermore, with no restrictions on the small-x behavior, it is *always* possible to find compatibility between experiment and lattice QCD by using both in the fit, as the fit is free to adjust the small-x behavior as necessary to match the lattice results. As a relevant example, even if δu from lattice QCD is greatly underestimated by the result from experiment alone, the result from lattice QCD can still be accommodated in the fit by making the u PDF very large below $x \approx 0.005$.

Thus, after including the lattice QCD data into the fit, is it necessary to check if the small-x behavior is meaningful. For this analysis we are in the fortunate situation to have a theoretical prediction for the small-x behavior. As discussed in Section 5.3.3, we use theoretical predictions [330] to constrain the small-x behavior by restricting

the α parameter of Eq. (2.40) to be $0.17 \pm 50\%$. We assume that this restriction on the small-x behavior ensures that the behavior of the PDFs is physical, and find that the contribution to the moments below $x \approx 0.005$ is small. Thus the fit no longer has the freedom to manipulate the PDFs at small-x in order to accommodate the lattice data, and a fit of both experimental and lattice data is a meaningful test of their agreement. An analogous discussion also applies to the high-x region, since experimental data only goes as high as $x \approx 0.3$. The Soffer Bound, which is included in all three phenomenological analyses discussed here, solves the issue by providing strong theoretical constraints at high-x that prevent the fit from seeking non-physical solutions to accommodate the lattice data.

In Fig. 5.13 we show our results for valence PDFs from the "Baseline," "Published," and "SIDIS only" scenarios. For the up quark, we see a consistent trend of the pp data pushing the up quarks towards slightly larger values in the valence region, with the "Published" result being larger than the "SIDIS only" result and the "Baseline" result being even larger. For the down quark, there seems to be less of a consistent trend. The SIDIS and preliminary STAR data seem to prefer a negative and (relatively) large down quark, while the published STAR data prefer a down quark that is closer to zero. For $x \lesssim 0.1$, we see agreement between all three scenarios.

In Fig. 5.14 we compare our "Baseline" results for the valence distributions to those from JAM3D without lattice QCD [34] and RB18 [32] (whose analysis did not consider the inclusion of lattice data). Antiquarks were not considered in these other analyses, so one has $q_v = q$. The two analyses find significantly different results for u_v, with RB18 finding a result less than half as large as the JAM3D result near the peak of the distribution. Our result ends up in the middle of the two analyses. This can be largely attributed to the data that was not available at the time of the RB18 analysis: the published $\sqrt{s} = 500$ GeV and preliminary $\sqrt{s} = 200$ GeV data from STAR. Both datasets push the up quark to be larger, the $\sqrt{s} = 200$ GeV data more so than the $\sqrt{s} = 500$ GeV. This upward push helps significantly when it comes to finding agreement with the lattice QCD data. For d_v, JAM3D and RB18 are in agreement, while our result is larger (but still within errors). As mentioned previously, both the SIDIS data and preliminary $\sqrt{s} = 200$ GeV STAR data push the fit towards a clearly negative down quark.

In Fig. 5.15 we show the "Baseline" result for the antiquark distributions, where we assume $h_1^{\bar{u}} = -h_1^{\bar{d}}$. As expected, the Soffer Bound forces the antiquarks to be very

small above $x \gtrsim 0.2$. Below where the Soffer Bound restricts the distributions, they still remain small and consistent with zero. $\bar{u}$ ($\bar{d}$) has a very slight preference to be negative (positive) around $x \approx 0.1$. Interestingly, this is where the sea asymmetry peaks in the case of helicity PDFs (see Chapter 4). More precise data is needed to determine if a "transversity sea asymmetry" also exists.

In Fig. 5.16 we show the valence distributions for both the "Baseline" and "Lattice" fits. We see a clear increase in the u_v distribution in the valence region when including the lattice data, which is a consequence of the lattice result for δu being significantly larger than the result of the "Baseline" (the tensor charge results will be discussed below). For d_v the "Baseline" result already agrees well with the tensor charge for lattice, and so the decrease around $x \approx 0.1$ is compensated by an increase around $x \approx 0.01$. This change in shape but not overall magnitude is likely a result of the changes seen in u_v, which d_v must compensate for in order to retain a good description of the experimental data.

We now move on to discussing the tensor charges, defined in Eqs. (5.1)–(5.2). Note that, as discussed previously, we use theoretical constraints at small $x \lesssim 0.005$ (α parameter within $0.17 \pm 50\%$) and high $x \gtrsim 0.3$ (the Soffer Bound) so that our result for the full moment is not subject to uncontrolled extrapolation errors. In Fig. 5.17 we show the tensor charges extracted from the "Baseline," "Published," "SIDIS only," and "RB18-like" fits and compare to JAM3D [34] and RB18 [32]. For δu and δd, we find a value that's in between RB18 and JAM3D for the former and a value that's lower than both for the latter, as expected from Fig. 5.14. We find agreement within errors with RB18 but not JAM3D, confirming the finding that the DiFF and TMD approaches to extracting transversity lead to different results for the tensor charges. For g_T we find agreement with all other phenomenological extractions due to large error bands on most extractions. Comparing the four fits from our own analysis, we see that the SIDIS only result is in very good agreement with the RB18 result. All of the STAR data prefer a larger δu, however the published and preliminary data pull δd in opposite directions, as was seen in Fig. 5.13. The "RB18-like" fit has much larger errors than the others, due to the missing bins for the SIA and SIDIS asymmetries (the lack of the bound Eq. (5.38) on $H_1^{\sphericalangle}$ does not significantly impact the errors). The resulting errors are similar to those from the RB18 analysis. The results for δd and g_T match extremely well, while our result for δu is larger but agrees within errors.

In Fig. 5.18 we show our results with the lattice data of Ref. [37] included and compare to other lattice QCD results [35, 36, 297]. Our fit has no issue in accommodating the lattice result for δd. This is not surprising given that the "Baseline" result for δd was already in good agreement. Our result for δu is slightly too small, but it has moved much closer to the lattice result after its inclusion in the fit. For g_T (which is not directly included in this fit) we find a result that is slightly smaller than the lattice results, which is an expected consequence of δu also being slightly smaller.

Although the "Baseline" result for δu is much lower than the value from Ref [37], we find that the fit is able to include it while only affecting the description of the experimental data marginally. The shift from the "Baseline" result to the "Lattice" result (particularly for δu) seems shocking at first. We argue the following: the "Baseline" result is the best solution given the experimental data. We find (upon including the lattice data) that there is another solution that describes the experimental data slightly worse, but that can also accommodate the lattice data. This shows that while the experimental data has a preferred solution for the tensor charge, this is a weak preference as the experimental data is not directly sensitive to the full moment. Before coming to a conclusion about the compatibility between lattice QCD and experimental data, one needs first to include both in the fit. One should only be concerned if the description of the lattice data remains poor even after including it in the fit. Both this analysis and the JAM3D analysis find that it *is* possible to include the lattice QCD data, and the resulting x-dependence of the transversity PDFs (from Fig. 5.14 and Fig. 5.16) is similar. Thus we conclude that there is no "transverse spin puzzle," and that the collinear DiFF, TMD, and lattice QCD approaches are all compatible. In the future, a simultaneous analysis of all three approaches can prove this definitively.

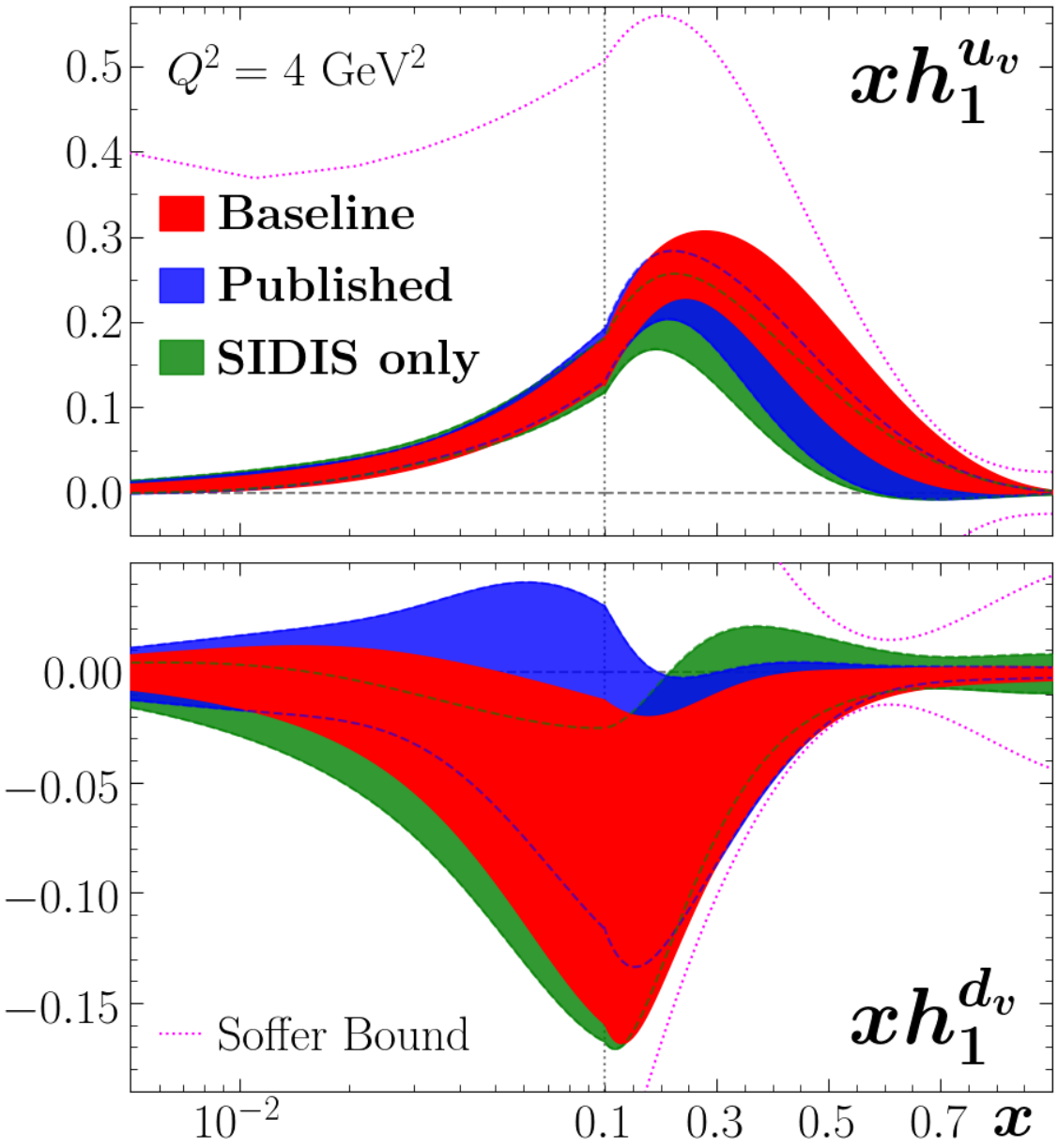

Figure 5.13: Extracted transversity valence PDFs. Transversity PDFs $h_1^{u_v}$ and $h_1^{d_v}$ plotted as a function of x at the scale $Q^2 = 4$ GeV^2. The "Baseline" result is shown in red, while the "Published only" and "SIDIS only" results are shown in blue and green, respectively (see Table 5.4). The Soffer Bound is indicated by the dashed pink lines. Note that the PDFs are multiplied by a factor of x.

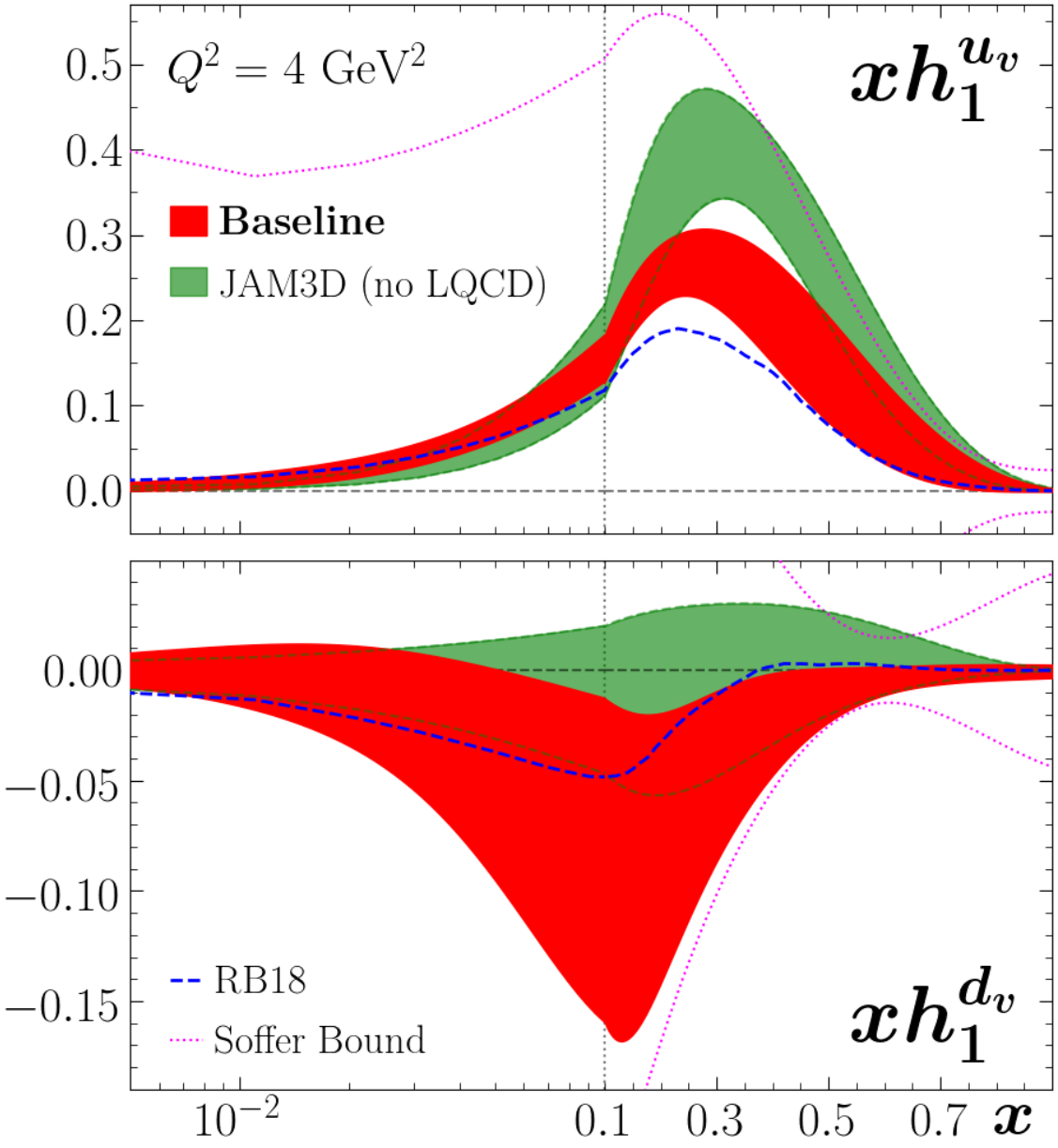

Figure 5.14: Extracted transversity valence PDFs compared to other groups. Transversity PDFs $h_1^{u_v}$ and $h_1^{d_v}$ plotted as a function of x at the scale $Q^2 = 4$ GeV^2. The "Baseline" result is shown in red, and is compared to the result from JAM3D (without lattice QCD) [34] (green bands) and the result of RB18 [32] (dashed blue lines). The Soffer Bound is indicated by the dashed pink lines. Note that the PDFs are multiplied by a factor of x.

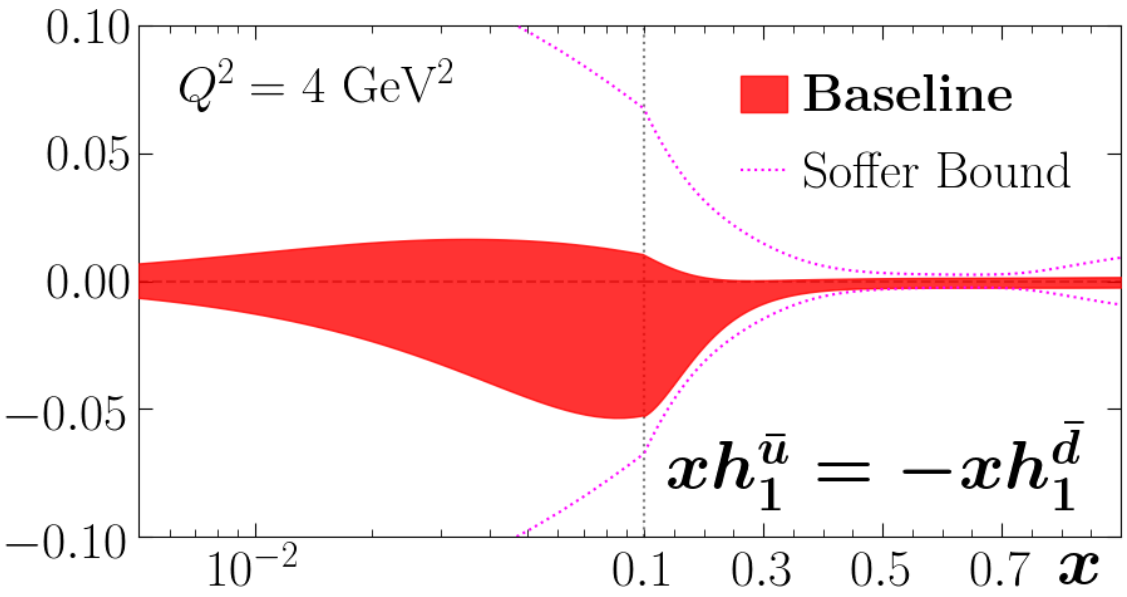

Figure 5.15: Extracted transversity antiquark PDFs. Transversity PDF $h_1^{\bar{u}} = -h_1^{\bar{d}}$ plotted as a function of x at the scale $Q^2 = 4$ GeV^2. In our analysis we assume this relationship between the antiquark PDFs (see Subsection 5.3.3). The "Baseline" result is shown in red, while the Soffer Bound is indicated by the dashed pink lines. Note that the PDFs are multiplied by a factor of x.

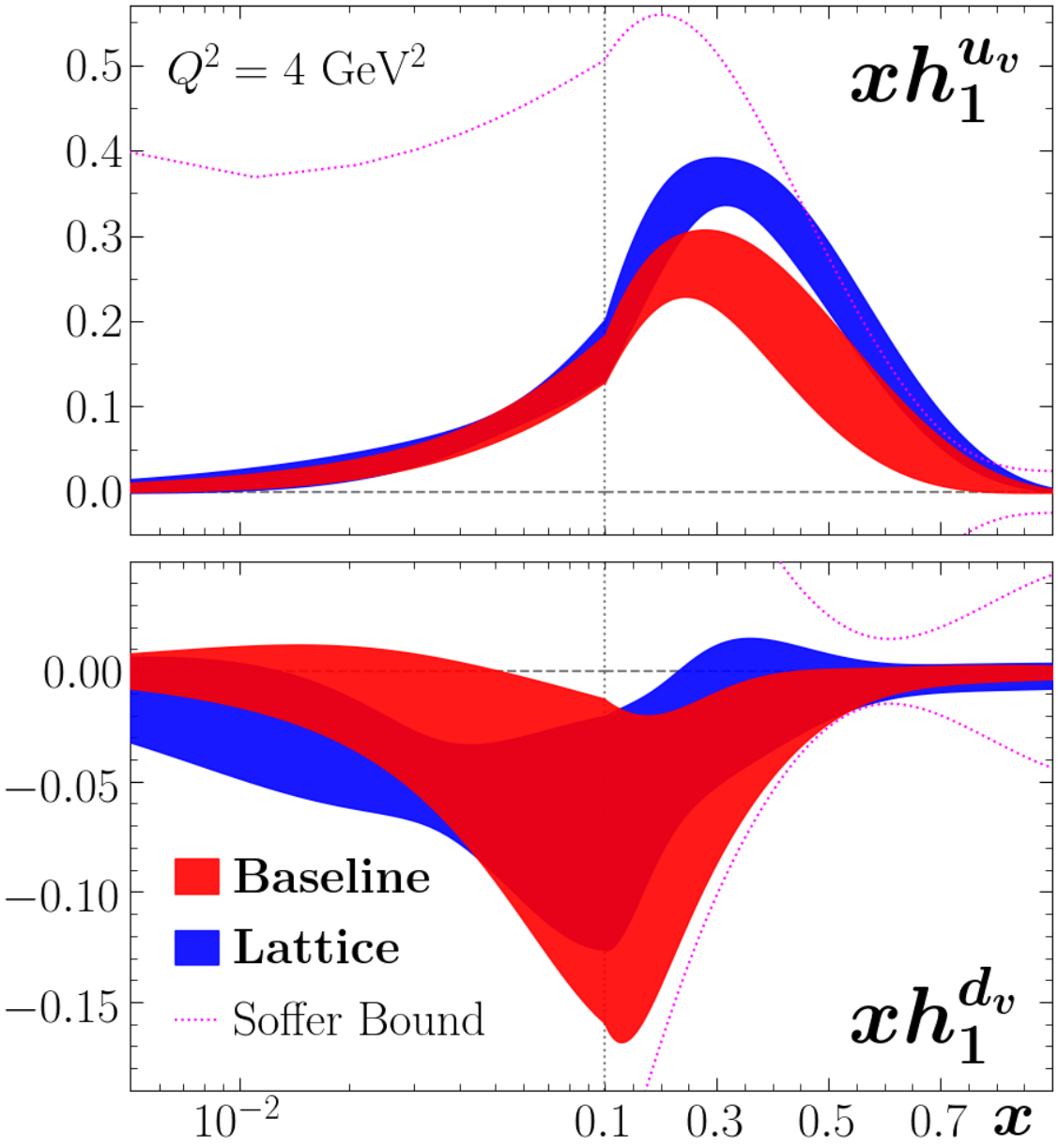

Figure 5.16: Extracted transversity valence PDFs with lattice QCD. Transversity PDFs $h_1^{u_v}$ and $h_1^{d_v}$ plotted as a function of x at the scale $Q^2 = 4$ GeV^2. The "Baseline" result is shown in red, while the "Lattice" result is shown in blue (see Table 5.4). The Soffer Bound is indicated by the dashed pink lines. Note that the PDFs are multiplied by a factor of x.

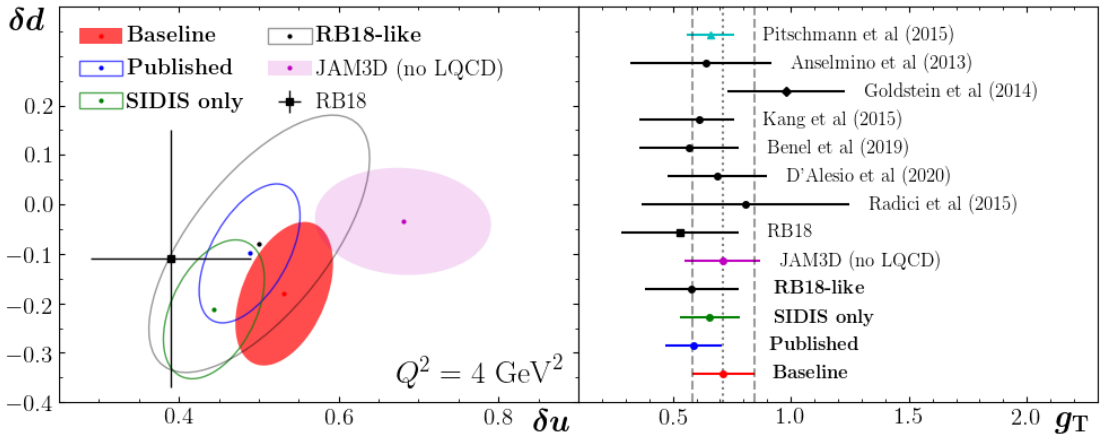

Figure 5.17: Extracted tensor charges compared to phenomenology. The tensor charges δu, δd, and g_T at the scale $Q^2 = 4$ GeV^2. The "Baseline" result is shown in red, while the "Published," "SIDIS only," and "RB18-like" results are shown in blue, green, and gray respectively. δu and δd are compared to the JAM3D [34] result without Lattice QCD (magenta) and the result of RB18 [32] (black point). g_T is also compared to other phenomenological extractions [31, 298–302] (black points) and the result from Dyson-Schwinger [340] (cyan).

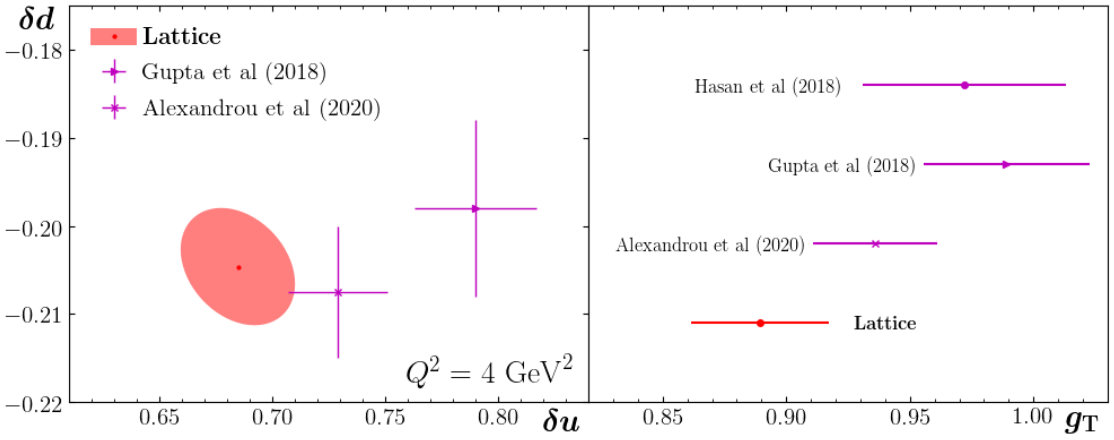

Figure 5.18: Extracted tensor charges compared to lattice QCD. The tensor charges δu, δd, and g_T at the scale $Q^2 = 4$ GeV^2. The "Lattice" result with δu and δd from Ref. [37] included in the fit is shown in red and is compared to lattice QCD results [35–37, 297] (magenta).

5.7 Summary

In this chapter we have presented the results of a simultaneous global QCD analysis of DiFFs, IFFs, and transversity PDFs [60]. We have also proposed a new definition for the unintegrated DiFFs that is compatible with the probability interpretation of collinear DiFFs and derive the LO evolution equations with this new definition [59]. We include, for the first time, the BELLE measurement on the e^+e^- cross section [333], the STAR measurement at $\sqrt{s} = 500$ GeV [339], and preliminary data from STAR at $\sqrt{s} = 200$ GeV [338]. The data from BELLE provide the first direct experimental information on the unpolarized DiFFs, and allow for global analyses to reduce (but not eliminate) dependence on event generators to constrain these functions. The new data from STAR cause the up (down) transversity PDF to increase (decrease), moving our phenomenological results closer to those from lattice QCD. Even though several aspects of our analysis differ from the approach used by the Pavia group, our extracted tensor charges from experiment are in good agreement with RB18 [32] which also used the DiFF channel to extract the transversity PDFs.

Upon including the lattice QCD measurements of δu and δd from ETMC [37], we find good agreement between the lattice QCD results and experiment. The remaining tension could be attributed to approximations in our analysis, particularly the fact that it is only LO in the strong coupling, or to underestimated systematic uncertainties on the lattice QCD results. We note that JAM3D [34] also found agreement with lattice QCD. This fact combined with our analysis strongly indicates that the three approaches (DiFF extraction used here and in RB18 [32], TMD extraction used by JAM3D, and lattice QCD) are in fact reconcilable, and that there is no "transverse spin puzzle." A simultaneous analysis of both the DiFF and TMD channels, with lattice QCD included, would provide a definitive test of this hypothesis.

CHAPTER 6

GENERALIZED PARTON DISTRIBUTION FUNCTIONS

First-principles calculations of PDFs using lattice QCD have remained challenging due to their explicit time-dependence. As a result, in the past almost all related studies in lattice QCD have focused on moments of PDFs which are defined through time-independent local operators. Meanwhile, the full dependence of PDFs on the parton momentum fraction x has remained elusive.

The recently proposed quasi parton distributions (quasi-PDFs) offer a way to directly access the x-dependence of the PDFs in lattice QCD [82, 341]. Quasi-PDFs are defined through spatial equal-time operators that can be computed on four-dimensional Euclidean lattices. They reduce to their corresponding light-cone PDFs if the hadron momentum $P^3 = |\vec{P}| \to \infty$, prior to renormalization. However, for lattice calculations one first renormalizes, and P^3 is finite. This leads to two sources of discrepancies between quasi-PDFs and light-cone PDFs: higher twist corrections that are suppressed by powers of $\frac{1}{P^3}$, and different ultraviolet (UV) behavior for these two types of PDFs. The UV disparities can be cured order by order in perturbative QCD through a so-called matching procedure — see for instance Refs. [342–344]. Other approaches for computing the x-dependence of PDFs and related quantities have also been suggested [83, 84, 345–354]. Some of them are closely related to the concept of quasi-PDFs.

By now there has been important progress in understanding the renormalization of quasi-PDFs [85, 355–366]. A variety of other aspects of quasi-PDFs and, generally, Euclidean correlators have also been extensively studied [367–394]. In particular,

the first lattice QCD results for quasi-PDFs and related quantities can be considered milestones in this field [37,84,85,359,362,395–433], with results on higher twists [434–437], parton distribution amplitudes [399, 403, 409, 438–442], and TMDs [443–448]. Additionally, the convergence of quasi-PDFs to the corresponding light-cone PDFs has been explored in several models [61, 449–455]. The progress in this field was recently reviewed in Refs. [456–460].

As already pointed out in Ref. [82], the concept of quasi-distributions is not limited to forward PDFs. For example, generalized parton distributions (GPDs) [461–465] could also be addressed in this approach. A number of compelling motivations to study GPDs exist, with many details to be found in various review articles [26–28, 466–469]. These reasons include the facts that GPDs provide 3D images of hadrons [470–473], provide access to the angular momenta of partons [462], and are related to pressure and shear forces inside hadrons [474–476]. While information on GPDs can be extracted from processes such as deep virtual Compton scattering [461–464, 477] and hard exclusive meson production [465, 478, 479], such an extraction presents complicated issues [480].

In this situation, reliable information from lattice QCD on GPDs using quasi distributions is very helpful. Previous lattice QCD calculations provided information only on the lowest moments of GPDs [481–485], but in recent years simulations at the physical point have become available [486–499]. The quasi-GPD approach can be used to gain information on the x dependence of GPDs. In Refs. [500–506] the perturbative matching for quasi-GPDs was studied, while model studies were performed in Refs. [61, 62, 507, 508]. The first lattice QCD studies have also been performed for the pion [509] and the nucleon [510–516].

In this chapter, which is based on Refs. [61, 62], we present the results for quasi-distributions in the Scalar Diquark Model (SDM), including the quasi-GPDs corresponding to the eight leading twist light-cone GPDs as well as their forward limits which, for specific quasi-GPDs, reduce to the quasi-PDFs. We present analytical and numerical results for the quasi-GPDs and quasi-PDFs. The convergence of the quasi-GPDs to their corresponding GPDs in the large momentum limit is tested, as well as the dependence of the results on the input parameters of the model. Finally, we compare the moments of the quasi-GPDs to predictions based on spin sum rules and momentum dependence.

6.1 Definition of GPDs

The definition for twist-2 light-cone GPDs of quarks within a spin-$\frac{1}{2}$ hadron is given by a Fourier transform of off-forward matrix elements of bi-local quark operators (see for instance Ref. [27])[1],

$$F^{[\Gamma]}(x,\Delta;\lambda,\lambda') = \frac{1}{2}\int \frac{dz^-}{2\pi} e^{ik\cdot z}\langle p',\lambda'|\bar{\psi}(-\tfrac{z}{2})\,\Gamma\,\mathcal{W}(-\tfrac{z}{2},\tfrac{z}{2})\psi(\tfrac{z}{2})|p,\lambda\rangle\Big|_{z^+=0,\vec{z}_\perp=\vec{0}_\perp}. \quad (6.1)$$

In Eq. (6.1), Γ denotes a generic gamma matrix, and the Wilson line

$$\mathcal{W}(-\tfrac{z}{2},\tfrac{z}{2})\Big|_{z^+=0,\vec{z}_\perp=\vec{0}_\perp} = \mathcal{P}\exp\Big(-ig_s\int_{-\frac{z^-}{2}}^{\frac{z^-}{2}} dy^- A^+(0^+,y^-,\vec{0}_\perp)\Big), \quad (6.2)$$

ensures the color gauge invariance of the operator, where $\mathcal{P}$ indicates path-ordering and g_s the strong coupling constant. The incoming (outgoing) hadron state in Eq. (6.1) is characterized by the 4-momentum $p\,(p')$ and the helicity $\lambda\,(\lambda')$. In this context, common kinematic variables are defined as

$$P \equiv \frac{1}{2}(p+p')\,, \qquad \Delta \equiv p'-p\,, \qquad t\equiv\Delta^2\,, \qquad \xi \equiv \frac{p'^+ - p^+}{p'^+ + p^+} = -\frac{\Delta^+}{2P^+}\,. \quad (6.3)$$

We limit our discussion of the skewness ξ to $\xi \geq 0$, as ξ is non-negative for all known physical processes that allow access to GPDs. We work in a symmetric frame of reference where $\vec{P}_\perp = 0$ and take P^3 to be positive and large. The variable t is related to ξ and $\vec{\Delta}_\perp$ through

$$t = -\frac{1}{1-\xi^2}(4\xi^2M^2 + \vec{\Delta}_\perp^2)\,, \quad (6.4)$$

where M is the nucleon mass. Eq. (6.1) represents a leading-twist matrix element if Γ contains one plus-index. The eight corresponding quark GPDs are then defined via (see for instance Refs. [27, 517])

$$F^{[\gamma^+]}(x,\Delta;\lambda,\lambda') = \frac{1}{2P^+}\bar{u}(p',\lambda')\Big[\gamma^+H(x,\xi,t) + \frac{i\sigma^{+\mu}\Delta_\mu}{2M}E(x,\xi,t)\Big]u(p,\lambda), \quad (6.5)$$

[1]For a generic four-vector v we denote the Minkowski components by (v^0,v^1,v^2,v^3) and the light-cone components by $(v^+,v^-,\vec{v}_\perp)$, with $v^+ = \frac{1}{\sqrt{2}}(v^0+v^3)$, $v^- = \frac{1}{\sqrt{2}}(v^0-v^3)$ and $\vec{v}_\perp = (v^1,v^2)$.

$$F^{[\gamma^+\gamma_5]}(x,\Delta;\lambda,\lambda') = \frac{1}{2P^+}\bar{u}(p',\lambda')\Big[\gamma^+\gamma_5\widetilde{H}(x,\xi,t) + \frac{\Delta^+\gamma_5}{2M}\widetilde{E}(x,\xi,t)\Big]u(p,\lambda), \quad (6.6)$$

$$\begin{aligned} F^{[i\sigma^{j+}\gamma_5]}(x,\Delta;\lambda,\lambda') = -\frac{i\varepsilon^{-+ij}}{2P^+}\bar{u}(p',\lambda')\Big[& i\sigma^{+i}H_T(x,\xi,t) + \frac{P^+\Delta^i_\perp}{M^2}\widetilde{H}_T(x,\xi,t) \\ &+\frac{\gamma^+\Delta^i_\perp - \Delta^+\gamma^i_\perp}{2M}E_T(x,\xi,t) \\ &-\frac{P^+\gamma^i_\perp}{M}\widetilde{E}_T(x,\xi,t)\Big]u(p,\lambda)\,, \end{aligned} \quad (6.7)$$

where $u(p,\lambda)$ ($\bar{u}(p',\lambda')$) is the helicity spinor for the incoming (outgoing) hadron and $\sigma^{\mu\nu} = \frac{i}{2}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)$. We adopt the convention of $\varepsilon^{0123} = 1$. The quarks are unpolarized in the case of the vector GPDs H and E, longitudinally polarized for $\widetilde{H}$ and $\widetilde{E}$, and transversely polarized for H_T, E_T, $\widetilde{H}_T$ and $\widetilde{E}_T$. In Eq. (6.7), because of the relation $i\sigma^{\mu\nu}\gamma_5 = -\frac{1}{2}\epsilon^{\mu\nu\alpha\beta}\sigma_{\alpha\beta}$, one may also work with the matrix $i\sigma^{j+}$ (instead of $i\sigma^{j+}\gamma_5$) to define chiral-odd quark GPDs. A generic GPD depends upon the average longitudinal momentum fraction $x = \frac{k^+}{P^+}$, as well as ξ and t. By means of Eq. (6.4) one can instead consider light-cone GPDs as function of x, ξ and $\vec{\Delta}_\perp$, and these variables are chosen for the numerical evaluations of the GPDs in Section 6.3. It is noted that in general the support region in x for the light-cone GPDs is $-1 \le x \le 1$.

Quasi-GPDs, on the other hand, are defined through an equal-time spatial correlation function [82],

$$\begin{aligned} F_{\rm Q}^{[\Gamma]}(x,\Delta;\lambda,\lambda';P^3) &= \frac{1}{2}\int\frac{dz^3}{2\pi}e^{ik\cdot z} \\ &\times \langle p',\lambda'|\bar{\psi}(-\tfrac{z}{2})\,\Gamma\,\mathcal{W}_{\rm Q}(-\tfrac{z}{2},\tfrac{z}{2})\psi(\tfrac{z}{2})|p,\lambda\rangle\Big|_{z^0=0,\vec{z}_\perp=\vec{0}_\perp}, \end{aligned} \quad (6.8)$$

where the Wilson line is given by,

$$\mathcal{W}_{\rm Q}(-\tfrac{z}{2},\tfrac{z}{2})\Big|_{z^0=0,\vec{z}_\perp=\vec{0}_\perp} = \mathcal{P}\exp\Big(-ig_s\int_{-\frac{z^3}{2}}^{\frac{z^3}{2}}dy^3A^3(0,\vec{0}_\perp,y^3)\Big)\,. \quad (6.9)$$

For a given light-cone GPD, we consider two distinct definitions of its corresponding quasi-GPD. The counterparts of Eqs. (6.5), (6.6) and (6.7) are

$$F^{[\gamma^0]}(x,\Delta;\lambda,\lambda';P^3) = \frac{1}{2P^0}\bar{u}(p',\lambda')\Big[\gamma^0 H_{{\rm Q}(0)}(x,\xi,t;P^3)$$

$$+\frac{i\sigma^{0\mu}\Delta_\mu}{2M}E_{\mathrm{Q}(0)}(x,\xi,t;P^3)\Big]u(p,\lambda), \qquad (6.10)$$

$$F^{[\gamma^3\gamma_5]}(x,\Delta;\lambda,\lambda';P^3) = \frac{1}{2P^0}\bar{u}(p',\lambda')\Big[\gamma^3\gamma_5\widetilde{H}_{\mathrm{Q}(3)}(x,\xi,t;P^3) + \frac{\Delta^3\gamma_5}{2M}\widetilde{E}_{\mathrm{Q}(3)}(x,\xi,t;P^3)\Big]u(p,\lambda), \qquad (6.11)$$

$$\begin{aligned} F^{[i\sigma^{j0}\gamma_5]}(x,\Delta;\lambda,\lambda';P^3) = -\frac{i\varepsilon^{03ij}}{2P^0}\bar{u}(p',\lambda')\Big[&i\sigma^{3i}H_{T,\mathrm{Q}(0)}(x,\xi,t;P^3) \\ &+\frac{\gamma^3\Delta^i_\perp-\Delta^3\gamma^i_\perp}{2M}E_{T,\mathrm{Q}(0)}(x,\xi,t;P^3) \\ &+\frac{P^3\Delta^i_\perp}{M^2}\widetilde{H}_{T,\mathrm{Q}(0)}(x,\xi,t;P^3) \\ &-\frac{P^3\gamma^i_\perp}{M}\widetilde{E}_{T,\mathrm{Q}(0)}(x,\xi,t;P^3)\Big]u(p,\lambda)\,. \end{aligned} \qquad (6.12)$$

One can define $H_{\mathrm{Q}(3)}$ and $E_{\mathrm{Q}(3)}$ through Eq. (6.10) using the replacement $0 \to 3$ (see also Ref. [61]), while $\widetilde{H}_{\mathrm{Q}(0)}$ and $\widetilde{E}_{\mathrm{Q}(0)}$ are defined through Eq. (6.11) with $0 \leftrightarrow 3$. The chiral-odd quasi-GPDs $H_{T,\mathrm{Q}(3)}$, $E_{T,\mathrm{Q}(3)}$, $\widetilde{H}_{T,\mathrm{Q}(3)}$, and $\widetilde{E}_{T,\mathrm{Q}(3)}$ are defined through Eq. (6.12) with $0 \to 3$, with the exception that ε^{03ij} should be left as is. The factor $\frac{1}{P^0}$ on the r.h.s. of Eq. (6.11) (which appears counterintuitive due to the $\gamma^3\gamma_5$ projection) is necessary to be consistent with the definition of the corresponding helicity quasi-PDF, such that the definitions of all (sixteen) quasi-GPDs are consistent with the corresponding forward limits. It has been argued that the gamma matrices used in Eq. (6.10), Eq. (6.11) and Eq. (6.12) provide optimal behavior of the associated operators under renormalization [358,378]. By taking the forward limit of Eqs. (6.10)–(6.12) one recovers the most frequently used definitions (as of the writing of Ref. [62]) of the quasi-PDFs $f_{1,\mathrm{Q}(0)}$, $g_{1,\mathrm{Q}(3)}$ and $h_{1,\mathrm{Q}(0)}$. We will return to this point in Section 6.5.

We now briefly discuss the behavior of GPDs under the replacement $\xi \to -\xi$. Hermiticity implies that all light-cone GPDs but $\widetilde{E}_T$ are even functions of ξ, while $\widetilde{E}_T$ is an odd function of ξ [27, 517]. We find the exact same (model-independent) behavior for the corresponding quasi-GPDs. Exploiting the symmetry of quasi-GPDs under $\xi \to -\xi$ provides more statistics for lattice calculations [512].

Apart from the dependence on ξ and t, quasi-GPDs are functions of $x = \frac{k^3}{P^3}$. The latter variable is of course different from the average plus-momentum $\frac{k^+}{P^+}$ that

appears for light-cone GPDs, and it is not possible to relate these two momentum fractions in a model-independent manner. In Section 6.3.2 we study the impact of their difference in the cut-graph approach in the diquark spectator model. Note that the support region for the quasi-GPDs is given by $-\infty < x < \infty$. For the calculations we also utilize the relation $P^0 = \delta P^3$ where the variable

$$\delta \equiv \sqrt{1 + \frac{M^2 - t/4}{(P^3)^2}}\,. \tag{6.13}$$

is frequently used below. Moreover, one has $P \cdot \Delta = 0$, from which one can obtain $\Delta^0 = -2\xi P^3$.

6.2 Analytical Results in Scalar Diquark Model

In this section we present the analytical results in the SDM. The SDM for a relativistic spin-$\frac{1}{2}$ particle is specified through the Lagrange density

$$\mathcal{L}_{\rm SDM} = \bar{\Psi}\left(i\,\partial\!\!\!/ - M\right)\Psi + \bar{\psi}\left(i\,\partial\!\!\!/ - m_q\right)\psi + \frac{1}{2}\left(\partial_\mu\varphi\,\partial^\mu\varphi - m_s^2\right) + g\big(\bar{\Psi}\,\psi\,\varphi + \bar{\psi}\,\Psi\,\varphi\big)\,, \tag{6.14}$$

with $\partial\!\!\!/ = \partial_\mu\gamma^\mu$. In Eq. (6.14), Ψ denotes the (fermionic) hadron field, ψ the quark field, and φ the scalar diquark field, while m_q and m_s are the quark and diquark masses, respectively. For the hadron to be stable the masses need to satisfy the relation $M < m_s + m_q$. The main ingredient of the model is the hadron-quark-diquark vertex with the coupling constant g. In this framework one can carry out perturbative calculations. All the model results for PDFs discussed below are of $\mathcal{O}(g^2)$, which is the lowest nontrivial order. We do not consider virtual diagrams which contribute for $x = 1$ only. Diquark spectator models have been used frequently to study various aspects of the nucleon structure — see for instance Refs. [517–521]. Often, scalar and vector diquarks have been involved simultaneously in order to obtain distributions of both up quarks and down quarks in the nucleon. In addition, the nucleon-quark-diquark vertices have frequently been multiplied by form factors. By so doing one can eliminate UV divergences of parton correlation functions, and the model becomes more flexible due to additional parameters. On the other hand, the model then no longer follows from a Lagrange density. The first model calculation of quasi-PDFs was carried out in such a type of diquark model [449] (see also Ref. [450]). We ultimately

find that the results of our study are largely insensitive to the type of the diquark. The following sections will focus on the scalar diquarks, while the axial vector diquarks will be discussed in Section 6.4. We take the model as defined through Eq. (6.14) and use a cutoff for the transverse quark momenta.

6.2.1 Results for light-cone GPDs

We begin with the results for the light-cone GPDs. To the lowest nontrivial order in the SDM, the correlator in Eq. (6.1) takes the form

$$\begin{aligned} F^{[\Gamma]}(x,\Delta;\lambda,\lambda') &= \frac{i\,g^2}{2(2\pi)^4}\int dk^-\,d^2\vec{k}_\perp \\ &\times \frac{\bar{u}(p',\lambda')\left(\not{k}+\frac{\not{\Delta}}{2}+m_q\right)\Gamma\left(\not{k}-\frac{\not{\Delta}}{2}+m_q\right)u(p,\lambda)}{D_{\rm GPD}}\,, \end{aligned} \tag{6.15}$$

where g denotes the strength of the nucleon-quark-diquark vertex, and

$$D_{\rm GPD} = \left[\left(k+\frac{\Delta}{2}\right)^2 - m_q^2 + i\varepsilon\right]\left[\left(k-\frac{\Delta}{2}\right)^2 - m_q^2 + i\varepsilon\right]\left[(P-k)^2 - m_s^2 + i\varepsilon\right]. \tag{6.16}$$

The light-cone GPDs are derived through Gordon identities and the k^- integral is performed via contour integration. The result for the GPD H can be cast in the form

$$H(x,\xi,t) = \begin{cases} 0 & (-1\le x\le -\xi)\,, \\ \dfrac{g^2(x+\xi)(1+\xi)(1-\xi^2)}{4(2\pi)^3}\displaystyle\int d^2\vec{k}_\perp \frac{N_H}{D_1\,D_2^{-\xi\le x\le\xi}} & (-\xi\le x\le\xi)\,, \\ \dfrac{g^2(1-x)(1-\xi^2)}{2(2\pi)^3}\displaystyle\int d^2\vec{k}_\perp \frac{N_H}{D_1\,D_2^{x\ge\xi}} & (\xi\le x\le 1)\,, \end{cases} \tag{6.17}$$

and corresponding expressions hold for the other GPDs. The following is a compilation of the numerators of all the leading-twist light-cone GPDs in the SDM:

$$N_H = \vec{k}_\perp^{\,2} + (m_q+xM)^2 + (1-x)^2\,\frac{t}{4} - (1-x)\xi t\,\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^{\,2}}\,, \tag{6.18}$$

$$N_E = 2(1-x)M\left[m_q + \left(x + 2\xi\,\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^{\,2}}\right)M\right], \tag{6.19}$$

$$
\begin{aligned}
N_{\widetilde{H}} &= -\vec{k}_\perp^2 + (m_q + xM)^2 - (1-x)^2\,\frac{t}{4} \\
&+ \xi\big[4M(m_q + xM) + (1-x)t\big]\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\,, \qquad (6.20)\\
\xi N_{\widetilde{E}} &= 2M\Bigg[(1-x)\xi(m_q+M) \\
&+ 2\big[(1-\xi^2)m_q + (x-\xi^2)M\big]\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\Bigg]\,, \qquad (6.21)\\
N_{H_T} &= \vec{k}_\perp^2 - 2\frac{(\vec{k}_\perp\cdot\vec{\Delta}_\perp)^2}{\vec{\Delta}_\perp^2} + (m_q+xM)^2 - (1-x)^2\,\frac{t}{4} \\
&+ \xi\big[4M(m_q + xM) + (1-x)t\big]\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\,, \qquad (6.22)\\
N_{E_T} &= 2M\Bigg[4M\,\frac{\vec{k}_\perp^2\vec{\Delta}_\perp^2 - 2\,(\vec{k}_\perp\cdot\vec{\Delta}_\perp)^2}{(\vec{\Delta}_\perp^2)^2} \\
&+ \left(1 - x - 2\xi\,\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\right)(m_q+M)\Bigg]\,, \qquad (6.23)\\
N_{\widetilde{H}_T} &= -M^2\Bigg[4(1-\xi^2)\,\frac{\vec{k}_\perp^2\vec{\Delta}_\perp^2 - 2\,(\vec{k}_\perp\cdot\vec{\Delta}_\perp)^2}{(\vec{\Delta}_\perp^2)^2} \\
&+ (1-x)\left(1 - x - 4\xi\,\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\right)\Bigg]\,, \qquad (6.24)\\
N_{\widetilde{E}_T} &= 4M\Bigg[2\xi M\,\frac{\vec{k}_\perp^2\vec{\Delta}_\perp^2 - 2\,(\vec{k}_\perp\cdot\vec{\Delta}_\perp)^2}{(\vec{\Delta}_\perp^2)^2} - (m_q+xM)\,\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\Bigg]\,. \qquad (6.25)
\end{aligned}
$$

The denominators in Eq. (6.17) are given by

$$
\begin{aligned}
D_1 &= (1+\xi)^2\vec{k}_\perp^2 + \frac{1}{4}(1-x)^2\vec{\Delta}_\perp^2 - (1-x)(1+\xi)\vec{k}_\perp\cdot\vec{\Delta}_\perp \\
&+ (1-x)(1+\xi)m_q^2 + (x+\xi)(1+\xi)m_s^2 - (1-x)(x+\xi)M^2\,, \\
D_2^{-\xi\le x\le\xi} &= \xi(1-\xi^2)\vec{k}_\perp^2 + \frac{1}{4}(1-x^2)\xi\vec{\Delta}_\perp^2 + x(1-\xi^2)\vec{k}_\perp\cdot\vec{\Delta}_\perp \qquad (6.26)\\
&+ \xi(1-\xi^2)m_q^2 - \xi(x^2-\xi^2)M^2\,, \\
D_2^{x\ge\xi} &= (1-\xi)^2\vec{k}_\perp^2 + \frac{1}{4}(1-x)^2\vec{\Delta}_\perp^2 + (1-x)(1-\xi)\vec{k}_\perp\cdot\vec{\Delta}_\perp
\end{aligned}
$$

$$+ \quad (1-x)(1-\xi)m_q^2 + (x-\xi)(1-\xi)m_s^2 - \ (1-x)(x-\xi)M^2 \,.$$

The light-cone GPDs in the SDM can also be extracted from the results for the generalized transverse momentum dependent parton distributions listed in Ref. [29], which are consistent with the results above. The light-cone GPDs vanish for $-1 \leq x \leq -\xi$ due to the absence of antiquarks to $\mathcal{O}(g^2)$ in the SDM. We emphasize that the positions of the k^- poles in Eq. (6.16) depend on x. This leads to different analytical expressions for the light-cone GPDs in the ERBL and DGLAP regions. The GPDs remain continuous at the boundaries $x = \pm\,\xi$ between these regions (see also Ref. [61]), though their derivatives are discontinuous. Note also that spectator models typically lead to discontinuous higher twist GPDs [522, 523]. The GPD $\widetilde{E}$ exhibits a singularity as $\xi \to 0$ which is why we show $\xi\widetilde{E}$ in Eq. (6.21) and later on for the numerics.

Our model results must satisfy the symmetry behavior under the replacement $\xi \to -\xi$ discussed in Section 6.1 above. In order to verify that the results pass this test, it is necessary to replace the integration variable $\vec{k}_\perp$ with $-\vec{k}_\perp$. One then finds that the numerators in Eqs. (6.18)–(6.25) are indeed even under $\xi \to -\xi$ except the one for $\widetilde{E}_T$, which is odd under this transformation. The analysis of the denominators requires more care. In order to locate the position of the poles in the complex k^--plane, and hence to arrive at the above expressions of the light-cone GPDs, we have considered $\xi > 0$. Keeping this in mind, one can verify that $\xi \to -\xi$ switches the position of the poles of the quark propagators only such that the denominators in the ERBL and DGLAP regions are even in ξ. We also note that our analytical results for the quasi-GPDs below show the exact same behavior under $\xi \to -\xi$ as the respective light-cone GPDs.

In the SDM to $\mathcal{O}(g^2)$, all the leading-twist light-cone GPDs are UV-finite, except H and $\widetilde{H}$. We consider the fact that the chiral-odd GPD H_T is UV-finite to be an artifact of the SDM. In the quark-target model in perturbative QCD this function shows the well-known UV-divergence [501, 517]. For the numerics we impose a cut-off on the transverse quark momenta on all the light-cone GPDs as well as the (UV-finite) quasi-GPDs.

6.2.2 Results for quasi-GPDs

The quasi-GPD correlator in Eq. (6.8) in the SDM reads

$$\begin{aligned} F_{\rm Q}^{[\Gamma]}(x,\Delta;\lambda,\lambda';P^3) &= \frac{i\,g^2}{2(2\pi)^4}\int dk^0\,d^2\vec{k}_\perp \\ &\times \frac{\bar{u}(p',\lambda')\left(\not{k}+\frac{\not{\Delta}}{2}+m_q\right)\Gamma\left(\not{k}-\frac{\not{\Delta}}{2}+m_q\right)u(p,\lambda)}{D_{\rm GPD}}\,. \end{aligned} \tag{6.27}$$

We again use Gordon identities to obtain the quasi-GPDs. Before carrying out the k^0 integral one has

$$H_{\rm Q(0/3)}(x,\xi,t;P^3) = \frac{i\,g^2P^3}{(2\pi)^4}\int dk^0\,d^2\vec{k}_\perp\,\frac{N_{H(0/3)}}{D_{\rm GPD}}\,, \tag{6.28}$$

and corresponding expressions for the other quasi-GPDs. The numerators for the unpolarized quasi-GPDs $H_{\rm Q(0/3)}$ and $E_{\rm Q(0/3)}$, derived in [61], are given by:

$$\begin{aligned} N_{H(0)} &= \delta(k^0)^2 - \frac{2}{P^3}\left[x(P^3)^2 - m_qM - x\,\frac{t}{4} - \frac{1}{2}\,\delta\xi t\,\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\right]k^0 \\ &+ \delta\left[x^2(P^3)^2 + \vec{k}_\perp^2 + m_q^2 + (1-2x)\,\frac{t}{4} - \delta\xi t\,\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\right], \end{aligned} \tag{6.29}$$

$$\begin{aligned} N_{H(3)} &= -(k^0)^2 + \frac{2}{\delta P^3}\left[x\big((P^3)^2+M^2\big) - \frac{t}{4}\right]k^0 - x^2(P^3)^2 + \vec{k}_\perp^2 \\ &+ m_q\big(m_q+2xM\big) + \frac{t}{4} - (1-x)\,\frac{\xi t}{\delta}\,\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\,, \end{aligned} \tag{6.30}$$

$$N_{E(0)} = -2M\delta\left(m_q + xM + 2M\delta\xi\,\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\right)\left(\frac{k^0}{\delta P^3}-1\right), \tag{6.31}$$

$$N_{E(3)} = 2(1-x)M\left(\frac{M}{\delta P^3}\,k^0 + m_q + 2\,\frac{M\xi}{\delta}\,\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\right). \tag{6.32}$$

The numerators for the case of longitudinal quark polarization, corresponding to the quasi-GPDs $\widetilde{H}_{\rm Q(0/3)}$ and $\widetilde{E}_{\rm Q(0/3)}$ and derived in [62], are given by:

$$N_{\widetilde{H}(0)} = -(k^0)^2 + 2k^0\left[x\delta P^3 - 2\xi\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}(1-\delta^2)P^3\right] - \vec{k}_\perp^2 - \frac{t}{4}$$

$$
\begin{aligned}
&- \quad x^2(P^3)^2 + 2x(1-\delta^2)(P^3)^2 + 2xM(m_q+M) + m_q^2 \\
&+ \quad 4\delta\xi\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\Big[(1-\delta^2)(P^3)^2 + M(m_q+M)\Big], \qquad (6.33)\\
N_{\widetilde{H}(3)} &= \delta(k^0)^2 + 2\frac{k^0}{P^3}\Big[(1-x-\delta^2)(P^3)^2 + M(m_q+M)\Big] \\
&+ \quad \delta\Big[-\vec{k}_\perp^2 - \frac{t}{4} + x^2(P^3)^2 + m_q^2\Big] \\
&+ \quad 4\xi\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\Big[(1-x)(1-\delta^2)(P^3)^2 + M(m_q+M)\Big], \qquad (6.34)\\
\xi N_{\widetilde{E}(0)} &= 4k^0M^2\frac{1}{P^3}\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2} + 2M\xi(1-x)(m_q+M) \\
&- \quad 4\delta M\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\Big[M\xi^2 - m_q(1-\xi^2)\Big], \qquad (6.35)\\
\xi N_{\widetilde{E}(3)} &= -2k^0\frac{\xi}{P^3}M(m_q+M) + 2\delta\xi M(m_q+M) \\
&- \quad 4M\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\Big[M(\xi^2-x) - m_q(1-\xi^2)\Big]. \qquad (6.36)
\end{aligned}
$$

Note that the quasi-GPDs $\widetilde{E}_{\mathrm{Q}(0/3)}$ have a pole at $\xi = 0$, just like their light-cone counterpart. We next list the numerators of the quasi-GPDs that appear for transverse quark polarization derived in [62]:

$$
\begin{aligned}
N_{H_T(0)} &= \delta(k^0)^2 - \frac{k^0}{P^3}\Big[\frac{\vec{\Delta}_\perp^2}{2} - 2m_qM + \Big(2x - 2\xi^2(1-\delta^2)\Big)(P^3)^2\Big] \\
&+ \quad \delta\Big[\vec{k}_\perp^2 - 2\frac{(\vec{k}_\perp\cdot\vec{\Delta}_\perp)^2}{\vec{\Delta}_\perp^2} + \frac{\vec{\Delta}_\perp^2}{4} + m_q^2\Big] + \delta\Big[x^2 - \xi^2(1-\delta^2)\Big](P^3)^2 \\
&- \quad 4\xi\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\Big[(x-\xi^2)(1-\delta^2)(P^3)^2 - m_qM + \frac{\vec{\Delta}_\perp^2}{4}\Big], \qquad (6.37)\\
N_{H_T(3)} &= -(k^0)^2 - 2k^0\Big[2\xi(1-\delta^2)P^3\,\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\Delta_\perp^2} - x\delta P^3\Big]\vec{k}_\perp^2 - 2\frac{(\vec{k}_\perp\cdot\vec{\Delta}_\perp)^2}{\vec{\Delta}_\perp^2} \\
&+ \quad (1-2x)\frac{\vec{\Delta}_\perp^2}{4} + m_q^2 + 2xm_qM - (P^3)^2\Big[x^2 + (1-2x)\xi^2(1-\delta^2)\Big]
\end{aligned}
$$

$$
\begin{aligned}
&- \quad 4\delta\xi\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\Delta_\perp^2}\left[\frac{\vec{\Delta}_\perp^2}{4} - m_q M - \xi^2(1-\delta^2)(P^3)^2\right], && (6.38)\\
N_{\widetilde{H}_T(0)} &= \delta N_{\widetilde{H}_T(3)}\,, && (6.39)\\
N_{\widetilde{H}_T(3)} &= -4M^2(1-\xi^2)\frac{\vec{k}_\perp^2\vec{\Delta}_\perp^2 - 2\,(\vec{k}_\perp\cdot\vec{\Delta}_\perp)^2}{(\vec{\Delta}_\perp^2)^2}\\
&- \quad \frac{k^0}{\delta P^3}\left[2\delta\xi\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2} - (1-x)\right]M^2\\
&+ \quad 2\frac{\xi}{\delta}\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\left[(1-x)+\delta^2\right]M^2 - (1-x)M^2\,, && (6.40)\\
N_{E_T(0)} &= 8\delta M^2\frac{\vec{k}_\perp^2\vec{\Delta}_\perp^2 - 2\,(\vec{k}_\perp\cdot\vec{\Delta}_\perp)^2}{(\vec{\Delta}_\perp^2)^2} - 2\delta M(m_q+M)\left[\frac{k^0}{\delta P^3}-1\right]\\
&- \quad 4\xi M(m_q+M)\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\,, && (6.41)\\
N_{E_T(3)} &= 8M^2\frac{\vec{k}_\perp^2\vec{\Delta}_\perp^2 - 2\,(\vec{k}_\perp\cdot\vec{\Delta}_\perp)^2}{(\vec{\Delta}_\perp^2)^2} + 2(1-x)M(m_q+M)\\
&- \quad 4\delta\xi M(m_q+M)\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\,, && (6.42)\\
N_{\widetilde{E}_T(0)} &= 8\delta^2\xi M^2\frac{\vec{k}_\perp^2\vec{\Delta}_\perp^2 - 2\,(\vec{k}_\perp\cdot\vec{\Delta}_\perp)^2}{(\vec{\Delta}_\perp^2)^2} - 4\delta M(m_q+xM)\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\,, && (6.43)\\
N_{\widetilde{E}_T(3)} &= 8\frac{\xi}{\delta}M^2\frac{\vec{k}_\perp^2\vec{\Delta}_\perp^2 - 2\,(\vec{k}_\perp\cdot\vec{\Delta}_\perp)^2}{(\vec{\Delta}_\perp^2)^2} - 4M(m_q+\frac{k^0}{\delta P^3}M)\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\,. && (6.44)
\end{aligned}
$$

The quasi-GPDs $\widetilde{H}_{T,\mathrm{Q}(0)}$ and $\widetilde{H}_{T,\mathrm{Q}(3)}$ corresponding to two different Dirac structures are related through Eq. (6.39). This is the only quasi-GPD whose two different projections have such a simple relation. We repeat that all quasi-GPDs have support in the range $-\infty < x < \infty$. However, for large P^3 they are all power-suppressed outside the region $-\xi \le x \le 1$. We also observe that the numerators of the quasi-GPDs $E_{T,\mathrm{Q}(3)}$ and $\widetilde{E}_{T,\mathrm{Q}(0)}$ are the only ones that do not depend on k^0.

The denominator D_{GPD} can be written as

$$
D_{\mathrm{GPD}} = (k^0-k^0_{1+})(k^0-k^0_{1-})(k^0-k^0_{2+})(k^0-k^0_{2-})(k^0-k^0_{3+})(k^0-k^0_{3-})\,, \quad (6.45)
$$

where the poles from the quark propagators, with 4-momenta $(k-\frac{\Delta}{2})$ and $(k+\frac{\Delta}{2})$,

and from the spectator propagator are given by

$$k^0_{1\pm} = -\xi P^3 \pm \sqrt{(x+\delta\xi)^2(P^3)^2 + \left(\vec{k}_\perp - \frac{\vec{\Delta}_\perp}{2}\right)^2 + m_q^2 - i\varepsilon}\,, \tag{6.46}$$

$$k^0_{2\pm} = \xi P^3 \pm \sqrt{(x-\delta\xi)^2(P^3)^2 + \left(\vec{k}_\perp + \frac{\vec{\Delta}_\perp}{2}\right)^2 + m_q^2 - i\varepsilon}\,, \tag{6.47}$$

$$k^0_{3\pm} = \delta P^3 \pm \sqrt{(1-x)^2(P^3)^2 + \vec{k}_\perp^{\,2} + m_s^2 - i\varepsilon}\,. \tag{6.48}$$

It is important to realize that while the positions of the poles depend on x, they never switch half planes. Specifically, k^0_{1-}, k^0_{2-} and k^0_{3-} always lie in the upper half plane, while the other three poles lie in the lower half plane. After performing the k^0 integral, one therefore has the same functional form for the quasi-GPDs for any x, which implies that all quasi-GPDs are continuous as a function of x — in this context, see also Ref. [61]. We have checked that for $P^3 \to \infty$ the analytical results of all quasi-GPDs reduce to the ones of the respective light-cone GPDs, the details of which can be found in Appendix B.

6.2.3 Results for quasi-PDFs

Starting from the expressions of the light-cone GPDs and taking $\Delta = 0$ (which implies $\xi = t = 0$), one obtains the following expressions for the light-cone PDFs:

$$\begin{aligned} f_1(x) = H(x,0,0) &= \frac{g^2(1-x)}{2(2\pi)^3}\int d^2\vec{k}_\perp \\ &\times \frac{\vec{k}_\perp^{\,2} + (m_q + xM)^2}{[\vec{k}_\perp^{\,2} + xm_s^2 + (1-x)m_q^2 - x(1-x)M^2]^2}\,, \end{aligned} \tag{6.49}$$

$$\begin{aligned} g_1(x) = \tilde{H}(x,0,0) &= \frac{g^2(1-x)}{2(2\pi)^3}\int d^2\vec{k}_\perp \\ &\times \frac{-\vec{k}_\perp^{\,2} + (m_q + xM)^2}{[\vec{k}_\perp^{\,2} + xm_s^2 + (1-x)m_q^2 - x(1-x)M^2]^2}\,, \end{aligned} \tag{6.50}$$

$$h_1(x) = H_T(x,0,0) = \frac{g^2(1-x)}{2(2\pi)^3}\int d^2\vec{k}_\perp$$

$$\times \quad \frac{(m_q + xM)^2}{[\vec{k}_\perp^2 + xm_s^2 + (1-x)m_q^2 - x(1-x)M^2]^2}\,. \qquad (6.51)$$

Only three GPDs survive in this limit — E, $\widetilde{E}$, E_T and $\widetilde{H}_T$ vanish because Δ appears in their prefactor in the parameterizations in Eqs. (6.5), (6.6) and (6.7), while $\widetilde{E}_T$ drops out since $\bar{u}'\gamma_T^i u$ vanishes in the forward limit. The GPD H reduces to the unpolarized PDF f_1, whereas $\widetilde{H}$ reduces to the helicity PDF g_1 and H_T to the transversity PDF h_1. Our results for the forward PDFs agree with the ones published in Ref. [517]. In general, like for light-cone GPDs, the region of support for PDFs is $-1 \le x \le 1$. In the SDM to $\mathcal{O}(g^2)$, they also vanish for $-1 \le x < 0$. In [62] we give a separate discussion for the point $x = 0$, where the forward PDFs in the SDM are discontinuous.

For the quasi-PDFs one has

$$f_{1,\mathrm{Q}(0/3)}(x; P^3) = \frac{ig^2 P^3}{(2\pi)^4} \int dk^0\, d^2\vec{k}_\perp\, \frac{N_{f1(0/3)}}{D_{\mathrm{PDF}}}\,, \qquad (6.52)$$

and corresponding expressions for the other quasi-PDFs. The numerators are given by

$$N_{f1(0)} = \delta_0 (k^0)^2 - \frac{2k^0}{P^3}\Big(x(P^3)^2 - m_q M\Big) + \delta_0\Big(\vec{k}_\perp^2 + x^2(P^3)^2 + m_q^2\Big)\,, \qquad (6.53)$$

$$N_{f1(3)} = -(k^0)^2 + k^0\Big(2x\delta_0 P^3\Big) + \vec{k}_\perp^2 - x^2(P^3)^2 + m_q\Big(m_q + 2xM\Big)\,, \qquad (6.54)$$

$$N_{g1(0)} = -(k^0)^2 + k^0\Big(2x\delta_0 P^3\Big) - \vec{k}_\perp^2 - x^2(P^3)^2 + m_q\Big(m_q + 2xM\Big)\,, \qquad (6.55)$$

$$N_{g1(3)} = \delta_0 (k^0)^2 - \frac{2k^0}{P^3}\Big(x(P^3)^2 - m_q M\Big) + \delta_0\Big(-\vec{k}_\perp^2 + x^2(P^3)^2 + m_q^2\Big)\,, \qquad (6.56)$$

$$N_{h1(0)} = \delta_0 (k^0)^2 - \frac{2k^0}{P^3}\Big(x(P^3)^2 - m_q M\Big) + \delta_0\Big(x^2(P^3)^2 + m_q^2\Big)\,, \qquad (6.57)$$

$$N_{h1(3)} = -(k^0)^2 + k^0\Big(2x\delta_0 P^3\Big) - x^2(P^3)^2 + m_q\Big(m_q + 2xM\Big)\,, \qquad (6.58)$$

and the denominator reads

$$D_{\mathrm{PDF}} = \big[k^2 - m_q^2 + i\varepsilon\big]^2\,\big[(P-k)^2 - m_s^2 + i\varepsilon\big]\,. \qquad (6.59)$$

In Eqs. (6.53)–(6.58) we have used $\delta_0 \equiv \delta(t=0)$. Like for quasi-GPDs, the support range of quasi-PDFs is $-\infty < x < \infty$. Results for the quasi-PDFs associated with

the gamma matrices $\gamma^3, \gamma^3\gamma_5$, and $i\sigma^{j3}\gamma_5$ were already presented in [449], but in the so-called cut-graph approximation. The full calculation with all contributions will be compared to the cut-graph approximation below. Note that we have calculated all the forward distributions independently using a trace technique, and have found complete agreement with the results obtained from the quasi-GPDs.

We now explicitly carry out some of the details of the contour integration for the k^0 integral in Eq. (6.52), where the poles are given by Eqs. (6.46) – (6.48) with $\xi = t = 0$. In the forward limit one has double poles at $k^0_{1\pm} = k^0_{2\pm}$. Closing the integration contour in the upper half plane gives contributions from the pole at k^0_{3-} and the double pole at $k^0_{1-} = k^0_{2-}$. In the case of $f_{1,Q(0)}$ the result of the k^0 integration reads

$$\begin{aligned} f_{1,Q(0)}(x,P^3) = -\frac{g^2P^3}{(2\pi)^3}\int d^2\vec{k}_\perp \Bigg[& \frac{N_{f1(0)}(k^0_{3-})}{(k^0_{3-}-k^0_{1+})^2(k^0_{3-}-k^0_{1-})^2(k^0_{3-}-k^0_{3+})} \\ & + \frac{N'_{f1(0)}(k^0_{1-})}{(k^0_{1-}-k^0_{1+})^2(k^0_{1-}-k^0_{3+})(k^0_{1-}-k^0_{3-})} \\ & - \frac{2N_{f1(0)}(k^0_{1-})}{(k^0_{1-}-k^0_{1+})^3(k^0_{1-}-k^0_{3+})(k^0_{1-}-k^0_{3-})} \\ & - \frac{N_{f1(0)}(k^0_{1-})}{(k^0_{1-}-k^0_{1+})^2(k^0_{1-}-k^0_{3+})^2(k^0_{1-}-k^0_{3-})} \\ & - \frac{N_{f1(0)}(k^0_{1-})}{(k^0_{1-}-k^0_{1+})^2(k^0_{1-}-k^0_{3+})(k^0_{1-}-k^0_{3-})^2}\Bigg], \end{aligned} \tag{6.60}$$

where in one of the terms the derivative $N'_{f1(0)} \equiv \frac{d}{dk^0}N_{f1(0)}$ enters. For $P^3 \to \infty$ one can recover the light-cone PDF f_1 in Eq. (6.49) by using the expression in Eq. (6.60). In this limit, in the region $0 \le x \le 1$ only the first term in the square brackets of Eq. (6.60) is leading. For $x > 1$ all terms are power-suppressed, while for $x < 0$ the first and last term are leading but the leading powers of the two terms cancel each other.

6.2.4 Cut-diagram Approach

In order to compute light-cone PDFs (and GPDs for $\xi = 0$) in diquark spectator models one can use a cut-diagram approach with a single on-shell particle (diquark) [517–521]. In this framework, one inserts in the PDF operator a sum over a complete set of states between the quark fields and, for the calculation of real graphs to $\mathcal{O}(g^2)$,

restricts this sum to a single diquark. One can verify that this technique provides the same result one finds by computing the correlator without inserting a complete set of states right from the start and then performing the k^- integration. On the other hand, care has to be taken for quasi-PDFs. To illustrate this point we consider as an example $f_{1,Q(0)}$ in the cut-diagram approach. One finds[2]

$$\begin{aligned} f_{1,Q(0),\mathrm{cut}}(x,P^3) &= \frac{g^2}{2(2\pi)^4}\int dk^0 d^2\vec{k}_\perp (2\pi)\delta((P-k)^2-m_s^2)\Theta(P^0-k^0) \\ &\times \frac{\bar{u}(P)(\not{k}+m_q)\gamma^0(\not{k}+m_q)u(P)}{[k^2-m_q^2+i\epsilon][k^2-m_q^2-i\epsilon]}, \end{aligned} \tag{6.61}$$

where the delta function and theta function ensure the on-shell diquark with positive energy. Working out the numerator in Eq. (6.61) and using

$$\delta((P-k)^2-m_s^2)\Theta(P^0-k^0) = \frac{1}{k_{3+}^0-k_{3-}^0}\delta(k^0-k_{3-}^0), \tag{6.62}$$

provides the result

$$\begin{aligned} f_{1,Q(0),\mathrm{cut}}(x,P^3) &= \frac{g^2P^3}{(2\pi)^3}\int d^2\vec{k}_\perp \frac{1}{k_{3+}^0-k_{3-}^0}\frac{N_{f1(0)}(k^0)}{[k^2-m_q^2+i\epsilon][k^2-m_q^2-i\epsilon]}\bigg|_{k^0=k_{3-}^0} \\ &= -\frac{g^2P^3}{(2\pi)^3}\int d^2\vec{k}_\perp \frac{N_{f1(0)}(k_{3-}^0)}{(k_{3-}^0-k_{1+}^0)^2(k_{3-}^0-k_{1-}^0)^2(k_{3-}^0-k_{3+}^0)}. \end{aligned} \tag{6.63}$$

Obvious modifications to Eq. (6.63) give the results for the other quasi-PDFs. This expression exactly agrees with the first term on the r.h.s. of Eq. (6.60), while the other four terms are missing. The discussion in the paragraph after Eq. (6.60) also implies that, for $P^3 \to \infty$, one can recover the light-cone PDF for $x \geq 0$, but not for $x < 0$, from the result in Eq. (6.63). In the case of quasi-PDFs, the cut-diagram approach [449, 450] is therefore a purely phenomenological model that could be used for $x \geq 0$.

[2]In the cut-diagram approach the sign of the $i\epsilon$ term in one of the quark propagators is different from Eq. (6.59). But the difference does not matter as the point $k^2 = m_q^2$ is not reached in this method.

6.3 Numerical Results in Scalar Diquark Model

In order to generate numerical results we must first specify the values of the input parameters to the SDM. We use $g = 1$ for the strength of the nucleon-quark-diquark coupling, with none of the general conclusions depending on the precise value of g. Our "standard values" for the mass parameters are $M = 0.939\,\mathrm{GeV}$, $m_s = 0.7\,\mathrm{GeV}$ and $m_q = 0.35\,\mathrm{GeV}$. Elaborating on our choice of parameters, we mention that our starting point comes from Ref. [521] where the value $m_q = 0.3\,\mathrm{GeV}$ was chosen and the value $m_s = 0.822\,\mathrm{GeV}$ was found from the phenomenological fits of Refs. [524, 525]. We have therefore chosen values similar to these, but have adjusted so that the convergence of the quasi distributions to the light-cone distributions is maximal. From Fig. 4 in [61], one can see that the relative difference between the quasi and light-cone f_1 is quite sensitive to adjustments in m_s. Thus it was important to reduce the value of m_s to $m_s = 0.7\,\mathrm{GeV}$ to achieve the best convergence. One can also see that the convergence is barely affected by changes in m_q, so we increase the value of m_q to $m_q = 0.35\,\mathrm{GeV}$ to satisfy $m_s + m_q > M$. In short, after exploring the sensitivity of our results to variations in m_s and m_q, we maintain that such a choice of the parameters, as discussed at length in Ref. [61], is "optimal" with regard to the question of convergence of the quasi-distributions to their light-cone counterparts. For most of our plots, the cut-off for the $|\vec{k}_\perp|$ integration is $\Lambda = 1\,\mathrm{GeV}$, and the transverse momentum transfer is $|\vec{\Delta}_\perp| = 0$. We also shall show some plots and comment extensively on the dependence of the various distributions on Λ and $|\vec{\Delta}_\perp|$. We begin with discussing the PDFs.

6.3.1 Results for quasi-PDFs

Results for the quasi-PDFs $f_{1,\mathrm{Q}(0/3)}$, $g_{1,\mathrm{Q}(0/3)}$, and $h_{1,\mathrm{Q}(0/3)}$ are shown in Fig. 6.1, Fig. 6.2, and Fig. 6.3, respectively. Comparing the three plots, one qualitatively observes the following features between them. First, for $P^3 = 2\,\mathrm{GeV}$ and above, there is not much difference between the two quasi-PDF definitions. Second, considerable differences appear between quasi-PDFs and light-cone PDFs as $x \to 0$ and $x \to 1$. As discussed in detail in Ref. [61], the discrepancy at small x is to be expected since the light-cone PDFs are discontinuous at $x = 0$. The quasi-PDFs are continuous, but for large P^3 must approach the corresponding light-cone PDF, which automatically results in large deviations in the region around $x = 0$. To better illustrate the

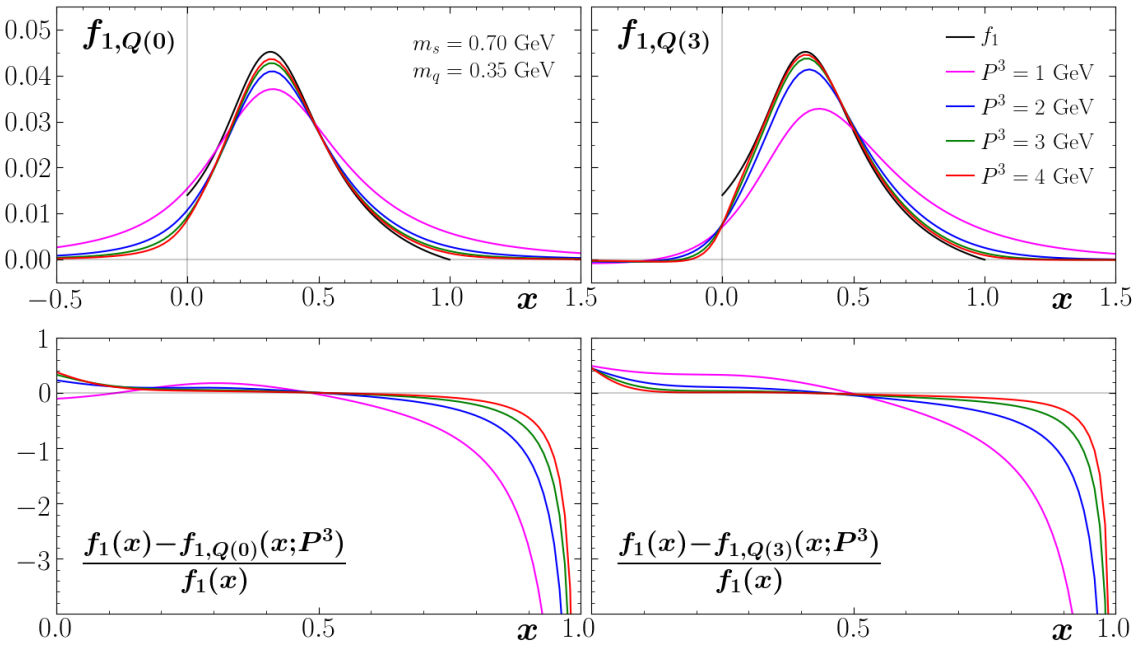

Figure 6.1: Quasi-PDF $f_{1,Q}$. It is shown as a function of x for different values of P^3 and with $m_s = 0.70$ GeV and $m_q = 0.35$ GeV. The top row shows the absolute values of $f_{1,Q(0)}$ (left) and $f_{1,Q(3)}$ (right) with the light-cone PDF f_1 shown for comparison. The bottom row shows the relative difference of the quasi-PDFs to the light-cone PDF.

discrepancy at large x we consider the relative difference, which in the case of f_1 we define as [61]

$$R_{f1(0/3)}(x; P^3) = \frac{f_1(x) - f_{1,\mathrm{Q}(0/3)}(x; P^3)}{f_1(x)}\,, \tag{6.64}$$

and analogously for all other quasi-PDFs and quasi-GPDs. This quantity is also shown in Figs. (6.1)–(6.3). For all three quasi-PDFs at $P^3 = 2\,\mathrm{GeV}$ one can hardly go above $x = 0.8$ for the relative difference to stay below 50%. We shed some more light on the origin of the large-x discrepancy in Section 6.3.2.

In Fig. 6.4 we show, as an example, the dependence of f_1 on the mass parameters m_s and m_q. Our general findings in this context can be summarized as follows. The impact of changing m_s is typically larger. Specifically, discrepancies get somewhat larger when increasing m_s, especially in the large-x region. This feature is partly related to the fact that increasing m_s increases the difference between the momentum fractions that enter the light-cone PDFs and quasi-PDFs. We refer to Section 6.3.2 for further discussion of this point.

Within the range $m_q \in [0.01,\ 0.35]\,\mathrm{GeV}$ which we have explored, we find only

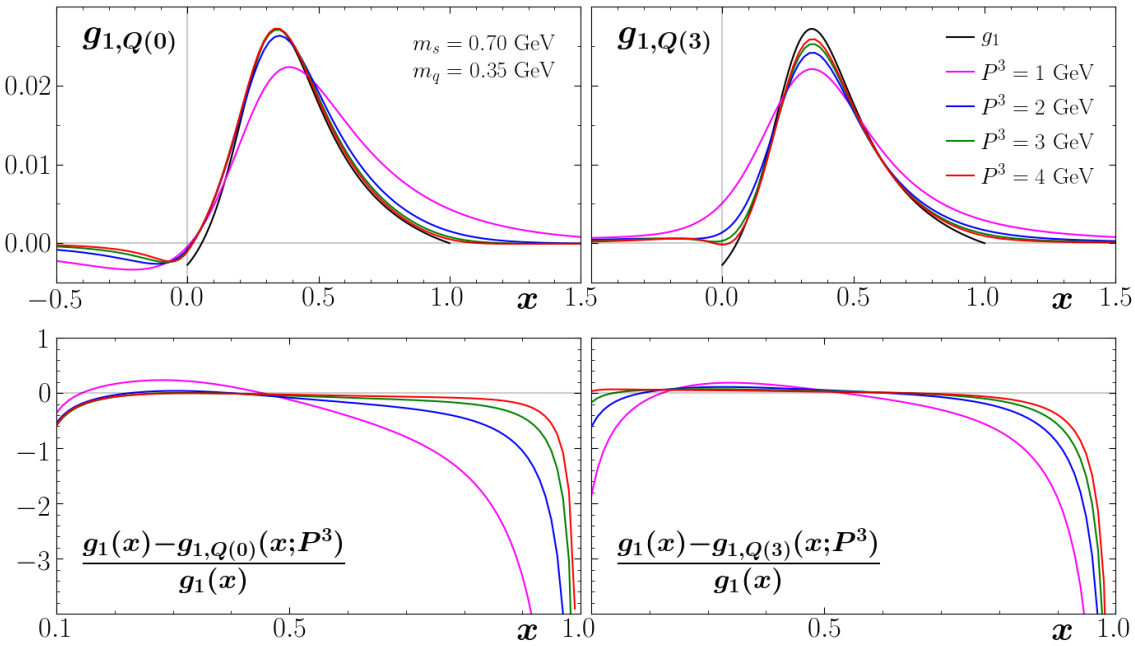

Figure 6.2: Quasi-PDF $g_{1,Q}$. Same as Fig. 6.1, except for the quasi-PDF $g_{1(Q)}$. Note that in the bottom row the relative difference is only shown down to $x = 0.1$ due to the pole in g_1.

a mild dependence on m_q. Analytically, this is caused by the fact that m_q is small compared to the other scales in the problem such as M, m_s, P^3, and Λ. Transversity is the only exception with regard to the m_q dependence especially in the small-x region. This can be understood from the analytical result in Eq. (6.51). For small x, the quark mass term in the numerator dominates resulting in a larger sensitivity to m_q of this distribution compared to f_1 and g_1. The latter distributions have a $\vec{k}_\perp^2$ in the numerator — in addition to the $(m_q + xM)^2$ term — which gives rise to the (standard) logarithmic UV-divergence and, in particular, a very mild dependence on m_q. As already discussed above, the absence of the UV divergence for the transversity is an artifact of the model, and therefore so is the stronger dependence of h_1 on m_q at small x. For the GPDs we find a very similar overall pattern upon variation of m_s and m_q. In the ERBL region there can be some deviations from this pattern. But the effects are not very significant, and we therefore refrain from further elaborating on them.

In Fig. 6.5, we show the relative difference for f_1 for two values (1 GeV and 4 GeV) of the cut-off Λ for the $k_\perp$-integration. For $x \lesssim 0.5$ the relative difference increases

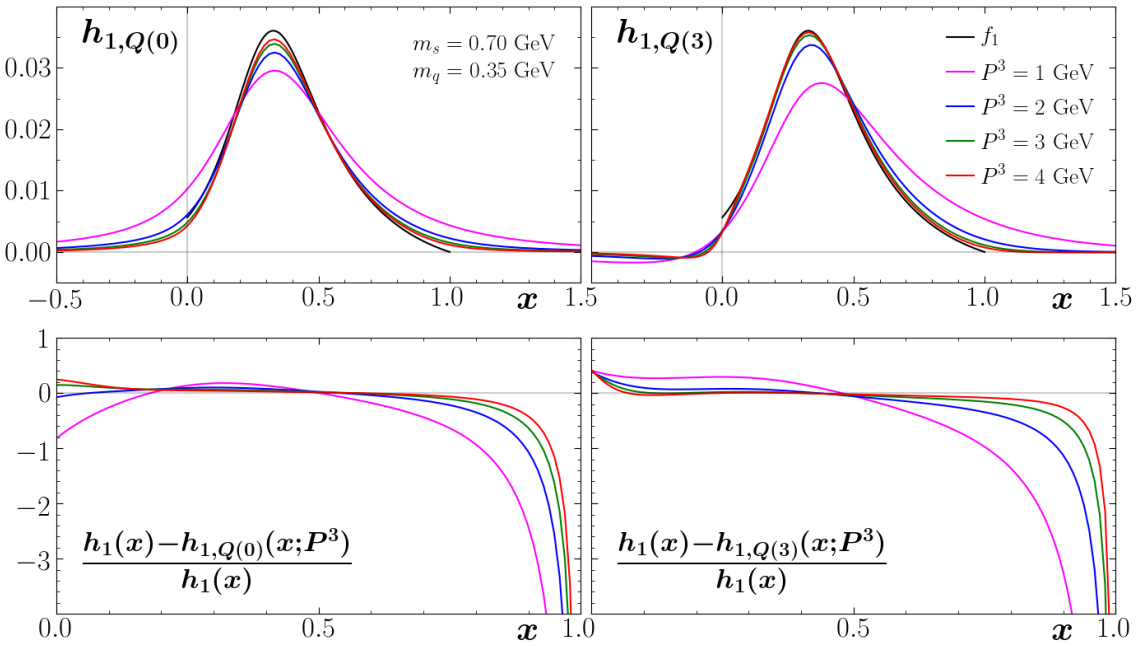

Figure 6.3: Quasi-PDF $h_{1,Q}$. Same as Fig. 6.1, except for the quasi-PDF $h_{1(Q)}$.

with an increase of Λ. But at least for $f_{1,\mathrm{Q(3)}}$ this effect is mild, given that the two values of Λ are very different. We find very similar results for the transversity distribution. On the other hand, for g_1 the impact (on the relative difference) of changing Λ is larger. This applies in particular in the region around the point at which g_1 changes sign — see Fig. 6.2. It is obvious from the definition in Eq. (6.64) that in such a case the relative difference is not a very good measure. A very similar situation occurs for GPDs if they switch sign. Overall, our choice $\Lambda = 1\,\mathrm{GeV}$ typically minimizes the difference between the quasi distributions and the light-cone distributions.

Also, the fact that some of the light-cone distributions have a logarithmic divergence does not necessarily lead to a much poorer convergence as Λ increases, unless one considers cut-off values much larger than 4 GeV. Thus our model can give a faithful description of these distributions with $1\,\mathrm{GeV} < \Lambda < 4\,\mathrm{GeV}$.

6.3.2 A particular higher twist contribution in the cut-diagram approximation

We repeat that the two momentum fractions $\frac{k^+}{P^+}$ and $\frac{k^3}{P^3}$ are different and that they cannot be related in a model-independent way. In this section we denote the latter by

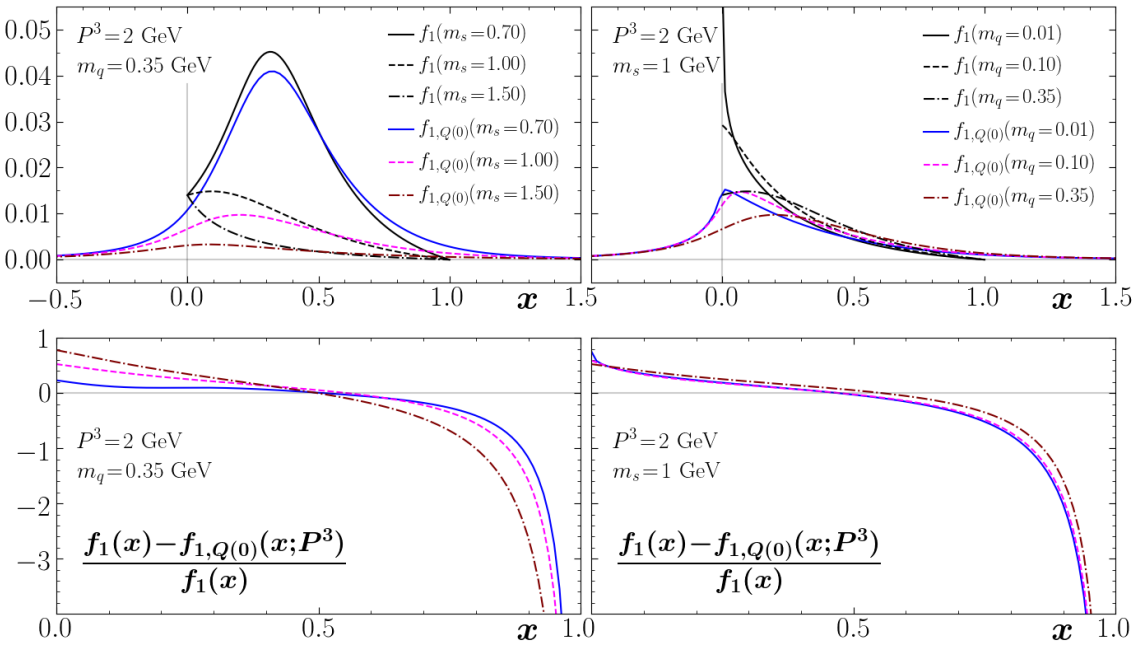

Figure 6.4: Quasi-PDF $f_{1,Q}$ dependence on m_s and m_q. It is shown as a function of x at $P^3 = 2$ GeV. On the left, $m_q = 0.35$ GeV is held constant while m_s is varied with values $0.70, 1.00, 1.50$ GeV. On the right, $m_s = 1$ GeV is held constant while m_q is varied with values $0.01, 0.10, 0.35$ GeV. The light-cone PDF f_1 is shown for comparison. The bottom row shows the relative difference of the quasi-PDF to the light-cone PDF with the various combinations of m_s and m_q.

$\tilde{x}$, and study the impact of the difference between x and $\tilde{x}$ in the (model-dependent) cut-graph approach in the SDM. In Fig. 6.6 we compare the full results for the quasi-PDF $f_{1,Q}$ with those obtained in the cut-diagram approach — see the discussion in Subsection 6.2.4. In the case of $f_{1,Q(0)}$ the analytical expressions are listed in Eq. (6.61) and Eq. (6.63). For positive and not too small x, making the approximation of keeping the spectator pole only does not have much influence. But more deviations occur as $x \to 0$, and the quasi-PDFs computed in the cut-diagram approach actually get closer to f_1. On the other hand, this method cannot be used for $x < 0$. We repeat that, even for large P^3, in the negative x region $f_{1,Q(0/3)}^{\text{spec.pole}}$ does not tend to zero.

In the cut-graph model one puts the di-quark spectator on-shell, that is, $(P-k)^2 = m_s^2$ (see Eq. (6.16)). One can then derive the relation

$$\tilde{x} \;=\; \frac{x}{2}(1+\delta_0) + \frac{\vec{k}_\perp^{\,2} + m_s^2 - (1-x)M^2}{2(1-x)(1+\delta_0)(P^3)^2}$$

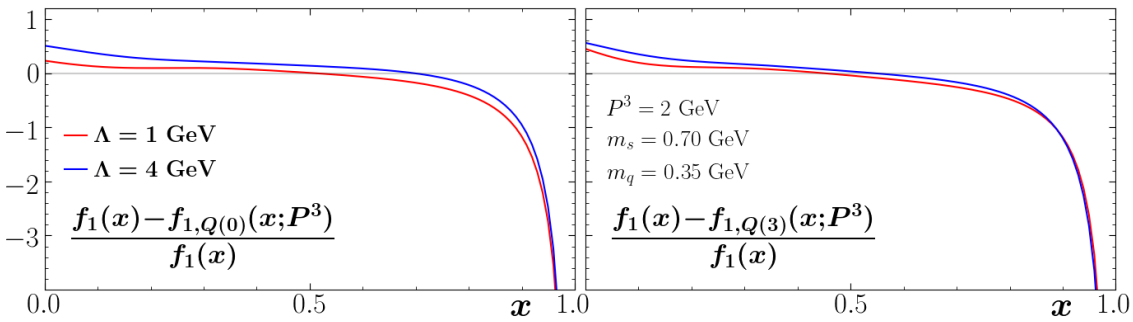

Figure 6.5: Quasi-PDF $f_{1,Q}$ dependence on Λ. Relative difference of the quasi-PDF $f_{1,Q}$ to the light-cone PDF f_1 as a function of x for different values of the cut-off Λ for the $k_\perp$-integration. The results are shown at $P^3 = 2$ GeV and with $m_s = 0.70$ GeV and $m_q = 0.35$ GeV, with $f_{1,Q(0)}$ on the left and $f_{1,Q(3)}$ on the right.

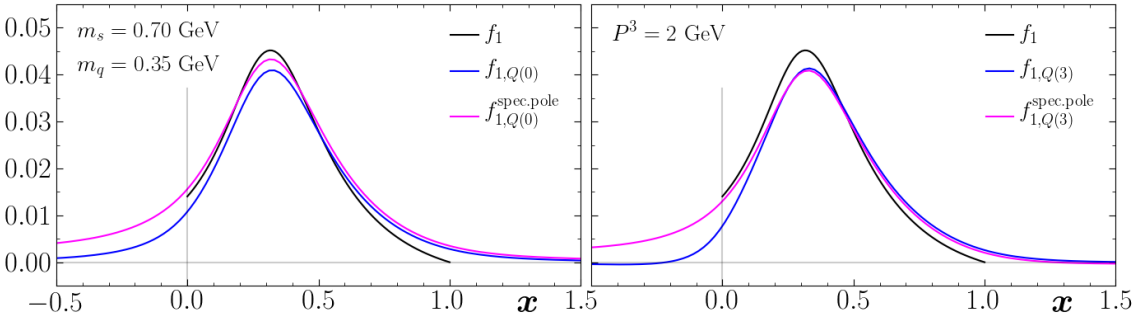

Figure 6.6: Quasi-PDF $f_{1,Q}$ in cut-diagram approach. Comparison between the full quasi-PDF $f_{1,Q}$ and the quasi-PDF in the cut-diagram approach $f_{1,Q}^{\text{spec.pole}}$ at $P^3 = 2$ GeV and with $m_s = 0.70$ GeV and $m_q = 0.35$ GeV. The quasi-PDF $f_{1,Q(0)}$ is shown on the left while $f_{1,Q(3)}$ is shown on the right. The light-cone PDF f_1 is shown for comparison.

$$= \quad x + \frac{1}{4(P^3)^2}\left(\frac{\vec{k}_\perp^{\,2} + m_s^2}{1-x} - (1-x)M^2\right) + \mathcal{O}\left(\frac{1}{(P^3)^4}\right). \tag{6.65}$$

Obviously, the difference between $\tilde{x}$ and x is of order $\mathcal{O}(1/(P^3)^2)$ and is therefore power-suppressed. A numerical comparison of the two variables can be found in Fig. 6.7. Their difference gets larger as m_s increases, as can also be expected based on Eq. (6.65). Most importantly, due to the $1/(1-x)$ factor, one finds $\tilde{x} \to \infty$ as $x \to 1$, which implies very large differences between the two momentum fractions at large x — see also Ref. [449]. One can therefore speculate that the considerable discrepancies between the quasi-distributions and the corresponding light-cone distributions at large x are mostly caused by the (huge) discrepancy between $\tilde{x}$ and x. In Fig. 6.8 we

explore this point for f_1. The quasi-PDF $f_{1,\mathrm{Q}(0)}$ indeed provides, at large x, a better agreement with the light-cone PDF, while this is not true for $f_{1,\mathrm{Q}(3)}$, unless one goes to extremely large x. In the case of g_1 and h_1 (not shown) we find that the "recipe" of distinguishing between $\tilde{x}$ and x works better for $g_{1,\mathrm{Q}(3)}$ and $h_{1,\mathrm{Q}(0)}$, respectively. The fact that, overall, this "recipe" does not lead to a much better agreement between quasi-PDFs and light-cone PDFs (at large x) can be traced back to other higher twist contributions in the cut-graph approach that also diverge for $x \to 1$.

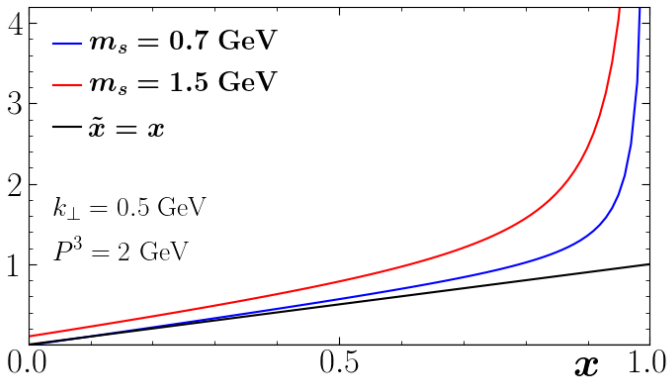

Figure 6.7: Momentum fraction $\tilde{x}$ as a function of x. $\tilde{x}$ is given by Eq. (6.65) in the cut-graph approach and is shown for different values of m_s and at $P^3 = 2$ GeV and $k_\perp = 0.5$ GeV.

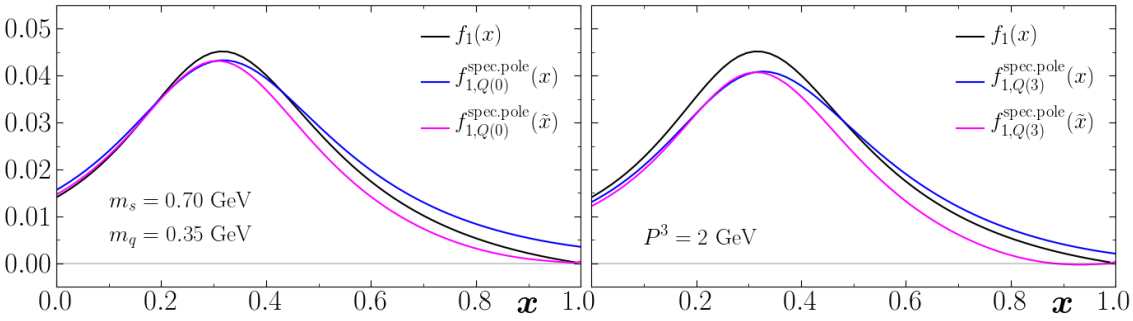

Figure 6.8: Quasi-PDF $f_{1,Q}$ in cut-diagram approach with x and $\tilde{x}$. Comparison between the quasi-PDF in the cut-diagram approach $f_{1,Q}^{\mathrm{spec.pole}}$ evaluated at x and at $\tilde{x}$ as defined in Eq. (6.65) at $P^3 = 2$ GeV and with $m_s = 0.70$ GeV and $m_q = 0.35$ GeV. The quasi-PDF $f_{1,Q(0)}$ is shown on the left while $f_{1,Q(3)}$ is shown on the right. The light-cone PDF f_1 is shown for comparison. Note that the curves for $\tilde{x}$ go to 0 as $x \to 1$, as the light-cone distributions do.

6.3.3 Results for quasi-GPDs

The results for the quasi-GPDs H_Q and E_Q are shown in Fig. 6.9 and Fig. 6.10, respectively. As mentioned above, for $P^3 \gtrsim 2$ GeV it does not matter very much whether one uses the definition involving γ^0 or γ^3. One finds that $E_{Q(3)}$ better matches the light-cone GPD E in the large-x region, while $E_{Q(0)}$ matches more closely for moderate x. Like for the quasi-PDFs, at very large x the quasi-GPDs do not converge well to the respective light-cone GPDs. One could have anticipated this outcome for H_Q (due to its relation to $f_{1,Q}$) but not necessarily for E_Q. To better visualize, especially for large x, how the quasi-GPDs and light-cone GPDs compare we show their relative difference at $\xi = 0.1$ in Fig. 6.11. At large x, the results for the relative difference are overall very similar to the PDF case. Moreover, the convergence behavior of H_Q and E_Q are mostly similar. (We do not read too much into the outcome that $E_{Q(0)}$ behaves poorer in the large x region than the other quasi-GPDs.)

The remaining six quasi-GPDs are shown in Figs. 6.12 – 6.17 at $\xi = 0.1$. For the skewness variable we have explored the range $0.01 \leq \xi \leq 0.4$ and below briefly comment on the ξ-dependence. Like in the case of quasi-PDFs, for $P^3 \gtrsim 2\,\text{GeV}$ there is no clear indication as to which of the two definitions (for each quasi-GPD) one should prefer. The convergence problem at large x persists for these quasi-GPDs as well. We emphasize that this outcome is a robust feature of our model calculation. In lattice calculations, the matching procedure could potentially improve the situation at large x, as was observed for the quasi-PDFs [404, 405]. Whether this is a robust feature of the lattice results after matching remains to be seen. We also note that, in general, there is a tendency of the discrepancies at large x to increase when ξ gets larger. The significance of this feature depends on the GPD under consideration, and it is most pronounced for the quasi-GPDs $\widetilde{E}_{\text{Q}}$ and $\widetilde{E}_{T,\text{Q}}$. This aspect is illustrated via Fig. 6.18 which clearly shows an increase in the relative difference at large x for larger values of ξ for the GPD $\widetilde{E}$ compared to the GPD H.

The plots in the Figs. 6.19 – 6.26 show the quasi-GPDs in the ERBL region for $\xi = 0.01$ and $\xi = 0.4$. Generally, for small ξ one finds significant deviations between the quasi-GPDs and the corresponding light-cone GPDs. This situation is the GPD counterpart of the problem for quasi-PDFs around $x = 0$. For small ξ, the light-cone GPDs rapidly approach zero at $x = -\xi$ in a very narrow x-range, whereas the quasi-GPDs are much smoother in that range. Once ξ is increased, we observe a (much) better agreement between quasi-GPDs and the light-cone GPDs for a large fraction

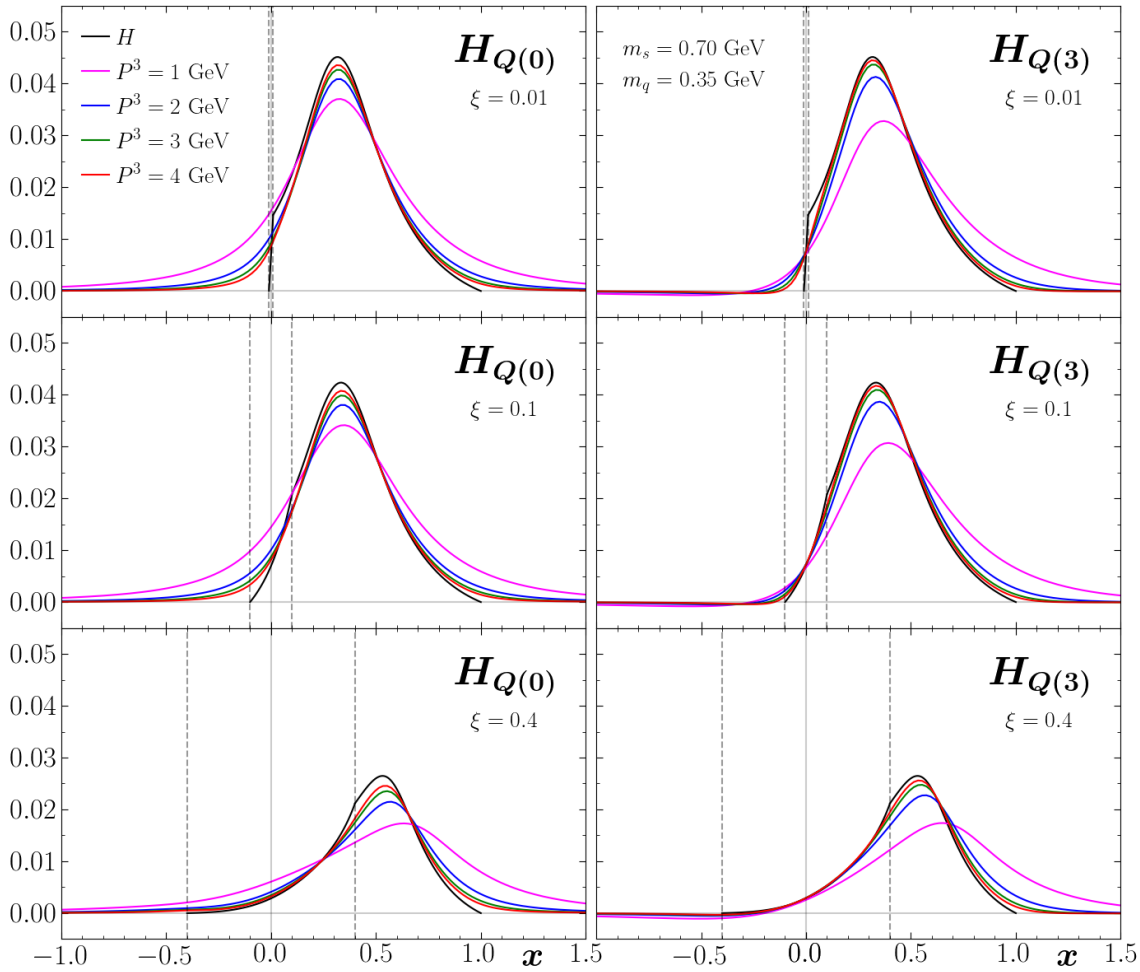

Figure 6.9: Quasi-GPD H_Q. It is shown as a function of x for different values of P^3 and ξ with $m_s = 0.70$ GeV and $m_q = 0.35$ GeV. The quasi-GPD $H_{Q(0)}$ is plotted on the left, while $H_{Q(3)}$ is plotted on the right. The light-cone GPD H is shown for comparison and the limits of the ERBL region ($\pm\xi$) are indicated by vertical dashed lines.

of the ERBL region. To be more quantitative, we look at $\widetilde{H}_{\mathrm{Q}(3)}$ as an example, with $\xi = 0.01$ at the point $x = 0.01$. From Fig. 6.21 we can see that the agreement is extremely poor for $P^3 = 4\,\mathrm{GeV}$. For decent agreement (relative difference less than 20 %), one must go to P^3 values as high as 18 GeV, which is well beyond the present reach of lattice QCD. On the other hand, if we instead choose $\xi = 0.4$ and focus on the point $x = 0.4$, one finds decent agreement (as defined above) with P^3 values as low as 1 GeV which are currently accessible in lattice QCD. This outcome suggests that lattice calculations could provide very valuable information in the ERBL region,

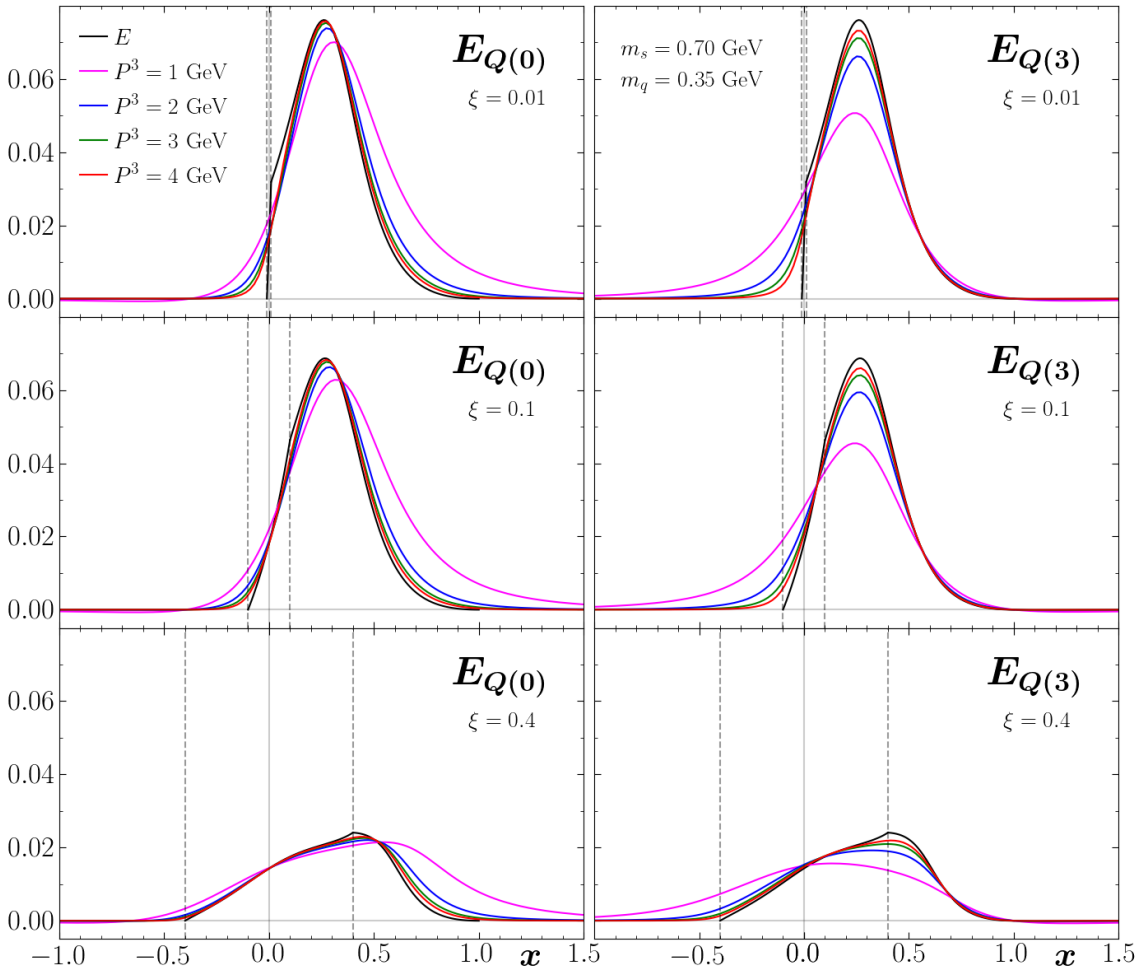

Figure 6.10: Quasi-GPD E_Q. Same as Fig. 6.9, except for the quasi-GPD E_Q.

provided that the skewness is not too small.

We have also studied the dependence of our results on the transverse momentum transfer to the hadron $|\vec{\Delta}_\perp|$ or t, where Fig. 6.27 shows results for $H_{\mathrm{Q(0)}}$ and $E_{\mathrm{Q(0)}}$. Apparently, at least at large x, the discrepancies get somewhat larger as $|\vec{\Delta}_\perp|$ is increased. However, we also find that the relative difference as defined in Eq. (6.64) is hardly affected at all when $|\vec{\Delta}_\perp|$ gets larger. This statement holds true for all the other quasi-GPDs. In fact none of the general conclusions discussed above are affected if $|\vec{\Delta}_\perp|$ is varied, where we have explored the range $0\,\mathrm{GeV} \leq |\vec{\Delta}_\perp| \leq 2\,\mathrm{GeV}$ or $0\,\mathrm{GeV}^2 \leq |t| \leq 4\,\mathrm{GeV}^2$.

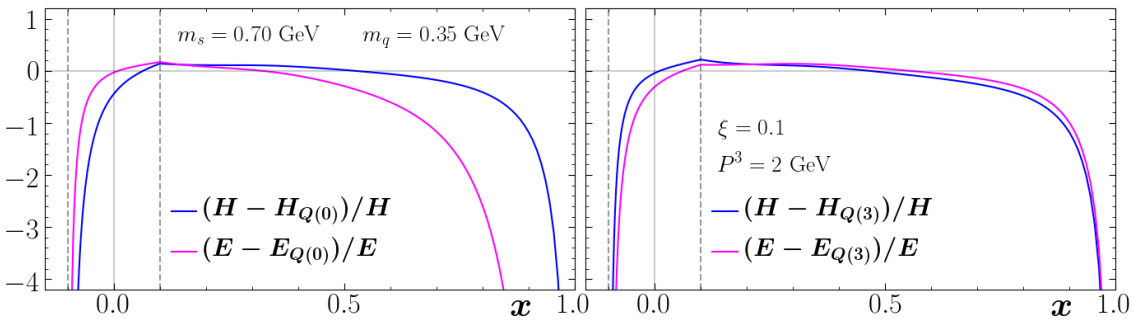

Figure 6.11: Quasi-GPDs H_Q and E_Q relative to light-cone GPDs H and E. Relative difference for the quasi-GPDs H_Q and E_Q to the light-cone GPDs H and E as a function of x at $P^3 = 2$ GeV and $\xi = 0.1$ with $m_s = 0.70$ GeV and $m_q = 0.35$ GeV. The differences for $H_{Q(0)}$ and $E_{Q(0)}$ are shown on the left, while the differences for $H_{Q(3)}$ and $E_{Q(3)}$ are shown on the right.

6.3.4 Exploring different skewness variables

So far we have used the same skewness variable ξ for both the light-cone GPDs and the quasi-GPDs. However, in the case of quasi-GPDs one could in principle consider different variables to describe the longitudinal momentum transfer to the hadron. Actually, in the matching calculations for quasi-GPDs the quantity $\tilde{\xi}_3 \equiv -\frac{\Delta^3}{2P^3}$ was used [500–502]. The two variables are related via $\tilde{\xi}_3 = \delta\xi$, with δ from Eq. (6.13). We emphasize that this relation is model-independent, which is in contrast to the situation for the parton momentum fractions $\frac{k^+}{P^+}$ and $\frac{k^3}{P^3}$ for which no model-independent relation exists. Another possible skewness variable is $\tilde{\xi}_0 \equiv -\frac{\Delta^0}{2P^0} = \frac{\xi}{\delta}$, though admittedly $\tilde{\xi}_0$ is somewhat less natural than $\tilde{\xi}_3$ due to the dependence of quasi-GPDs on $\frac{k^3}{P^3}$. In any case, the difference between the three variables is a higher twist effect that vanishes for $P^3 \to \infty$. For finite P^3, however, the differences can be substantial as illustrated in Fig. 6.28, and they are largest for large ξ. Note that as $\xi \to 1$ one has $|t| \to \infty$, and therefore also $\delta \to \infty$. Here we want to explore the impact of the difference between ξ, $\tilde{\xi}_3$, and $\tilde{\xi}_0$ on the quasi-GPDs.

In order to calculate quasi-GPDs using $\tilde{\xi}_{0/3}$ one can no longer use Eq. (6.4), but rather needs

$$\begin{aligned} t(\tilde{\xi}_0) &= -\frac{2}{\tilde{\xi}_0^2}\Big[(1-\tilde{\xi}_0^2)(P^3)^2 - 2\tilde{\xi}_0^2 M^2 \\ &- \sqrt{(1-\tilde{\xi}_0^2)^2(P^3)^4 - \tilde{\xi}_0^2(4M^2 + \vec{\Delta}_\perp^2)(P^3)^2}\Big], \end{aligned} \tag{6.66}$$

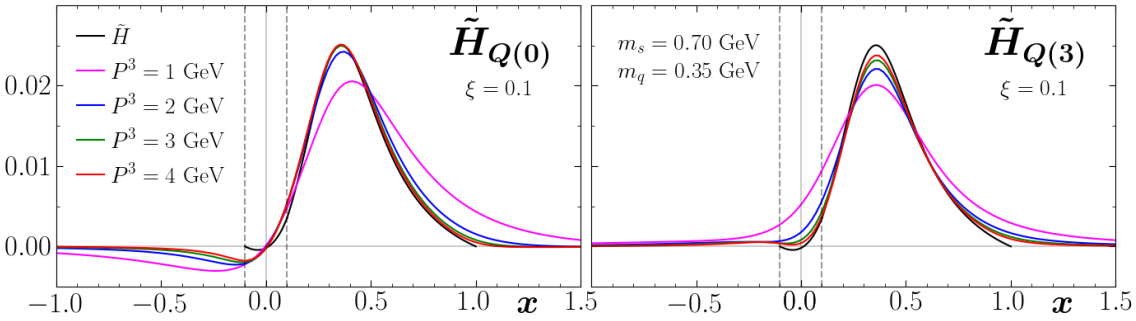

Figure 6.12: Quasi-GPD $\tilde{H}_Q$. It is shown as a function of x for different values of P^3 with $\xi = 0.1$, $m_s = 0.70$ GeV and $m_q = 0.35$ GeV. The quasi-GPD $\tilde{H}_{Q(0)}$ is plotted on the left, while $\tilde{H}_{Q(3)}$ is plotted on the right. The light-cone GPD $\tilde{H}$ is shown for comparison and the limits of the ERBL region ($\pm\xi$) are indicated by vertical dashed lines.

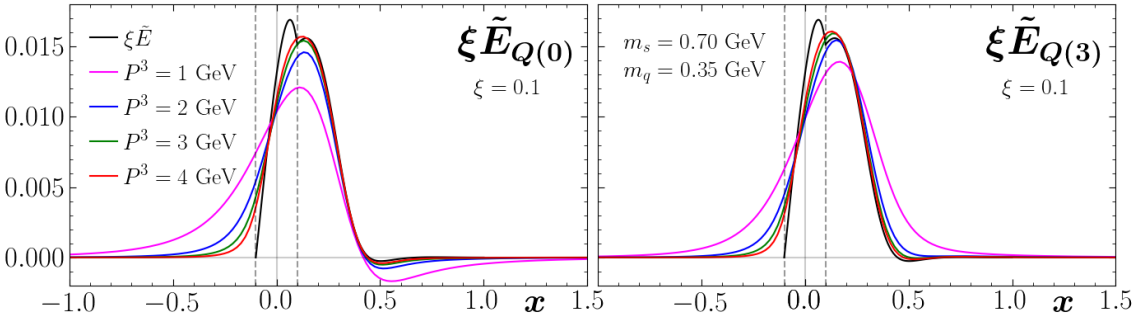

Figure 6.13: Quasi-GPD $\xi\tilde{E}_Q$. Same as Fig. 6.12, except for the quasi-GPD $\xi\tilde{E}_Q$.

$$t(\tilde{\xi}_3) \;=\; 2\left[(1-\tilde{\xi}_3^2)(P^3)^2 + M^2 - \frac{\vec{\Delta}_\perp^2}{4}\right] \tag{6.67}$$
$$- 2\sqrt{(1-\tilde{\xi}_3^2)^2(P^3)^4 + 2(1+\tilde{\xi}_3^2)\left(M^2 + \frac{\vec{\Delta}_\perp^2}{4}\right)(P^3)^2 + \left(M^2 + \frac{\vec{\Delta}_\perp^2}{4}\right)^2},$$

to compute the Mandelstam variable t. For $P^3 \to \infty$, both Eq. (6.66) and Eq. (6.67) reduce to Eq. (6.4), while non-negligible numerical differences exist when P^3 is finite. From Fig. 6.28 one finds that the allowed range for $\tilde{\xi}_0$ is smaller than $[0, 1]$. As a consequence, $t(\tilde{\xi}_0)$ would become imaginary if in Eq. (6.66) one uses a value for $\tilde{\xi}_0$ that is too large.

In Fig. 6.29 we examine the impact of choosing the different skewness variables as follows. The light-cone GPDs, which enter the relative difference R in Eq. (6.64),

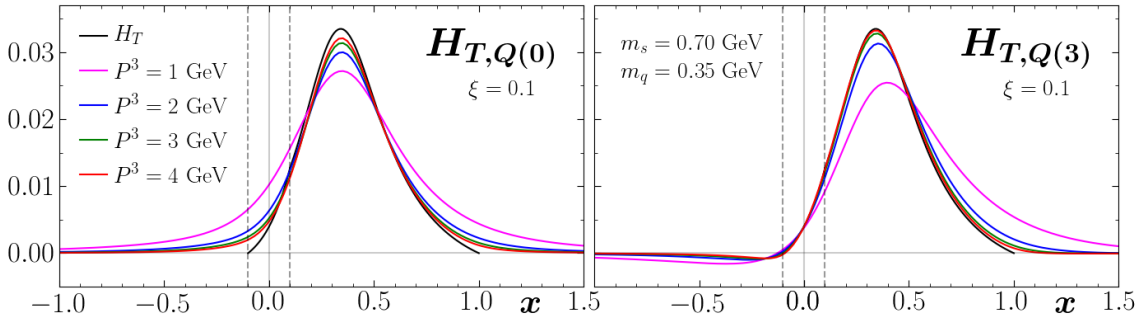

Figure 6.14: Quasi-GPD $H_{T,Q}$. Same as Fig. 6.12, except for the quasi-GPD $H_{T,Q}$.

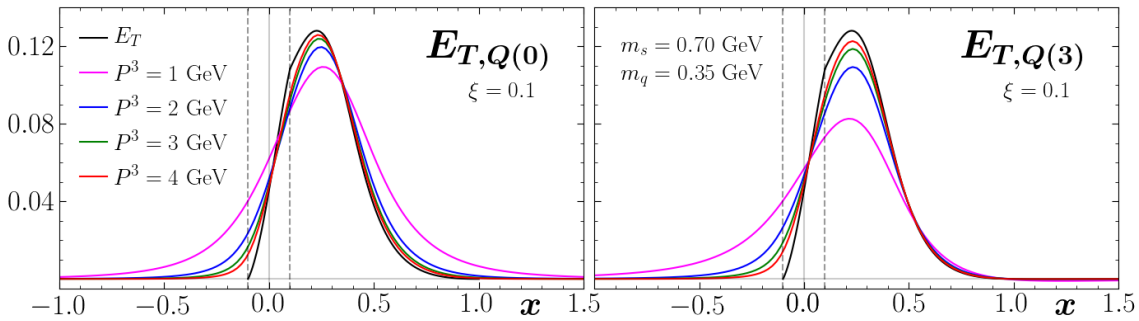

Figure 6.15: Quasi-GPD $E_{T,Q}$. Same as Fig. 6.12, except for the quasi-GPD $E_{T,Q}$.

are all evaluated for $\xi = 0.4$, while the quasi-GPDs are calculated using the three different skewness variables ξ, $\tilde{\xi}_0$ and $\tilde{\xi}_3$ all at the value 0.4. One observes considerable differences between the three cases, especially once P^3 is relatively low. Interestingly, in the case of $H_{\mathrm{Q}(0)}$ the relative difference is smaller for most of the DGLAP region (in particular, in the range where the GPDs have their maximum) if one uses $\tilde{\xi}_3$ instead of ξ. We find this pattern for most of the quasi-GPDs, while in the ERBL region no general pattern exists. The only outliers in that regard are $E_{\mathrm{Q}(0)}$, $\widetilde{E}_{\mathrm{Q}(0/3)}$ and $E_{T,\mathrm{Q}(0)}$, where $E_{\mathrm{Q}(0)}$ is shown in Fig. 6.29 as a representative case. We also observe that using the variable $\tilde{\xi}_0$ typically gives poorer convergence for the quasi-GPDs. This feature is again most pronounced in the range where the GPDs are largest. Our conclusions also hold for even larger values of ξ, where the numerical discrepancy between the three skewness variables increases further — see Fig. 6.28.

We take a moment to briefly discuss the nature of two distinct higher twist effects that we encounter in our model study. The higher twist effect associated with the longitudinal momentum transfer to the target (relating $\tilde{\xi}_{0/3}$ to ξ) is kinematical. Such an effect, expressed through the parameter δ, is model-independent and simply

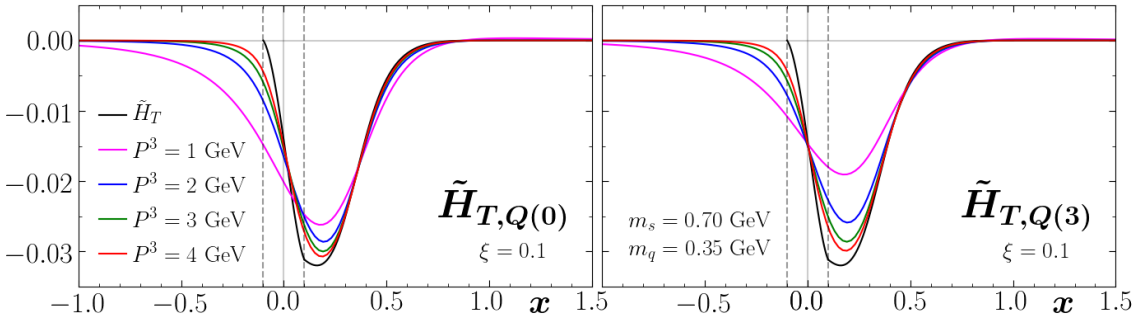

Figure 6.16: Quasi-GPD $\tilde{H}_{T,Q}$. Same as Fig. 6.12, except for the quasi-GPD $\tilde{H}_{T,Q}$.

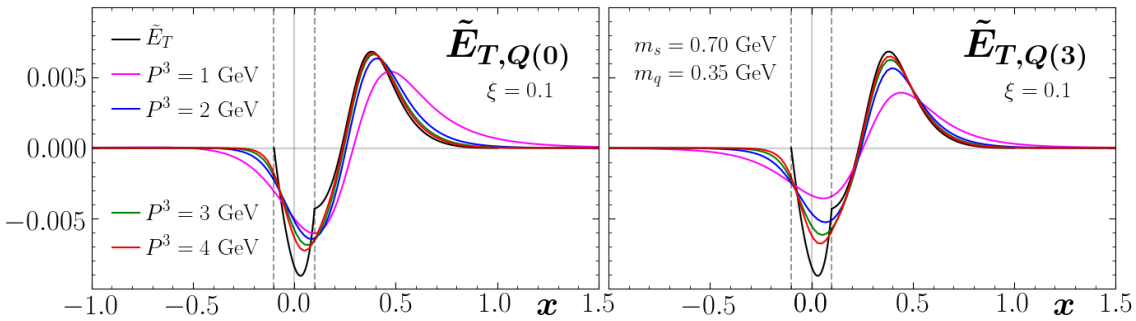

Figure 6.17: Quasi-GPD $\tilde{E}_{T,Q}$. Same as Fig. 6.12, except for the quasi-GPD $\tilde{E}_{T,Q}$.

describes the relationship between the variables P^0 and P^3. The impact of this effect, however, on the individual GPDs is model-dependent. On the other hand, the higher twist effect associated with the longitudinal parton momenta (relating $\tilde{x}$ to x) is a dynamical one which stems from modeling the spectator as an onshell diquark. We note that these effects are entirely separate from the higher twist effects associated with QCD higher twist operators.

6.4 Axial-vector Diquark Results

It is well known that both a scalar diquark and an axial-vector diquark are needed to describe the phenomenology of up quarks and down quarks in the nucleon — see, for instance, Refs. [526–528]. In this section, we therefore explore contributions from the axial-vector diquark. We repeat that we do not aim at a fine-tuned quantitative description of the standard GPDs, but rather just focus on how well the quasi-GPDs compare with their corresponding light-cone GPDs.

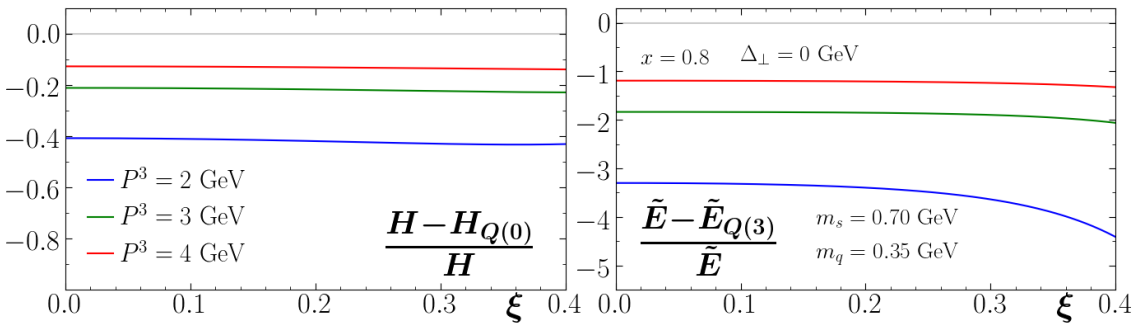

Figure 6.18: Quasi-GPDs as a function of ξ at $x = 0.8$. Relative difference between quasi-GPDs and light-cone GPDs as a function of ξ for different values of P^3 at $x = 0.8$ and with $\Delta_\perp = 0$, $m_s = 0.7$ GeV, and $m_q = 0.35$ GeV. The difference for $H_{Q(0)}$ is shown on the left, while $\tilde{E}_{Q(3)}$ is shown on the right. The relative difference at $P^3 = 1$ GeV is much larger than for $P^3 \geq 2$ GeV and is not shown.

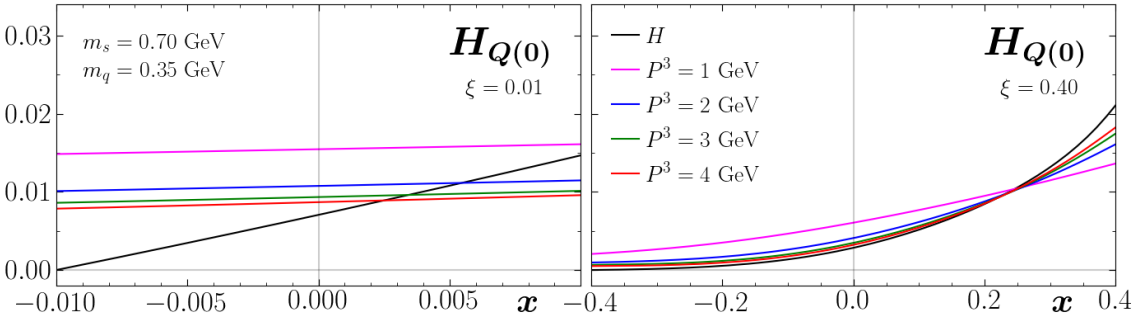

Figure 6.19: Quasi-GPD H_Q in ERBL region. It is shown as a function of x for different values of P^3 and with $m_s = 0.70$ GeV and $m_q = 0.35$ GeV. The results on the left are at $\xi = 0.01$, while the results on the right are at $\xi = 0.4$. The light-cone GPD H is shown for comparison.

In order to study the impact of the axial-vector diquark, we examine in detail the effects on the GPD H. For the scalar diquark, the vertex factor is given by ig_s , where g_s is the scalar coupling constant, and the propagator is given by $\frac{i}{(P-k)^2 - m_s^2 + i\varepsilon}$. In contrast, for the axial-vector diquark the vertex factor is given by $\frac{ig_a}{\sqrt{2}}\gamma^\mu\gamma_5$, where g_a is the axial vector coupling constant, and the propagator by $\frac{id^{\mu\nu}}{(P-k)^2 - m_a^2 + i\varepsilon}$, where $d^{\mu\nu}$ and m_a are, respectively, the polarization tensor and mass of the axial-vector diquark. There are several possible choices for the polarization tensor (see Ref. [521]), but we choose the definition

$$d^{\mu\nu} = -g^{\mu\nu} + \frac{P^\mu P^\nu}{m_a^2}\,, \tag{6.68}$$

which was analyzed in Ref. [518]. The other choices for the polarization tensor are

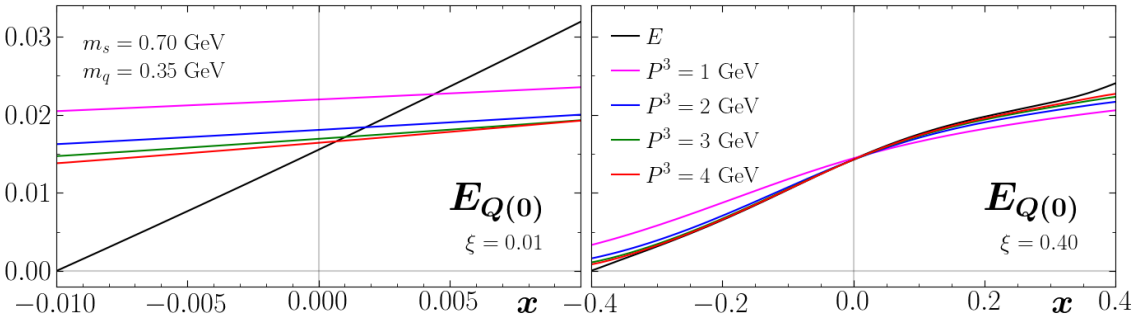

Figure 6.20: Quasi-GPD E_Q in ERBL region. Same as Fig. 6.19, except for the quasi-GPD $E_{Q(0)}$.

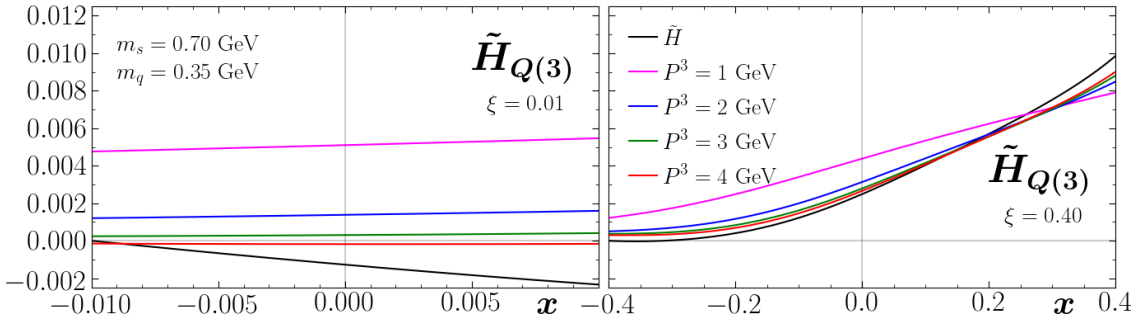

Figure 6.21: Quasi-GPD $\tilde{H}_Q$ in ERBL region. Same as Fig. 6.19, except for the quasi-GPD $\tilde{H}_{Q(3)}$.

explored in Refs. [520, 521, 529, 530]. We now replace the scalar diquark vertex and propagator in the light-cone correlation function with the axial vector diquark vertex and propagator. Note that with this replacement the light-cone correlation function no longer follows from a Lagrange density, and thus this direct replacement is also a part of our model for the axial vector diquark. The result is

$$\begin{aligned} F^{a[\Gamma]}(x,\Delta) &= \frac{ig_a^2}{4(2\pi)^4}\int dk^- d^2\vec{k}_\perp d_{\mu\nu} \\ &\times \frac{\bar{u}(p',\lambda')\gamma^\mu\gamma_5(\not{k}+\frac{\not{\Delta}}{2}+m_q)\Gamma(\not{k}+\frac{\not{\Delta}}{2}-m_q)\gamma^\nu\gamma_5 u(p,\lambda)}{D^a_{\rm GPD}}, \end{aligned} \tag{6.69}$$

where $D^a_{\rm GPD}$ is the same as $D_{\rm GPD}$ in Eq. (6.16) with the replacement $m_s \to m_a$. For the figures below we always choose $g_a = g_s = 1$ and $\vec{\Delta}_\perp = 0$ GeV. We also use $m_a = 1$ GeV as our standard value, as when quarks couple to a higher spin-state, the resulting state tends to have a larger mass [521].

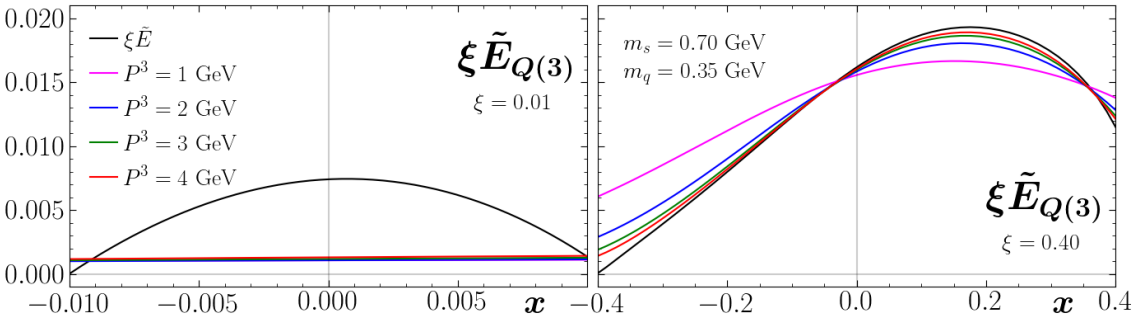

Figure 6.22: Quasi-GPD $\xi\tilde{E}_Q$ in ERBL region. Same as Fig. 6.19, except for the quasi-GPD $\xi\tilde{E}_{Q(3)}$.

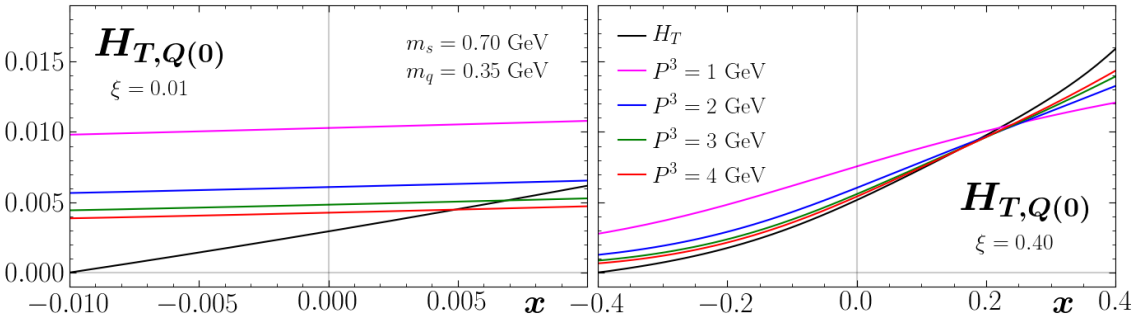

Figure 6.23: Quasi-GPD $H_{T,Q}$ in ERBL region. Same as Fig. 6.19, except for the quasi-GPD $H_{T,Q(0)}$.

For the light-cone GPD H, one again uses $\Gamma = \gamma^+$. The result for the axial-vector diquark is given by Eq. (6.17) with the replacement $N_H \to N_H^a$, where

$$\begin{aligned} 2N_H^a &= \left(2 + \frac{M^2}{m_a^2} + \frac{t}{4m_a^2}\right)\left(\vec{k}_\perp^{\,2} + m_q^2 + x^2M^2\right) + 2\left[4x - x\frac{M^2}{m_a^2} + (2+x)\frac{t}{4m_a^2}\right]m_qM \\ &+ (1+x)\left[2(1+x) + (1-3x)\frac{M^2}{m_a^2} + (1+x)\frac{t}{4m_a^2}\right]\frac{t}{4} \\ &+ (1+x)\left[2 - \frac{M^2}{m_a^2} + \frac{t}{4m_a^2}\right]\xi t\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^{\,2}}\,. \end{aligned} \tag{6.70}$$

The quasi–GPD correlator for the axial-vector diquark is given by

$$F^{a[\Gamma]}(x,\Delta;P^3) = \frac{ig_a^2}{4(2\pi)^4}\int dk^0 d^2\vec{k}_\perp d_{\mu\nu} \tag{6.71}$$

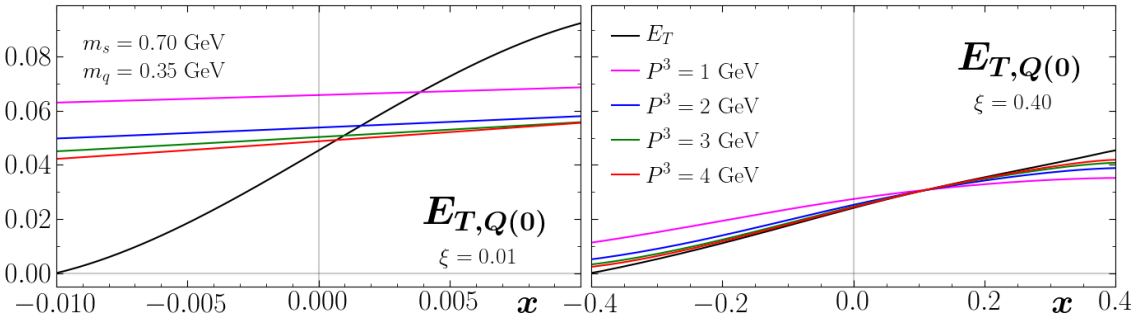

Figure 6.24: Quasi-GPD $E_{T,Q}$ in ERBL region. Same as Fig. 6.19, except for the quasi-GPD $E_{T,Q(0)}$.

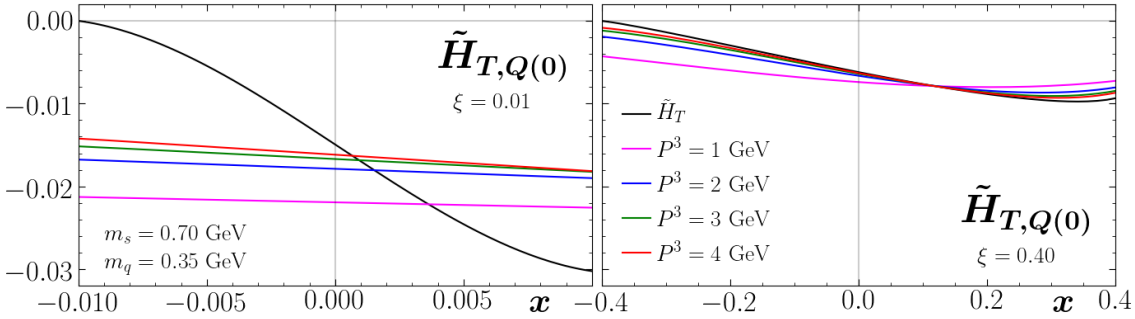

Figure 6.25: Quasi-GPD $\tilde{H}_{T,Q}$ in ERBL region. Same as Fig. 6.19, except for the quasi-GPD $\tilde{H}_{T,Q(0)}$.

$$\times \quad \frac{\bar{u}(p',\lambda')\gamma^\mu\gamma_5(\not{k}+\frac{\not{\Delta}}{2}+m_q)\Gamma(\not{k}+\frac{\not{\Delta}}{2}-m_q)\gamma^\nu\gamma_5 u(p,\lambda)}{D^a_{\rm GPD}},$$

and the results for the axial-vector diquark are given by Eq. (6.28) with the replacement $N_{H(0/3)} \to N^a_{H(0/3)}$, where

$$\begin{aligned}
\frac{2}{\delta}N^a_{H(0)} &= \left(2+\frac{M^2}{m_a^2}+\frac{t}{4m_a^2}\right)(k^0)^2 \\
&+ \frac{2}{\delta P^3}\Bigg[-x\left(2+\frac{M^2}{m_a^2}+\frac{t}{4m_a^2}\right)(P^3)^2+\left(4-\frac{M^2}{m_a^2}+\frac{t}{4m_a^2}\right)m_qM \\
&+ \left(2-\frac{M^2}{m_a^2}+\frac{t}{4m_a^2}\right)\left(x\frac{t}{4}+\frac{\delta}{2}\xi t\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\right)\Bigg]k^0 \\
&+ \left(2+\frac{M^2}{m_a^2}+\frac{t}{4m_a^2}\right)\left(\vec{k}_\perp^2+\frac{t}{4}+m_q^2+x^2(P^3)^2\right)+\frac{m_qM}{m_a^2}t
\end{aligned}$$

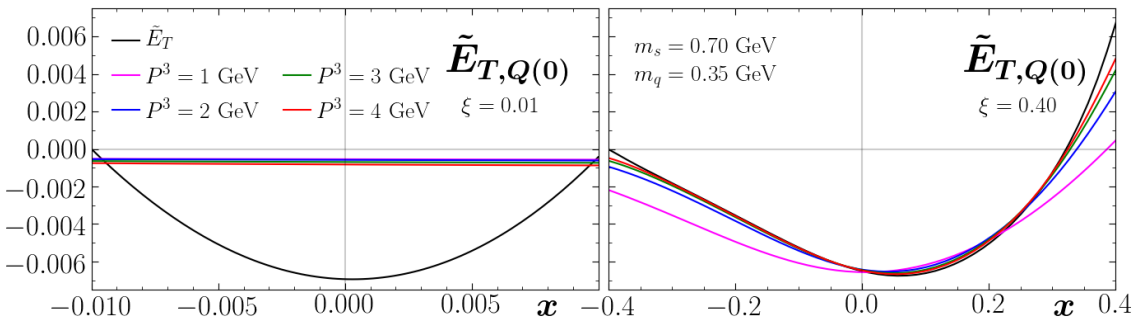

Figure 6.26: Quasi-GPD $\tilde{E}_{T,Q}$ in ERBL region. Same as Fig. 6.19, except for the quasi-GPD $\tilde{E}_{T,Q(0)}$.

$$+ \quad 2\left(2 - \frac{M^2}{m_a^2} + \frac{t}{4m_a^2}\right)\left(x\frac{t}{4} + \frac{\delta}{2}\xi t\frac{\vec{k}_\perp \cdot \vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\right), \tag{6.72}$$

$$\begin{aligned} 2N^a_{H(3)} \;=\; & -\left(2 + \frac{M^2}{m_a^2} + \frac{t}{4m_a^2}\right)(k^0)^2 \\ & + 2\delta P^3\left[\left(2 + \frac{M^2}{m_a^2} + \frac{t}{4m_a^2}\right)x + \left(2 - \frac{M^2}{m_a^2} + \frac{t}{4m_a^2}\right)\frac{1+x}{\delta^2(P^3)^2}\frac{t}{4}\right]k^0 \\ & + \left(2 + \frac{M^2}{m_a^2} + \frac{t}{4m_a^2}\right)\left(\vec{k}_\perp^2 + \frac{t}{4} + m_q^2 - x^2(P^3)^2\right) \\ & + 2\left[4x - x\frac{M^2}{m_a^2} + (2+x)\frac{t}{4m_a^2}\right]m_q M \\ & + \left(2 - \frac{M^2}{m_a^2} + \frac{t}{4m_a^2}\right)\frac{1+x}{\delta}\xi t\frac{\vec{k}_\perp \cdot \vec{\Delta}_\perp}{\vec{\Delta}_\perp^2}\,. \end{aligned} \tag{6.73}$$

In Figs. 6.30 and 6.31 we show the results obtained for the quasi-GPD $H_{\rm Q}$. Comparison with Figs. 6.9 and 6.19 above shows that the qualitative features of the GPD H remain the same regardless of the type of diquark. Fig. 6.31 shows that while convergence in the ERBL region is poor at extremely small values of ξ, it is reasonable at larger values.

Based on this, we conclude that with the polarization tensor chosen as in Eq. (6.68), there are non-negligible contributions to the GPDs and PDFs from the axial vector diquark. However, these contributions have the exact same qualitative features as those of the scalar diquark contributions, and thus our conclusions based on the scalar diquark contributions alone are robust. Our general findings therefore also apply for faithful GPDs for up and down quarks in a spectator model.

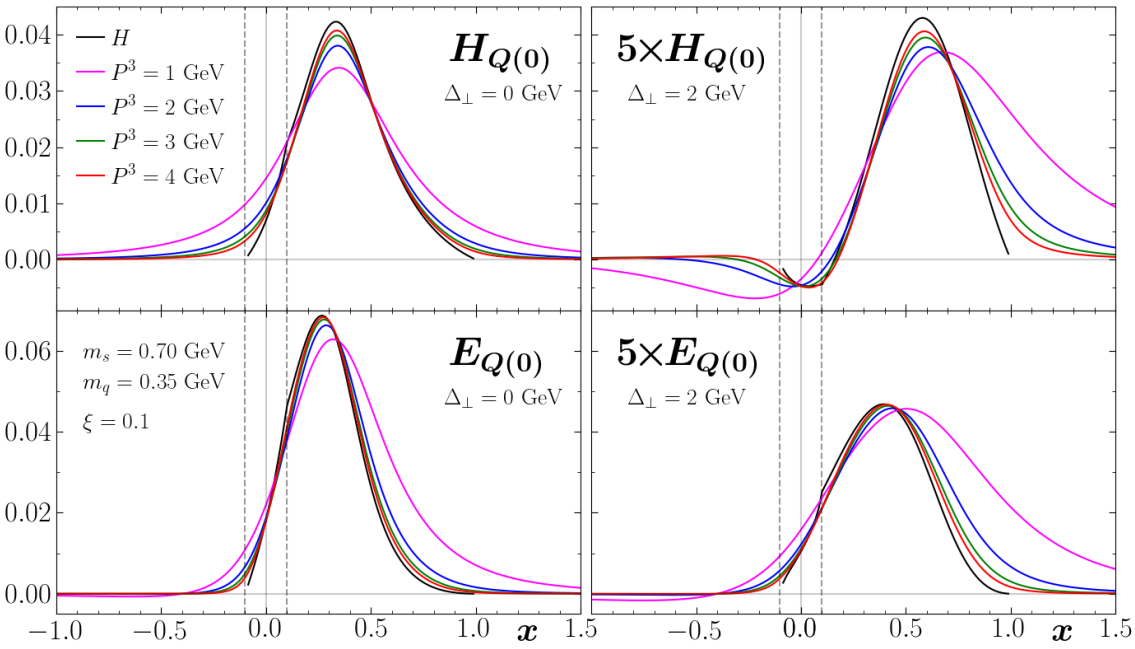

Figure 6.27: Quasi-GPDs H_Q and E_Q dependence on $\Delta_\perp$. They are shown as a function of x for different values of P^3 with $\xi = 0.1$, $m_s = 0.70$ GeV and $m_q = 0.35$ GeV. The quasi-GPD $H_{Q(0)}$ is plotted on the top, with $\Delta_\perp = |\vec{\Delta}_\perp| = 0$ GeV on the left and $\Delta_\perp = 2$ GeV on the right. Similarly, $E_{Q(0)}$ is plotted on the bottom. Note that in the right column the GPDs are multiplied by a factor of 5. The light-cone GPDs H and E are shown for comparison and the limits of the ERBL region ($\pm\xi$) are indicated by vertical dashed lines.

6.5 Moments of Quasi Distributions

Recently, moments of quasi-PDFs have attracted some attention [375, 386, 387, 389]. Specifically, in Refs [375, 386] concerns have been raised over divergences of moments of quasi-PDFs, while Ref. [387] argues that the two lowest moments are well defined. While the whole point of exploring quasi-PDFs is to go beyond the calculation of moments, it can still be instructive to look at them.

We first consider the lowest moments of quasi-GPDs and recall also the well-known results for the lowest moments of the corresponding light-cone GPDs. Including a flavor index 'q' one finds the model-independent relations

$$\int_{-1}^{1} dx H^q(x,t) = \int_{-\infty}^{\infty} dx \frac{1}{\delta} H^q_{\mathrm{Q}(0)}(x,t;P^3) = \int_{-\infty}^{\infty} dx\, H^q_{\mathrm{Q}(3)}(x,t;P^3) = F_1^q(t), \tag{6.74}$$

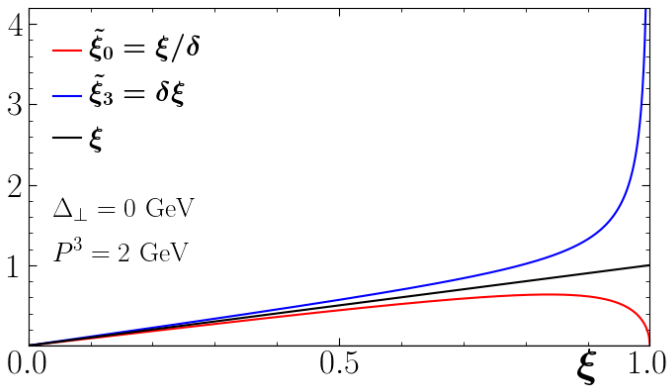

Figure 6.28: Comparison of the skewness variables ξ, $\tilde{\xi}_0$, and $\tilde{\xi}_3$. They are shown for $P^3 = 2$ GeV and $\Delta_\perp = |\vec{\Delta}_\perp| = 0$ GeV.

$$\int_{-1}^{1} dx E^q(x,t) = \int_{-\infty}^{\infty} dx \frac{1}{\delta} E^q_{\mathrm{Q}(0)}(x,t;P^3) = \int_{-\infty}^{\infty} dx E^q_{\mathrm{Q}(3)}(x,t;P^3) = F_2^q(t), \tag{6.75}$$

$$\int_{-1}^{1} dx \widetilde{H}^q(x,t) = \int_{-\infty}^{\infty} dx \widetilde{H}^q_{\mathrm{Q}(0)}(x,t;P^3) = \int_{-\infty}^{\infty} dx \frac{1}{\delta} \widetilde{H}^q_{\mathrm{Q}(3)}(x,t;P^3) = G_A^q(t), \tag{6.76}$$

$$\int_{-1}^{1} dx \widetilde{E}^q(x,t) = \int_{-\infty}^{\infty} dx \widetilde{E}^q_{\mathrm{Q}(0)}(x,t;P^3) = \int_{-\infty}^{\infty} dx \frac{1}{\delta} \widetilde{E}^q_{\mathrm{Q}(3)}(x,t;P^3) = G_P^q(t), \tag{6.77}$$

$$\int_{-1}^{1} dx H_T^q(x,t) = \int_{-\infty}^{\infty} dx \frac{1}{\delta} H^q_{T,\mathrm{Q}(0)}(x,t;P^3) = \int_{-\infty}^{\infty} dx H^q_{T,\mathrm{Q}(3)}(x,t;P^3) = F^q_{1,T}(t), \tag{6.78}$$

$$\int_{-1}^{1} dx\, E_T^q(x,t) = \int_{-\infty}^{\infty} dx \frac{1}{\delta} E^q_{T,\mathrm{Q}(0)}(x,t;P^3) = \int_{-\infty}^{\infty} dx E^q_{T,\mathrm{Q}(3)}(x,t;P^3) = 2F^q_{2,T}(t), \tag{6.79}$$

$$\int_{-1}^{1} dx \widetilde{H}_T^q(x,t) = \int_{-\infty}^{\infty} dx \frac{1}{\delta} \widetilde{H}^q_{T,\mathrm{Q}(0)}(x,t;P^3) = \int_{-\infty}^{\infty} dx \widetilde{H}^q_{T,\mathrm{Q}(3)}(x,t;P^3) = F^q_{3,T}(t), \tag{6.80}$$

where the dependence on ξ on the l.h.s. of the equations has been suppressed. In the above equations, F_1 is the Dirac form factor, F_2 the Pauli form factor, G_A the axial form factor, G_P the pseudo-scalar form factor, and $F_{1,T}$, $F_{2,T}$ and $F_{3,T}$ are the form factors of the local tensor current [531]. Note that time-reversal invariance leads to a vanishing first moment for $\widetilde{E}_T$ [27]. The results for the moments of the forward

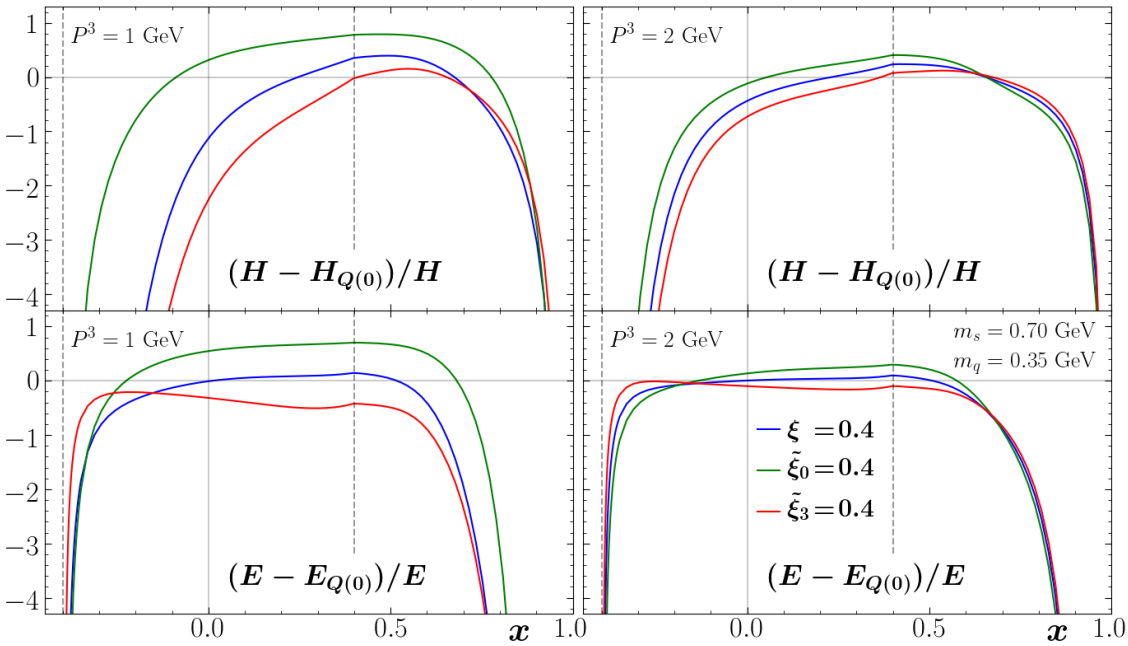

Figure 6.29: Quasi-GPDs H_Q and E_Q for different definitions of skewness. Relative difference for the quasi-GPDs relative to the light-cone GPDs as a function of x with $m_s = 0.70$ GeV and $m_q = 0.35$ GeV. The difference for $H_{Q(0)}$ is shown in the top row, with $P^3 = 1$ GeV on the left and $P^3 = 2$ GeV on the right, evaluated for three different definitions of the skewness variable (see text for more details) at 0.4. Similarly, the quasi-GPD $E_{Q(0)}$ is shown in the bottom row.

PDFs f_1, g_1 and h_1 can be extracted from Eqs. (6.74), (6.76) and (6.78), respectively. The lowest moment of light-cone GPDs depends on t, but does not depend on ξ. The quasi-GPDs depend in addition on P^3, but it is remarkable that also that dependence drops out in the lowest moment. (A corresponding discussion for $f_{1,\mathrm{Q}}$ can be found in Ref. [387].) However, according to Eqs. (6.74)–(6.80) one must divide half of the quasi-GPDs by the (simple) kinematical factor δ in order to arrive at this result. (Our numerical results in the SDM comply with Eqs. (6.74)–(6.80).) For $P^3 \lesssim 2\,\mathrm{GeV}$ inclusion of this factor leads to a visible difference for the numerics. Of course δ describes a higher twist effect, and therefore including this factor is in principle a matter of taste. But the moment analysis suggests that taking into account δ like in Eqs. (6.74)–(6.80) appears natural. In the case of quasi-PDFs, such a definition implies that $f_{1,\mathrm{Q}(0)}$, $g_{1,\mathrm{Q}(3)}$ and $h_{1,\mathrm{Q}(0)}$ are to be divided by δ_0 in comparison to what so far has been used mostly in the literature (as of the writing of Ref. [62]).

We now turn our attention to the second moment of quasi-distributions, but con-

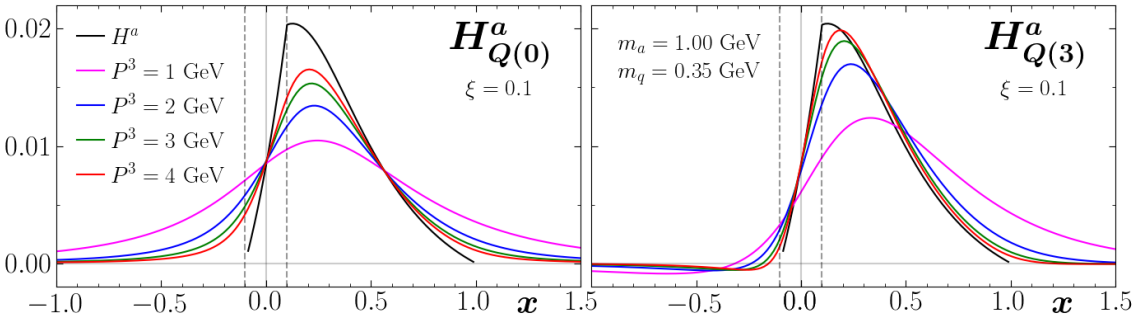

Figure 6.30: Axial-vector quasi-GPD H^a_Q. It is shown as a function of x for different values of P^3 at $\xi = 0.1$ and with $m_a = 1$ GeV and $m_q = 0.35$ GeV. The quasi-GPD $H_{Q(0)}$ is plotted on the left, while $H_{Q(3)}$ is plotted on the right. The light-cone GPD H^a is shown for comparison and the limits of the ERBL region ($\pm\xi$) are indicated by vertical dashed lines.

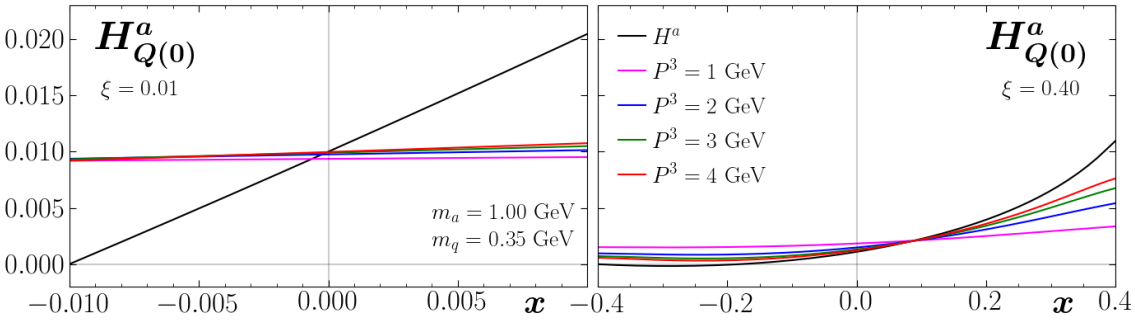

Figure 6.31: Axial-vector quasi-GPD H^a_Q in ERBL region. It is shown as a function of x for different values of P^3 and with $m_a = 1$ GeV and $m_q = 0.35$ GeV. The results on the left are at $\xi = 0.01$, while the results on the right are at $\xi = 0.4$. The light-cone GPD H^a is shown for comparison.

sider the vector operator $\bar{\psi}^q \gamma^\mu \psi^q$ only. It is well known that the corresponding local operators are related to the form factors of the energy momentum tensor (EMT) $T^{\mu\nu}$. The EMT, when evaluated between different hadron states, has five independent structures [532],

$$\begin{aligned}\langle p', \lambda' | T^{\mu\nu,q}(0) | p, \lambda \rangle &= \bar{u}(p', \lambda') \Bigg[\frac{P^\mu P^\nu}{M} A^q(t) + \frac{\Delta^\mu \Delta^\nu - g^{\mu\nu}\Delta^2}{M} C^q(t) \\ &+ M g^{\mu\nu} \bar{C}^q(t) + \frac{P^{\{\mu} i\sigma^{\nu\}\alpha} \Delta_\alpha}{4M} \big(A^q(t) + B^q(t) \big) \\ &+ \frac{P^{[\mu} i\sigma^{\nu]\alpha} \Delta_\alpha}{4M} D^q(t) \Bigg] u(p, \lambda) \,, \end{aligned} \tag{6.81}$$

where $T^{\mu\nu,q}(0) = \bar{\psi}^q(0)\gamma^\mu \frac{i}{2}\overleftrightarrow{D^\nu}\psi^q(0)$ with D^μ the covariant derivative, $a^{\{\mu}b^{\nu\}} = a^\mu b^\nu + a^\nu b^\mu$ and $a^{[\mu}b^{\nu]} = a^\mu b^\nu - a^\nu b^\mu$. The connection between the quasi-GPDs and the form factors of the EMT is established through the equation

$$(P^3)^2 \int_{-\infty}^{\infty} dx\, x \int_{-\infty}^{\infty} \frac{dz^3}{2\pi} e^{ixP^3z^3} \langle p', \lambda'|\bar{\psi}^q(-\tfrac{z^3}{2})\gamma^\mu \mathcal{W}_{\rm Q}(-\tfrac{z}{2}, \tfrac{z}{2})\, \psi^q(\tfrac{z^3}{2})|p, \lambda\rangle \Big|_{z^0=0, \vec{z}_\perp=\vec{0}_\perp}$$
$$= \langle p', \lambda'|T^{\mu 3,q}(0)|p, \lambda\rangle\,, \tag{6.82}$$

where the index μ can be 0 or 3. In close analogy to the celebrated expression for the second moment of $H + E$, namely $\int_{-1}^{1} dx\, x \left(H^q(x,\xi,t) + E^q(x,\xi,t)\right) = A^q(t) + B^q(t)$ where $A^q(0) + B^q(0) = J^q$ is the total angular momentum for the quark flavor q [462], one then finds for the quasi-GPDs

$$\int_{-\infty}^{\infty} dx\, x\, \frac{1}{\delta}\left(H^q_{\rm Q(0)}(x,\xi,t;P^3) + E^q_{\rm Q(0)}(x,\xi,t;P^3)\right) = \frac{1}{2}(\delta^2+1)\big(A^q(t) + B^q(t)\big)$$
$$+ \frac{1}{2}(\delta^2 - 1)D^q(t)\,, \tag{6.83}$$

$$\int_{-\infty}^{\infty} dx\, x \left(H^q_{\rm Q(3)}(x,\xi,t;P^3) + E^q_{\rm Q(3)}(x,\xi,t;P^3)\right) = A^q(t) + B^q(t)\,. \tag{6.84}$$

Note that in Eq. (6.83) the form factor D^q of the anti-symmetric part of the EMT enters. One can conclude that the second moment of $H_{\rm Q(3)} + E_{\rm Q(3)}$ is directly related to the angular momentum of quarks, while for $H_{\rm Q(0)} + E_{\rm Q(0)}$ this relation contains a higher twist "contamination." Our numerics are in accord with these two equations in the sense that the l.h.s. of Eq. (6.84) is independent of P^3 and agrees with what we find from the second moment of $H + E$, while the l.h.s. of Eq. (6.83) does depend on P^3 and converges to $A^q(t) + B^q(t)$ for large P^3. In Table 6.1, we show that the first moments of $H_{\rm Q(3)}$ and $E_{\rm Q(3)}$ match with the first moments of H and E, respectively, and are independent of P^3. We also show that Ji's spin-sum rule holds for the γ^3 projection regardless of the value of P^3. The conclusions remain the same if one uses γ^0 as the projection.

We now briefly take up the case of the second moment for f_1. In that case one

	$\int dx H_{\mathrm{Q}(3)}$	$\int dx E_{\mathrm{Q}(3)}$	$\int dx\, x\,(H_{\mathrm{Q}(3)} + E_{\mathrm{Q}(3)})$
$P^3 = 1$ GeV	0.0105746	0.0136164	0.00904580
$P^3 = 2$ GeV	0.0105743	0.0136165	0.00904583
$P^3 = 3$ GeV	0.0105744	0.0136164	0.00904580
$P^3 = 4$ GeV	0.0105743	0.0136163	0.00904584
Light-cone	0.0105741	0.0136164	0.00904572

Table 6.1: Moments of GPDs H and E. Left column: First moment of $H_{\mathrm{Q}(3)}$ for various values of P^3. Middle column: First moment of $E_{\mathrm{Q}(3)}$ for various values of P^3. Right column: Ji's spin-sum rule for the γ^3 projection. The bottom row shows the corresponding moments of the standard GPDs. All moments have been obtained with $\xi = 0.1$, $t = -1$ GeV^2, $m_s = 0.70$ GeV, and $m_q = 0.35$ GeV.

has

$$\int_{-1}^{1} dx\, x\, f_1(x) = A^q(0)\,, \tag{6.85}$$

and the corresponding equations for the quasi-PDFs read

$$\int_{-\infty}^{\infty} dx\, x\, \frac{1}{\delta_0}\, f_{1,\mathrm{Q}(0)}(x; P^3) \;=\; A^q(0)\,, \tag{6.86}$$

$$\int_{-\infty}^{\infty} dx\, x\, f_{1,\mathrm{Q}(3)}(x; P^3) \;=\; A^q(0) - (\delta_0^2 - 1)\bar{C}^q(0)\,. \tag{6.87}$$

The second moment of the quasi-PDF $f_{1,\mathrm{Q}(0)}$ is independent of P^3 only if the function is divided by δ_0, which agrees with the situation for the lowest moment. On the other hand, the second moment of $f_{1,\mathrm{Q}(3)}$ does depend on P^3. Once again, our numerical results align with these analytical results. We also find that the third moments of the quasi-PDFs $f_{1,\mathrm{Q}}$, $g_{1,\mathrm{Q}}$ and $h_{1,\mathrm{Q}}$ and their corresponding quasi-GPDs (H_{Q}, $\widetilde{H}_{\mathrm{Q}}$, and $H_{T,\mathrm{Q}}$) diverge. On the other hand, the third moments of the quasi-GPDs without forward counterparts do not diverge.

We emphasize again that these moments relations are model-independent. For the regulated results, the moments are finite for both the light-cone and quasi-distributions. Of course, renormalization of the quasi-distributions needs to be considered as well. However, this point is equally relevant for the moments of the light-cone GPDs. The model-independent expressions for the moments of the quasi-distributions

are potentially significant as they may be useful for studying the systematic uncertainties of results from lattice QCD, especially due to the fact that the P^3-dependence of the moments is either computable or nonexistent.

6.6 Summary

In this chapter we have presented the results for all of the quasi-GPDs corresponding to the eight leading-twist light-cone GPDs in the Scalar Diquark Model. For each light-cone GPD we have studied two quasi-GPDs. Taking the forward limit, we have also obtained the quasi-PDFs $f_{1,\mathrm{Q}}$, $g_{1,\mathrm{Q}}$ and $h_{1,\mathrm{Q}}$. In the limit $P^3 \to \infty$, all quasi-GPDs analytically reduce to the respective light-cone GPDs, as required. This outcome further supports the idea of computing quasi-GPDs in lattice QCD to get information on light-cone GPDs. Though the forward PDFs (in the SDM) are discontinuous at $x = 0$, for $P^3 \to \infty$ they are exactly reproduced by the corresponding quasi-PDFs. For the PDFs we have found significant numerical discrepancies between the quasi-distributions and the light-cone distributions near $x = 0$ and $x = 1$. It was found that the differences near $x = 1$ are due to higher twist effects that grow as $x \to 1$. For GPDs these disparities tend to increase with an increase of the skewness ξ. On the other hand, for large ξ we have found quite good agreement between quasi-GPDs and light-cone GPDs for a significant part of the ERBL region. Furthermore, at least in the DGLAP region, we have observed for most quasi-GPDs a better agreement with the light-cone GPDs if ξ is replaced by $\tilde{\xi}_3 = -\frac{\Delta^3}{2P^3}$. The difference between ξ and $\tilde{\xi}_3$ is a higher twist effect.

We have also identified the robustness of the results in the SDM by studying the dependence on the free parameters of the model — the quark mass m_q, the spectator mass m_s, the cut-off for the integration upon the transverse quark momentum Λ, and the momentum transfer to the hadron. We have further found that the contributions from the axial diquark are qualitatively similar to those of the scalar diquark. Moreover, we have derived model-independent results for the first and second moments of quasi-distributions, finding that these moments either do not depend on P^3, or their P^3-dependence can be computed. A particularly interesting case is the second moment of $H_{\mathrm{Q}}+E_{\mathrm{Q}}$, which is related to the total angular momentum of quarks, where the results for the moments suggest a preferred definition of several quasi-distributions.

Moments of quasi-distributions might allow one to explore systematic uncertainties of results in lattice QCD.

CHAPTER 7

SUMMARY AND OUTLOOK

A description of hadron structure is possible through the process of QCD factorization, allowing one to separate a high energy scattering process into perturbatively calculable hard coefficients and nonperturbative functions. Information on these nonperturbative functions can be gleaned from experimental measurements serving as high energy probes, and, since they are universal, a wide range of experimental measurements can be used. Global QCD analyses combine the theory of factorization with many experimental processes in order to extract these functions and thus provide information on the fundamental composition of hadrons.

Many questions still remain about the inner structure of hadrons; even when considering just the simplest 1D PDFs. The proton spin puzzle, which originated in 1988 from the EMC result [533] which found, against expectations, a small contribution to the proton's total spin from the spin of its valence quarks, to some extent still persists to this day. Furthermore, there is still little information on the proton's transverse spin structure, and questions remain about the spin-averaged strange PDF. Moreover, when it comes to the 3D parton distributions, TMDs and GPDs, many studies are still in their early stages. These include model calculations, lattice QCD calculations, and global QCD analyses. In this dissertation, we used a global QCD framework developed in previous JAM analyses [112–116] and model calculations to gain new insight into these mysteries.

In Chapter 2 we presented the results of a global QCD analysis of spin-averaged PDFs including the SeaQuest measurement of pp and pD Drell-Yan cross sections [95] and the latest STAR measurement of the W-lepton production process [96] and examined their impact on the $\bar{d} - \bar{u}$ asymmetry. We found a large impact from the

SeaQuest data and a modest impact from the STAR data, with the SeaQuest data keeping the asymmetry positive up to $x \approx 0.4$. We also found some tension between the SeaQuest and NuSea measurements [93, 94], and some difficulty in describing the STAR data at extreme pseudorapidities. The shape and magnitude of $\bar{d} - \bar{u}$ from the global analysis is consistent with expectations from nonperturbative models, such as those based on chiral symmetry breaking in QCD. Finally, we compared our extracted PDFs to those of other analyses and find some notable differences, particularly for the strange quark PDF.

In the future, SIDIS data from the EIC could also provide constraints on the spin-averaged strange and light sea quark distribution at low x [534]. With regards to the strange PDF, there are several improvements that could be made to our analysis. The first is to parameterize the nonperturbative charm, and to improve the heavy quark treatment using the ACOT renormalization scheme [535] rather than the Zero Mass Variable Flavor Number Scheme [536]. The inclusion of W/Z boson production data from ATLAS [170–176] may also provide important constraints on the strange distribution. Data on neutrino DIS from heavy nuclei from CHORUS [537] and NuTeV [538] are also available and are known to be relevant for the strange distribution, however their calculation involves knowledge of final-state nuclear effects [539,540] which are unknown experimentally and thus are a large source of systematic uncertainty [167]. Another improvement on our analysis of spin-averaged PDFs includes the resummation of large logarithms at high x for DIS and DY [541].

In Chapter 3 we presented the results of a global QCD analysis of spin-averaged PDFs and nuclear PDFs including the latest data from the MARATHON collaboration on helium and tritium targets [129], finding the first indication of an isovector effect in nuclear structure functions. We found that the MARATHON ^{3}He/^{3}H data is particularly sensitive to off-shell corrections at high $x_{\rm bj}$, making it vital for their extraction. Our extraction also shows disagreement with the KP model [194] when it comes to the EMC ratios for deuterium, helium, and tritium as well as the F_2^n/F_2^p ratio. We have also studied the impact of TMCs and higher twist contributions, and found that our results for the PDFs and off-shell functions are stable regardless of choices relating to TMCs and higher twists.

Additional information on the nuclear EMC effects in ^{3}He and ^{3}H separately will come from ^{3}He/D and ^{3}H/D ratios measured by MARATHON, which are expected to be analyzed in the near future. Constraints on the neutron structure, and the d/u

PDF ratio at large x, will come from the BONuS experiment at Jefferson Lab, which tags spectator protons in semi-inclusive DIS from the deuteron.

In Chapter 4 we performed the first simultaneous extraction of helicity PDFs, spin-averaged PDFs, and pion, kaon, and unidentified hadron FFs. This analysis included the latest polarized W-lepton production data from the STAR collaboration [96], and led to a data-driven extraction of a nonzero polarized sea asymmetry at intermediate x that was consistent with predictions from the statistical [252] and chiral soliton models [251]. The truncated contributions to the proton's spin were also extracted, finding for the first time a nonzero contribution to the proton's spin from $\Delta\bar{u}$. It was also found that there are two solutions for the gluon helicity, dependent on the imposition of positivity constraints. Finally the extracted pion, kaon, and unidentified hadron FFs were compared to those of other recent extractions.

More information on the helicity sea distribution may come soon from the JLab12 upgrade. In particular, the high-luminosity CLAS12 SIDIS experiment using K production [542] will provide precise SIDIS data to complement the W-lepton production data from RHIC. The puzzle of the two gluon solutions may be aided in the future by SIDIS data from Jefferson Lab [285], jet production data from the EIC [286], and information from lattice QCD [287–289]. Generally, the EIC [39,543] will allow global QCD analyses to access new information on the spin structure of the proton [63,74]. The EIC can measure the DIS double longitudinal DIS asymmetry at higher values of Q^2 and lower values of x compared to what is currently available. Impact studies have shown that this will provide significant constraints on the gluon's helicity at $x < 0.01$, where no information is currently available, due to the large scaling violations at low x [38,544,545]. With its ability to obtain larger statistics, the EIC can also measure new observables, such as parity-violating DIS (PVDIS) [546,547]. With the nucleon polarized and lepton unpolarized, PVDIS could provide information on the helicity strange distribution [38], and with the lepton polarized and the nucleon unpolarized it could provide information on the weak mixing angle [548]. Inclusive jet and di-jet production from DIS could also be measured at the EIC, providing a new process capable of constraining both the quark and gluon helicity PDFs [549,550]. Heavy quark production could also provide information on the gluon helicity [551]. Tagged DIS with deuterons could also provide insight on the neutron spin structure [552]. The increased precision and kinematic range at the EIC and in other future experiments require that the global QCD formalism keep pace, and advancements continue in that

arena. NNLO calculations have become available for jets [553] and W-boson production [554] that will allow global QCD analyses to extend beyond NLO calculations. A recent global analysis implemented a new formalism for helicity PDF evolution at small $x < 0.01$, which puts constraints on the helicity PDFs below the lower x limit of experimental data [555].

In Chapter 5 we have presented the results of a simultaneous global QCD analysis of DiFFs, IFFs, and transversity PDFs [60]. We have also proposed a new definition for the unintegrated DiFFs that is compatible with the probability interpretation of collinear DiFFs and derived the LO evolution equations with this new definition [59]. We include, for the first time, the BELLE measurement on the e^+e^- cross section [333], the STAR measurement at $\sqrt{s} = 500$ GeV [339], and preliminary data from STAR at $\sqrt{s} = 200$ GeV [338]. Our extracted tensor charges are in good agreement with the 2018 analysis of Radici and Bacchetta [32] which also used the DiFF channel to extract the transversity PDFs. This is a remarkable result, given that the two analyses differ in a number of aspects. Upon including lattice QCD information, we find good agreement between the lattice QCD result and experiment. The remaining tension could be attributed to approximations in our analysis, particularly the fact that it is only LO in the strong coupling, or to underestimated systematic uncertainties on the lattice QCD results. Our results indiciate that the three approaches used to extract transversity PDFs (DiFF channel, TMD channel, and lattice QCD) are in fact reconcilable.

In the future, this analysis could be improved by extending it to NLO in the strong coupling. Currently, all of the relevant NLO formulas still need to be derived, although the result for the SIA cross section differential in z_1 and z_2 is available already [312]. If SIA were extended to NLO, it would also allow the proper treatment of the thrust cut on the data. New data could allow less dependence on PYTHIA or other event generators for constraining D_1. Multiplicity data from SIDIS from COMPASS [556] on proton and deuteron targets as well as cross section data from pp collisions [557] could be used to constrain the other D_1 functions, including the gluon. The current analysis only involves $\pi^+\pi^-$ DiFFs, but could in principle be extended to include others if the data were available. In fact, Belle already has SIA cross section data available for $\pi^+\pi^+$, π^+K^+, π^+K^-, K^+K^-, and K^+K^+ production [333], while COMPASS has SIDIS asymmetry data available for π^+K^-, π^-K^+, and K^+K^- production [318]. Using these other channels could potentially provide new information

on the transversity sea distributions. However, until data are available to constrain the IFFs for these new channels, this data cannot be used to gain information on the transversity PDFs. In the future, the EIC can provide crucial measurements to help reduce uncertainties on the tensor charges [558]. Finally, a simultaneous analysis of the DiFF and TMD channels will provide a definitive test of whether the two methods of extracting transversity PDFs are compatible.

In Chapter 6 we have presented the results for all of the quasi-GPDs corresponding to the eight leading-twist light-cone GPDs in the Scalar Diquark Model. For each light-cone GPD we have studied two quasi-GPDs. Taking the forward limit, we have also obtained the quasi-PDFs $f_{1,\mathrm{Q}}$, $g_{1,\mathrm{Q}}$ and $h_{1,\mathrm{Q}}$. In the limit $P^3 \to \infty$, all quasi-GPDs (quasi-PDFs) analytically reduce to the respective light-cone GPDs (PDFs), as required. We have studied in detail the agreement between the quasi and light-cone distributions for finite P^3 and its dependence on x and ξ. We have further found that the contributions from the axial diquark are qualitatively similar to those of the scalar diquark. Moreover, we have derived model-independent results for the first and second moments of quasi-distributions, finding that these moments either do not depend on P^3, or their P^3-dependence can be computed. Moments of quasi-distributions might allow one to explore systematic uncertainties of results in lattice QCD. In the future, these model results could be extended to GPDs of sub-leading twist, where lattice QCD results are beginning to be calculated [514].

Future experimental data, particularly from the JLab 12 program and the EIC, are primed to provide valuable new input to global QCD analyses. In the meantime, global analyses are also being expanded to extract higher dimensional objects such as TMDs and GPDs. To mention just two examples, the JAM collaboration has recently extracted transversity PDFs alongside the Collins and Sivers TMDs [33, 34], and the worm-gear TMD has also been extracted [559]. Advancements in lattice QCD are also important for comparisons to and inclusion in global analyses.

BIBLIOGRAPHY

[1] E. Rutherford. The scattering of alpha and beta particles by matter and the structure of the atom. *Phil. Mag. Ser. 6*, 21:669–688, 1911.

[2] J. Chadwick. Possible Existence of a Neutron. *Nature*, 129:312, 1932.

[3] J. Chadwick. The Existence of a Neutron. *Proc. Roy. Soc. Lond. A*, 136(830):692–708, 1932.

[4] Murray Gell-Mann. A Schematic Model of Baryons and Mesons. *Phys. Lett.*, 8:214–215, 1964.

[5] G. Zweig. An SU(3) model for strong interaction symmetry and its breaking. Version 1. 1 1964.

[6] R. P. Feynman. Photon-hadron interactions. 1973.

[7] Peter W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Lett.*, 12:132–133, 1964.

[8] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, 1964.

[9] Peter W. Higgs. Spontaneous Symmetry Breakdown without Massless Bosons. *Phys. Rev.*, 145:1156–1163, 1966.

[10] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, 1964.

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

[12] Jon Butterworth. The Standard Model: How far can it go and how can we tell? *Phil. Trans. Roy. Soc. Lond. A*, 374(2075):20150260, 2016.

[13] Gerard 't Hooft. Renormalization of Massless Yang-Mills Fields. *Nucl. Phys. B*, 33:173–199, 1971.

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

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

[16] R. P. Feynman. The behavior of hadron collisions at extreme energies. *Conf. Proc. C*, 690905:237–258, 1969.

[17] 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.

[18] 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.

[19] John Collins. *Foundations of perturbative QCD*, volume 32. Cambridge University Press, 11 2013.

[20] P. J. Mulders and R. D. Tangerman. The Complete tree level result up to order 1/Q for polarized deep inelastic leptoproduction. *Nucl. Phys. B*, 461:197–237, 1996. [Erratum: Nucl.Phys.B 484, 538–540 (1997)].

[21] John C. Collins. What exactly is a parton density? *Acta Phys. Polon. B*, 34:3103, 2003.

[22] Markus Diehl. Introduction to GPDs and TMDs. *Eur. Phys. J. A*, 52(6):149, 2016.

[23] Ted C. Rogers. An overview of transverse-momentum–dependent factorization and evolution. *Eur. Phys. J. A*, 52(6):153, 2016.

[24] A. V. Radyushkin. Nonforward parton distributions. *Phys. Rev. D*, 56:5524–5557, 1997.

[25] Xiang-Dong Ji. Off forward parton distributions. *J. Phys. G*, 24:1181–1205, 1998.

[26] A. V. Belitsky and A. V. Radyushkin. Unraveling hadron structure with generalized parton distributions. *Phys. Rept.*, 418:1–387, 2005.

[27] M. Diehl. Generalized parton distributions. *Phys. Rept.*, 388:41–277, 2003.

[28] Kresimir Kumericki, Simonetta Liuti, and Herve Moutarde. GPD phenomenology and DVCS fitting: Entering the high-precision era. *Eur. Phys. J. A*, 52(6):157, 2016.

[29] Stephan Meissner, Andreas Metz, and Marc Schlegel. Generalized parton correlation functions for a spin-1/2 hadron. *JHEP*, 08:056, 2009.

[30] R. L. Jaffe and Aneesh Manohar. The G(1) Problem: Fact and Fantasy on the Spin of the Proton. *Nucl. Phys. B*, 337:509–546, 1990.

[31] Marco Radici, A. Courtoy, Alessandro Bacchetta, and Marco Guagnelli. Improved extraction of valence transversity distributions from inclusive dihadron production. *JHEP*, 05:123, 2015.

[32] Marco Radici and Alessandro Bacchetta. First Extraction of Transversity from a Global Analysis of Electron-Proton and Proton-Proton Data. *Phys. Rev. Lett.*, 120(19):192001, 2018.

[33] Justin Cammarota, Leonard Gamberg, Zhong-Bo Kang, Joshua A. Miller, Daniel Pitonyak, Alexei Prokudin, Ted C. Rogers, and Nobuo Sato. Origin of single transverse-spin asymmetries in high-energy collisions. *Phys. Rev. D*, 102(5):054002, 2020.

[34] 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.

[35] Rajan Gupta, Yong-Chull Jang, Boram Yoon, Huey-Wen Lin, Vincenzo Cirigliano, and Tanmoy Bhattacharya. Isovector Charges of the Nucleon from 2+1+1-flavor Lattice QCD. *Phys. Rev. D*, 98:034503, 2018.

[36] Nesreen Hasan, Jeremy Green, Stefan Meinel, Michael Engelhardt, Stefan Krieg, John Negele, Andrew Pochinsky, and Sergey Syritsyn. Nucleon axial, scalar, and tensor charges using lattice QCD at the physical pion mass. *Phys. Rev. D*, 99(11):114505, 2019.

[37] 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.

[38] Y. Zhou, C. Cocuzza, F. Delcarro, W. Melnitchouk, A. Metz, and N. Sato. Revisiting quark and gluon polarization in the proton at the EIC. *Phys. Rev. D*, 104(3):034028, 2021.

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

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

[41] 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.

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

[43] V. N. Gribov and L. N. Lipatov. e+ e- pair annihilation and deep inelastic e p scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15:675–684, 1972.

[44] J. F. Owens. On the Q**2 Dependence of Parton Fragmentation Functions. *Phys. Lett. B*, 76:85–88, 1978.

[45] Tsuneo Uematsu. Q**2 Dependence of Quark and Gluon Fragmentation Functions in a Parton Picture Based on QCD. *Phys. Lett. B*, 79:97–102, 1978.

[46] Howard Georgi and H. David Politzer. Quark Decay Functions and Heavy Hadron Production in QCD. *Nucl. Phys. B*, 136:445–460, 1978.

[47] G. Curci, W. Furmanski, and R. Petronzio. Evolution of Parton Densities Beyond Leading Order: The Nonsinglet Case. *Nucl. Phys. B*, 175:27–92, 1980.

[48] W. Furmanski and R. Petronzio. Singlet Parton Densities Beyond Leading Order. *Phys. Lett. B*, 97:437–442, 1980.

[49] E. G. Floratos, C. Kounnas, and R. Lacaze. Higher Order QCD Effects in Inclusive Annihilation and Deep Inelastic Scattering. *Nucl. Phys. B*, 192:417–462, 1981.

[50] J. Kalinowski, K. Konishi, and T. R. Taylor. JET CALCULUS BEYOND LEADING LOGARITHMS. *Nucl. Phys. B*, 181:221–252, 1981.

[51] J. Kalinowski, K. Konishi, P. N. Scharbach, and T. R. Taylor. RESOLVING QCD JETS BEYOND LEADING ORDER: QUARK DECAY PROBABILITIES. *Nucl. Phys. B*, 181:253–276, 1981.

[52] M. Stratmann and W. Vogelsang. Next-to-leading order evolution of polarized and unpolarized fragmentation functions. *Nucl. Phys. B*, 496:41–65, 1997.

[53] A. Mitov, S. Moch, and A. Vogt. Next-to-Next-to-Leading Order Evolution of Non-Singlet Fragmentation Functions. *Phys. Lett. B*, 638:61–67, 2006.

[54] S. Moch and A. Vogt. On third-order timelike splitting functions and top-mediated Higgs decay into hadrons. *Phys. Lett. B*, 659:290–296, 2008.

[55] A. A. Almasy, S. Moch, and A. Vogt. On the Next-to-Next-to-Leading Order Evolution of Flavour-Singlet Fragmentation Functions. *Nucl. Phys. B*, 854:133–152, 2012.

[56] C. Cocuzza, W. Melnitchouk, A. Metz, and N. Sato. Bayesian Monte Carlo extraction of the sea asymmetry with SeaQuest and STAR data. *Phys. Rev. D*, 104(7):074031, 2021.

[57] C. Cocuzza, C. E. Keppel, H. Liu, W. Melnitchouk, A. Metz, N. Sato, and A. W. Thomas. Isovector EMC Effect from Global QCD Analysis with MARATHON Data. *Phys. Rev. Lett.*, 127(24):242001, 2021.

[58] C. Cocuzza, W. Melnitchouk, A. Metz, and N. Sato. Polarized antimatter in the proton from a global QCD analysis. *Phys. Rev. D*, 106(3):L031502, 2022.

[59] C. Cocuzza, W. Melnitchouk, A. Metz, D. Pitonyak, A. Prokudin, and N. Sato. On the definition and evolution of dihadron fragmentation functions. *In preparation*, 2023.

[60] C. Cocuzza, W. Melnitchouk, A. Metz, D. Pitonyak, A. Prokudin, N. Sato, and R Seidl. Tensor charge of the nucleon from dihadron measurements. *In preparation*, 2023.

[61] Shohini Bhattacharya, Christopher Cocuzza, and Andreas Metz. Generalized quasi parton distributions in a diquark spectator model. *Phys. Lett. B*, 788:453–463, 2019.

[62] Shohini Bhattacharya, Christopher Cocuzza, and Andreas Metz. Exploring twist-2 GPDs through quasidistributions in a diquark spectator model. *Phys. Rev. D*, 102(5):054021, 2020.

[63] D. F. Geesaman and P. E. Reimer. The sea of quarks and antiquarks in the nucleon. *Rept. Prog. Phys.*, 82(4):046301, 2019.

[64] D. A. Ross and Christopher T. Sachrajda. Flavor Symmetry Breaking in anti-Quark Distributions. *Nucl. Phys. B*, 149:497–516, 1979.

[65] Anthony William Thomas. A Limit on the Pionic Component of the Nucleon Through SU(3) Flavor Breaking in the Sea. *Phys. Lett. B*, 126:97–100, 1983.

[66] Andreas W. Schreiber, Piet J. Mulders, A. I. Signal, and Anthony William Thomas. The Pion cloud of the nucleon and its effect on deep inelastic structure. *Phys. Rev. D*, 45:3069–3078, 1992.

[67] E. M. Henley and G. A. Miller. Excess of anti-D over anti-U in the proton sea quark distribution. *Phys. Lett. B*, 251:453–454, 1990.

[68] S. Kumano and J. T. Londergan. Origin of SU(2) flavor symmetry breaking in anti-quark distributions. *Phys. Rev. D*, 44:717–724, 1991.

[69] W. Melnitchouk, Anthony William Thomas, and A. I. Signal. Gottfried sum rule and the shape of F2(p) - F2(n). *Z. Phys. A*, 340:85–92, 1991.

[70] J. Speth and Anthony William Thomas. Mesonic contributions to the spin and flavor structure of the nucleon. *Adv. Nucl. Phys.*, 24:83–149, 1997.

[71] S. Kumano. Flavor asymmetry of anti-quark distributions in the nucleon. *Phys. Rept.*, 303:183–257, 1998.

[72] W. Melnitchouk, J. Speth, and Anthony William Thomas. Dynamics of light anti-quarks in the proton. *Phys. Rev. D*, 59:014033, 1998.

[73] P. V. Pobylitsa, Maxim V. Polyakov, K. Goeke, T. Watabe, and C. Weiss. Isovector unpolarized quark distribution in the nucleon in the large N(c) limit. *Phys. Rev. D*, 59:034024, 1999.

[74] Wen-Chen Chang and Jen-Chieh Peng. Flavor Structure of the Nucleon Sea. *Prog. Part. Nucl. Phys.*, 79:95–135, 2014.

[75] Mary Alberg and Gerald A. Miller. Chiral Light Front Perturbation Theory and the Flavor Dependence of the Light-Quark Nucleon Sea. *Phys. Rev. C*, 100(3):035205, 2019.

[76] R. D. Field and R. P. Feynman. Quark Elastic Scattering as a Source of High Transverse Momentum Mesons. *Phys. Rev. D*, 15:2590–2616, 1977.

[77] Andreas W. Schreiber, A. I. Signal, and Anthony William Thomas. Structure functions in the bag model. *Phys. Rev. D*, 44:2653–2662, 1991.

[78] Claude Bourrely and Jacques Soffer. Parton distributions from W+- and Z production in polarized p p and p n collisions at RHIC. *Nucl. Phys. B*, 423:329–348, 1994.

[79] C. Bourrely and Jacques Soffer. Experimental evidence for simple relations between unpolarized and polarized parton distributions. *Phys. Rev. D*, 51:2108–2113, 1995.

[80] Fernanda Monti Steffens and Anthony William Thomas. The Flavor asymmetry of the nucleon sea. *Phys. Rev. C*, 55:900–908, 1997.

[81] David J. Broadhurst, A. L. Kataev, and C. J. Maxwell. Comparison of the Gottfried and Adler sum rules within the large N(c) expansion. *Phys. Lett. B*, 590:76–85, 2004.

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

[83] Yan-Qing Ma and Jian-Wei Qiu. Extracting Parton Distribution Functions from Lattice QCD Calculations. *Phys. Rev. D*, 98(7):074021, 2018.

[84] Kostas Orginos, Anatoly Radyushkin, Joseph Karpie, and Savvas Zafeiropoulos. Lattice QCD exploration of parton pseudo-distribution functions. *Phys. Rev. D*, 96(9):094503, 2017.

[85] 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.

[86] A. S. Ito et al. Measurement of the Continuum of Dimuons Produced in High-Energy Proton - Nucleus Collisions. *Phys. Rev. D*, 23:604–633, 1981.

[87] P. Amaudruz et al. The Gottfried sum from the ratio F2(n) / F2(p). *Phys. Rev. Lett.*, 66:2712–2715, 1991.

[88] M. Arneodo et al. A Reevaluation of the Gottfried sum. *Phys. Rev. D*, 50:R1–R3, 1994.

[89] Kurt Gottfried. Sum rule for high-energy electron - proton scattering. *Phys. Rev. Lett.*, 18:1174, 1967.

[90] A. Baldit et al. Study of the isospin symmetry breaking in the light quark sea of the nucleon from the Drell-Yan process. *Phys. Lett. B*, 332:244–250, 1994.

[91] K. Ackerstaff et al. The Flavor asymmetry of the light quark sea from semiinclusive deep inelastic scattering. *Phys. Rev. Lett.*, 81:5519–5523, 1998.

[92] Jason C. Webb. *Measurement of continuum dimuon production in 800-GeV/C proton nucleon collisions*. PhD thesis, New Mexico State U., 2003.

[93] E. A. Hawker et al. Measurement of the light anti-quark flavor asymmetry in the nucleon sea. *Phys. Rev. Lett.*, 80:3715–3718, 1998.

[94] R. S. Towell et al. Improved measurement of the anti-d / anti-u asymmetry in the nucleon sea. *Phys. Rev. D*, 64:052002, 2001.

[95] J. Dove et al. The asymmetry of antimatter in the proton. *Nature*, 590(7847):561–565, 2021.

[96] Jaroslav Adam et al. Measurements of W and Z/γ^* cross sections and their ratios in p+p collisions at RHIC. *Phys. Rev. D*, 103(1):012001, 2021.

[97] Jacob J. Ethier and Emanuele R. Nocera. Parton Distributions in Nucleons and Nuclei. *Ann. Rev. Nucl. Part. Sci.*, 70:43–76, 2020.

[98] J. Pumplin, D. R. Stump, and W. K. Tung. Multivariate fitting and the error matrix in global analysis of data. *Phys. Rev. D*, 65:014011, 2001.

[99] D. Stump, J. Pumplin, R. Brock, D. Casey, J. Huston, J. Kalk, H. L. Lai, and W. K. Tung. Uncertainties of predictions from parton distribution functions. 1. The Lagrange multiplier method. *Phys. Rev. D*, 65:014012, 2001.

[100] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.

[101] Thomas Becher, Matthias Neubert, and Gang Xu. Dynamical Threshold Enhancement and Resummation in Drell-Yan Production. *JHEP*, 07:030, 2008.

[102] P. J. Ehlers, A. Accardi, L. T. Brady, and W. Melnitchouk. Nuclear effects in the proton-deuteron Drell-Yan process. *Phys. Rev. D*, 90(1):014010, 2014.

[103] Felix Ringer and Werner Vogelsang. Single-Spin Asymmetries in W Boson Production at Next-to-Leading Order. *Phys. Rev. D*, 91(9):094033, 2015.

[104] B. Jager, M. Stratmann, and W. Vogelsang. Single inclusive jet production in polarized pp collisions at $O(alpha_s^3)$. *Phys. Rev. D*, 70:034010, 2004.

[105] Zhong-Bo Kang, Felix Ringer, and Wouter J. Waalewijn. The Energy Distribution of Subjets and the Jet Shape. *JHEP*, 07:064, 2017.

[106] Marco Stratmann and Werner Vogelsang. Towards a global analysis of polarized parton distributions. *Phys. Rev. D*, 64:114007, 2001.

[107] D. de Florian, M. Stratmann, and W. Vogelsang. QCD analysis of unpolarized and polarized Lambda baryon production in leading and next-to-leading order. *Phys. Rev. D*, 57:5811–5824, 1998.

[108] Guido Altarelli, R. Keith Ellis, G. Martinelli, and So-Young Pi. Processes Involving Fragmentation Functions Beyond the Leading Order in QCD. *Nucl. Phys. B*, 160:301–329, 1979.

[109] P. Nason and B. R. Webber. Scaling violation in e+ e- fragmentation functions: QCD evolution, hadronization and heavy quark mass effects. *Nucl. Phys. B*, 421:473–517, 1994. [Erratum: Nucl.Phys.B 480, 755 (1996)].

[110] W. Furmanski and R. Petronzio. Lepton - Hadron Processes Beyond Leading Order in Quantum Chromodynamics. *Z. Phys. C*, 11:293, 1982.

[111] Dirk Graudenz. One particle inclusive processes in deeply inelastic lepton - nucleon scattering. *Nucl. Phys. B*, 432:351–376, 1994.

[112] Nobuo Sato, W. Melnitchouk, S. E. Kuhn, J. J. Ethier, and A. Accardi. Iterative Monte Carlo analysis of spin-dependent parton distributions. *Phys. Rev. D*, 93(7):074005, 2016.

[113] Nobuo Sato, J. J. Ethier, W. Melnitchouk, M. Hirai, S. Kumano, and A. Accardi. First Monte Carlo analysis of fragmentation functions from single-inclusive e^+e^- annihilation. *Phys. Rev. D*, 94(11):114004, 2016.

[114] J. J. Ethier, N. Sato, and W. Melnitchouk. First simultaneous extraction of spin-dependent parton distributions and fragmentation functions from a global QCD analysis. *Phys. Rev. Lett.*, 119(13):132001, 2017.

[115] N. Sato, C. Andres, J. J. Ethier, and W. Melnitchouk. Strange quark suppression from a simultaneous Monte Carlo analysis of parton distributions and fragmentation functions. *Phys. Rev. D*, 101(7):074020, 2020.

[116] Eric Moffat, Wally Melnitchouk, T. C. Rogers, and Nobuo Sato. Simultaneous Monte Carlo analysis of parton densities and fragmentation functions. *Phys. Rev. D*, 104(1):016015, 2021.

[117] M. Gluck and E. Reya. Operator Mixing and Scaling Deviations in Asymptotically Free Field Theories. *Phys. Rev. D*, 14:3034–3044, 1976.

[118] M. Gluck, E. Reya, and A. Vogt. Radiatively generated parton distributions for high-energy collisions. *Z. Phys. C*, 48:471–482, 1990.

[119] David A. Kosower. Evolution of parton distributions. *Nucl. Phys. B*, 506:439–467, 1997.

[120] Stefan Weinzierl. Fast evolution of parton distributions. *Comput. Phys. Commun.*, 148:314–326, 2002.

[121] A. Vogt. Efficient evolution of unpolarized and polarized parton distributions with QCD-PEGASUS. *Comput. Phys. Commun.*, 170:65–92, 2005.

[122] A. C. Benvenuti et al. A High Statistics Measurement of the Proton Structure Functions F(2) (x, Q**2) and R from Deep Inelastic Muon Scattering at High Q**2. *Phys. Lett. B*, 223:485–489, 1989.

[123] M. Arneodo et al. Measurement of the proton and deuteron structure functions, F2(p) and F2(d), and of the ratio sigma-L / sigma-T. *Nucl. Phys. B*, 483:3–43, 1997.

[124] M. Arneodo et al. Accurate measurement of F2(d) / F2(p) and R**d - R**p. *Nucl. Phys. B*, 487:3–26, 1997.

[125] L. W. Whitlow, E. M. Riordan, S. Dasu, Stephen Rock, and A. Bodek. Precise measurements of the proton and deuteron structure functions from a global analysis of the SLAC deep inelastic electron scattering cross-sections. *Phys. Lett. B*, 282:475–482, 1992.

[126] S. P. Malace et al. Applications of quark-hadron duality in F(2) structure function. *Phys. Rev. C*, 80:035207, 2009.

[127] S. Tkachenko et al. Measurement of the structure function of the nearly free neutron using spectator tagging in inelastic ^{2}H(e, e'p)X scattering with CLAS. *Phys. Rev. C*, 89:045206, 2014. [Addendum: Phys.Rev.C 90, 059901 (2014)].

[128] J. Seely et al. New measurements of the EMC effect in very light nuclei. *Phys. Rev. Lett.*, 103:202301, 2009.

[129] D. Abrams et al. Measurement of the Nucleon F_2^n/F_2^p Structure Function Ratio by the Jefferson Lab MARATHON Tritium/Helium-3 Deep Inelastic Scattering Experiment. *Phys. Rev. Lett.*, 128(13):132003, 2022.

[130] H. Abramowicz et al. Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data. *Eur. Phys. J. C*, 75(12):580, 2015.

[131] Timo Antero Aaltonen et al. Measurement of $d\sigma/dy$ of Drell-Yan e^+e^- pairs in the Z Mass Region from $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Lett. B*, 692:232–239, 2010.

[132] V. M. Abazov et al. Measurement of the Shape of the Boson Rapidity Distribution for $p\bar{p} \to Z/\gamma^* \to e^+e^- + X$ Events Produced at $\sqrt{s}$ of 1.96-TeV. *Phys. Rev. D*, 76:012003, 2007.

[133] T. Aaltonen et al. Direct Measurement of the W Production Charge Asymmetry in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. Lett.*, 102:181801, 2009.

[134] Victor Mukhamedovich Abazov et al. Measurement of the W Boson Production Charge Asymmetry in $p\bar{p} \to W + X \to e\nu + X$ Events at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. Lett.*, 112(15):151803, 2014. [Erratum: Phys.Rev.Lett. 114, 049901 (2015)].

[135] Serguei Chatrchyan et al. Measurement of the lepton charge asymmetry in inclusive W production in pp collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 04:050, 2011.

[136] Serguei Chatrchyan et al. Measurement of the Electron Charge Asymmetry in Inclusive W Production in pp Collisions at $\sqrt{s} = 7$ TeV. *Phys. Rev. Lett.*, 109:111806, 2012.

[137] Serguei Chatrchyan et al. Measurement of the Muon Charge Asymmetry in Inclusive $pp \to W + X$ Production at $\sqrt{s} = 7$ TeV and an Improved Determination of Light Parton Distribution Functions. *Phys. Rev. D*, 90(3):032004, 2014.

[138] Vardan Khachatryan et al. Measurement of the differential cross section and charge asymmetry for inclusive pp $\to \mathrm{W}^{\pm} + X$ production at $\sqrt{s} = 8$ TeV. *Eur. Phys. J. C*, 76(8):469, 2016.

[139] Roel Aaij et al. Measurement of the forward W boson cross-section in pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$. *JHEP*, 12:079, 2014.

[140] Roel Aaij et al. Measurement of forward W and Z boson production in association with jets in proton-proton collisions at $\sqrt{s} = 8$ TeV. *JHEP*, 05:131, 2016.

[141] Victor Mukhamedovich Abazov et al. Measurement of the inclusive jet cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. D*, 85:052006, 2012.

[142] A. Abulencia et al. Measurement of the Inclusive Jet Cross Section using the k_{T} algorithmin $p\overline{p}$ Collisions at $\sqrt{s} = 1.96$ TeV with the CDF II Detector. *Phys. Rev. D*, 75:092006, 2007. [Erratum: Phys.Rev.D 75, 119901 (2007)].

[143] B. I. Abelev et al. Longitudinal double-spin asymmetry and cross section for inclusive jet production in polarized proton collisions at s**(1/2) = 200-GeV. *Phys. Rev. Lett.*, 97:252001, 2006.

[144] C. Adolph et al. Multiplicities of charged pions and charged hadrons from deep-inelastic scattering of muons off an isoscalar target. *Phys. Lett. B*, 764:1–10, 2017.

[145] C. Adolph et al. Multiplicities of charged kaons from deep-inelastic muon scattering off an isoscalar target. *Phys. Lett. B*, 767:133–141, 2017.

[146] J. V. Guerrero, J. J. Ethier, A. Accardi, S. W. Casper, and W. Melnitchouk. Hadron mass corrections in semi-inclusive deep-inelastic scattering. *JHEP*, 09:169, 2015.

[147] R. Brandelik et al. Evidence for charged primary partons in e+ e- $\rightarrow$ 2 jets. *Physics Letters B*, 100(4):357363, 1981.

[148] M. Althoff et al. Jet Production and Fragmentation in e+ e- Annihilation at 12-GeV to 43-GeV. *Z. Phys. C*, 22:307–340, 1984.

[149] R. Brandelik et al. Charged Pion, Kaon, Proton and anti-Proton Production in High-Energy e+ e- Annihilation. *Phys. Lett. B*, 94:444–449, 1980.

[150] M. Althoff et al. Charged Hadron Composition of the Final State in e+ e- Annihilation at High-Energies. *Z. Phys. C*, 17:5–15, 1983.

[151] W. Braunschweig et al. Pion, Kaon and Proton Cross-sections in e^+e^- Annihilation at 34-GeV and 44-GeV Center-of-mass Energy. *Z. Phys. C*, 42:189, 1989.

[152] H. Albrecht et al. Inclusive Production of Charged Pions, Charged and Neutral Kaons and Anti-protons in e^+e^- Annihilation at 10-GeV and in Direct Υ Decays. *Z. Phys. C*, 44:547, 1989.

[153] Xing-Qi Lu. HEAVY QUARK JETS FROM E+ E- ANNIHILATION AT 29-GeV. Other thesis, 1986.

[154] H. Aihara et al. Charged hadron production in e^+e^- annihilation at 29-GeV. *Phys. Rev. Lett.*, 52:577, 1984.

[155] H. Aihara et al. Charged hadron inclusive cross-sections and fractions in e^+e^- annihiliation $\sqrt{s} = 29$ GeV. *Phys. Rev. Lett.*, 61:1263, 1988.

[156] M. Derrick et al. Hadron Production in e^+e^- Annihilation at $\sqrt{s}$ = 29-GeV. *Phys. Rev. D*, 35:2639, 1987.

[157] Koya Abe et al. Production of π^+, π^-, K^+, K^-, p and $\bar{\text{p}}$ in Light (uds), c and b Jets from Z^0 Decays. *Phys. Rev. D*, 69:072003, 2004.

[158] J. P. Lees et al. Production of charged pions, kaons, and protons in e^+e^- annihilations into hadrons at $\sqrt{s}$=10.54 GeV. *Phys. Rev. D*, 88:032011, 2013.

[159] R. Akers et al. Measurement of the production rates of charged hadrons in e+ e- annihilation at the Z0. *Z. Phys. C*, 63:181–196, 1994.

[160] G. Abbiendi et al. Leading particle production in light flavor jets. *Eur. Phys. J. C*, 16:407–421, 2000.

[161] D. Buskulic et al. Inclusive pi+-, K+- and (p, anti-p) differential cross-sections at the Z resonance. *Z. Phys. C*, 66:355–366, 1995.

[162] P. Abreu et al. pi+-, K+-, p and anti-p production in Z0 —> q anti-q, Z0 —> b anti-b, Z0 —> u anti-u, d anti-d, s anti-s. *Eur. Phys. J. C*, 5:585–620, 1998.

[163] R. Itoh et al. Measurement of inclusive particle spectra and test of MLLA prediction in e+ e- annihilation at s**(1/2) = 58-GeV. *Phys. Lett. B*, 345:335–342, 1995.

[164] M. Leitgab et al. Precision Measurement of Charged Pion and Kaon Differential Cross Sections in e+e- Annihilation at s=10.52 GeV. *Phys. Rev. Lett.*, 111:062002, 2013.

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

[166] 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.

[167] A. Accardi, L. T. Brady, W. Melnitchouk, J. F. Owens, and N. Sato. Constraints on large-x parton distributions from new weak boson production and deep-inelastic scattering data. *Phys. Rev. D*, 93(11):114017, 2016.

[168] 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.

[169] 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.

[170] Georges Aad et al. Measurement of the inclusive $W^{\pm}$ and Z/gamma cross sections in the electron and muon decay channels in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys. Rev. D*, 85:072004, 2012.

[171] Morad Aaboud et al. Precision measurement and interpretation of inclusive W^+ , W^- and Z/γ^* production cross sections with the ATLAS detector. *Eur. Phys. J. C*, 77(6):367, 2017.

[172] Georges Aad et al. Measurement of the low-mass Drell-Yan differential cross section at $\sqrt{s} = 7$ TeV using the ATLAS detector. *JHEP*, 06:112, 2014.

[173] Georges Aad et al. Measurement of the high-mass Drell–Yan differential cross-section in pp collisions at sqrt(s)=7 TeV with the ATLAS detector. *Phys. Lett. B*, 725:223–242, 2013.

[174] 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.

[175] M. Aaboud et al. Measurement of the Drell-Yan triple-differential cross section in *pp* collisions at $\sqrt{s} = 8$ TeV. *JHEP*, 12:059, 2017.

[176] Georges Aad et al. Measurement of the cross-section and charge asymmetry of W bosons produced in proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 79(9):760, 2019.

[177] Richard D. Ball et al. Parton distributions from high-precision collider data. *Eur. Phys. J. C*, 77(10):663, 2017.

[178] J. Dove et al. Measurement of flavor asymmetry of light-quark sea in the proton with Drell-Yan dimuon production in $p + p$ and $p + d$ collisions at 120 GeV. 12 2022.

[179] Sanghwa Park, Alberto Accardi, Xiaoxian Jing, and J. F. Owens. CJ15 global PDF analysis with new electroweak data from the STAR and SeaQuest experiments. In *28th International Workshop on Deep Inelastic Scattering and Related Subjects*, 8 2021.

[180] Marco Guzzi et al. NNLO constraints on proton PDFs from the SeaQuest and STAR experiments and other developments in the CTEQ-TEA global analysis. *SciPost Phys. Proc.*, 8:005, 2022.

[181] Yusupujiang Salamu, Chueng-Ryong Ji, W. Melnitchouk, and P. Wang. $\bar{d} - \bar{u}$ asymmetry in the proton in chiral effective theory. *Phys. Rev. Lett.*, 114:122001, 2015.

[182] Y. Salamu, Chueng-Ryong Ji, W. Melnitchouk, A. W. Thomas, P. Wang, and X. G. Wang. Parton distributions from nonlocal chiral SU(3) effective theory: Flavor asymmetries. *Phys. Rev. D*, 100(9):094026, 2019.

[183] M. Burkardt, K. S. Hendricks, Chueng-Ryong Ji, W. Melnitchouk, and A. W. Thomas. Pion momentum distributions in the nucleon in chiral effective theory. *Phys. Rev. D*, 87(5):056009, 2013.

[184] 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.

[185] J. R. McKenney, Nobuo Sato, W. Melnitchouk, and Chueng-Ryong Ji. Pion structure function from leading neutron electroproduction and SU(2) flavor asymmetry. *Phys. Rev. D*, 93(5):054011, 2016.

[186] P. C. Barry, N. Sato, W. Melnitchouk, and Chueng-Ryong Ji. First Monte Carlo Global QCD Analysis of Pion Parton Distributions. *Phys. Rev. Lett.*, 121(15):152001, 2018.

[187] W. Melnitchouk and Anthony William Thomas. Neutron / proton structure function ratio at large x. *Phys. Lett. B*, 377:11–17, 1996.

[188] S. Kuhlmann, J. Huston, J. Morfin, Fredrick I. Olness, J. Pumplin, J. F. Owens, W. K. Tung, and J. J. Whitmore. Large x parton distributions. *Phys. Lett. B*, 476:291–296, 2000.

[189] S. I. Alekhin, S. A. Kulagin, and R. Petti. Nuclear Effects in the Deuteron and Constraints on the d/u Ratio. *Phys. Rev. D*, 96(5):054005, 2017.

[190] J. Arrington, J. G. Rubin, and W. Melnitchouk. How Well Do We Know The Neutron Structure Function? *Phys. Rev. Lett.*, 108:252001, 2012.

[191] Iraj Ruhi Afnan, Francois Rene Pierre Bissey, J. Gomez, A. T. Katramatou, W. Melnitchouk, G. G. Petratos, and Anthony William Thomas. Neutron structure function and A = 3 mirror nuclei. *Phys. Lett. B*, 493:36–42, 2000.

[192] Iraj Ruhi Afnan, Francois Rene Pierre Bissey, J. Gomez, A. T. Katramatou, Simonetta Liuti, W. Melnitchouk, G. G. Petratos, and Anthony William Thomas. Deep inelastic scattering from A = 3 nuclei and the neutron structure function. *Phys. Rev. C*, 68:035201, 2003.

[193] J. J. Aubert et al. The ratio of the nucleon structure functions $F2_n$ for iron and deuterium. *Phys. Lett. B*, 123:275–278, 1983.

[194] Sergey A. Kulagin and R. Petti. Global study of nuclear structure functions. *Nucl. Phys. A*, 765:126–187, 2006.

[195] J. Gomez et al. Measurement of the A-dependence of deep inelastic electron scattering. *Phys. Rev. D*, 49:4348–4372, 1994.

[196] S. A. Kulagin and R. Petti. Structure functions for light nuclei. *Phys. Rev. C*, 82:054614, 2010.

[197] Ingo Schienbein et al. A Review of Target Mass Corrections. *J. Phys. G*, 35:053101, 2008.

[198] E. Moffat, T. C. Rogers, W. Melnitchouk, N. Sato, and F. Steffens. What does kinematical target mass sensitivity in DIS reveal about hadron structure? *Phys. Rev. D*, 99(9):096008, 2019.

[199] R. Ruiz et al. Target mass corrections in lepton-nucleus DIS: theory and applications to nuclear PDFs. 1 2023.

[200] M. A. G. Aivazis, John C. Collins, Fredrick I. Olness, and Wu-Ki Tung. Leptoproduction of heavy quarks. 2. A Unified QCD formulation of charged and neutral current processes from fixed target to collider energies. *Phys. Rev. D*, 50:3102–3118, 1994.

[201] Otto Nachtmann. Positivity constraints for anomalous dimensions. *Nucl. Phys. B*, 63:237–247, 1973.

[202] Howard Georgi and H. David Politzer. Freedom at Moderate Energies: Masses in Color Dynamics. *Phys. Rev. D*, 14:1829, 1976.

[203] L. T. Brady, A. Accardi, T. J. Hobbs, and W. Melnitchouk. Next-to leading order analysis of target mass corrections to structure functions and asymmetries. *Phys. Rev. D*, 84:074008, 2011. [Erratum: Phys.Rev.D 85, 039902 (2012)].

[204] W. Melnitchouk, Andreas W. Schreiber, and Anthony William Thomas. Deep inelastic scattering from off-shell nucleons. *Phys. Rev. D*, 49:1183–1198, 1994.

[205] W. Melnitchouk, Andreas W. Schreiber, and Anthony William Thomas. Relativistic deuteron structure function. *Phys. Lett. B*, 335:11–16, 1994.

[206] Sergey A. Kulagin, G. Piller, and W. Weise. Shadowing, binding and off-shell effects in nuclear deep inelastic scattering. *Phys. Rev. C*, 50:1154–1169, 1994.

[207] Sergey A. Kulagin, W. Melnitchouk, G. Piller, and W. Weise. Spin dependent nuclear structure functions: General approach with application to the deuteron. *Phys. Rev. C*, 52:932–946, 1995.

[208] P. Fernandez de Cordoba, E. Marco, H. Muther, E. Oset, and Amand Faessler. Deep inelastic lepton scattering in nuclei at x $>$ 1 and the nucleon spectral function. *Nucl. Phys. A*, 611:514–538, 1996.

[209] A. J. Tropiano, J. J. Ethier, W. Melnitchouk, and N. Sato. Deep-inelastic and quasielastic electron scattering from $A = 3$ nuclei. *Phys. Rev. C*, 99(3):035201, 2019.

[210] J. T. Londergan, J. C. Peng, and A. W. Thomas. Charge Symmetry at the Partonic Level. *Rev. Mod. Phys.*, 82:2009–2052, 2010.

[211] M. Lacombe, B. Loiseau, R. Vinh Mau, J. Cote, P. Pires, and R. de Tourreil. Parametrization of the deuteron wave function of the Paris n-n potential. *Phys. Lett. B*, 101:139–140, 1981.

[212] A. Kievsky, E. Pace, G. Salme, and M. Viviani. Neutron electromagnetic form-factors and inclusive scattering of polarized electrons by polarized He-3 and He-3 targets. *Phys. Rev. C*, 56:64–75, 1997.

[213] Robert B. Wiringa, V. G. J. Stoks, and R. Schiavilla. An Accurate nucleon-nucleon potential with charge independence breaking. *Phys. Rev. C*, 51:38–51, 1995.

[214] R. Machleidt. The High precision, charge dependent Bonn nucleon-nucleon potential (CD-Bonn). *Phys. Rev. C*, 63:024001, 2001.

[215] Franz Gross and Alfred Stadler. Covariant spectator theory of np scattering: Phase shifts obtained from precision fits to data below 350-MeV. *Phys. Rev. C*, 78:014005, 2008.

[216] Franz Gross and Alfred Stadler. Covariant spectator theory of *np* scattering: Effective range expansions and relativistic deuteron wave functions. *Phys. Rev. C*, 82:034004, 2010.

[217] R. W. Schulze and P. U. Sauer. Inelastic electron scattering from the three-nucleon bound states with polarization. *Phys. Rev. C*, 48:38–63, 1993.

[218] K. Kovarik et al. nCTEQ15 - Global analysis of nuclear parton distributions with uncertainties in the CTEQ framework. *Phys. Rev. D*, 93(8):085037, 2016.

[219] Kari J. Eskola, Petja Paakkinen, Hannu Paukkunen, and Carlos A. Salgado. EPPS16: Nuclear parton distributions with LHC data. *Eur. Phys. J. C*, 77(3):163, 2017.

[220] Rabah Abdul Khalek et al. A first determination of parton distributions with theoretical uncertainties. *Eur. Phys. J.*, C:79:838, 2019.

[221] Marina Walt, Ilkka Helenius, and Werner Vogelsang. Open-source QCD analysis of nuclear parton distribution functions at NLO and NNLO. *Phys. Rev. D*, 100(9):096015, 2019.

[222] J. Ashman et al. An Investigation of the Spin Structure of the Proton in Deep Inelastic Scattering of Polarized Muons on Polarized Protons. *Nucl. Phys. B*, 328:1, 1989.

[223] B. Adeva et al. Spin asymmetries A(1) and structure functions g1 of the proton and the deuteron from polarized high-energy muon scattering. *Phys. Rev. D*, 58:112001, 1998.

[224] B. Adeva et al. Spin asymmetries A(1) of the proton and the deuteron in the low x and low Q**2 region from polarized high-energy muon scattering. *Phys. Rev. D*, 60:072004, 1999. [Erratum: Phys.Rev.D 62, 079902 (2000)].

[225] V. Yu. Alexakhin et al. The Deuteron Spin-dependent Structure Function g1(d) and its First Moment. *Phys. Lett. B*, 647:8–17, 2007.

[226] M. G. Alekseev et al. The Spin-dependent Structure Function of the Proton g_1^p and a Test of the Bjorken Sum Rule. *Phys. Lett. B*, 690:466–472, 2010.

[227] C Adolph et al. Longitudinal double spin asymmetries in single hadron quasi-real photoproduction at high p_T. *Phys. Lett. B*, 753:573–579, 2016.

[228] Guenter Baum et al. A New Measurement of Deep Inelastic e p Asymmetries. *Phys. Rev. Lett.*, 51:1135, 1983.

[229] P. L. Anthony et al. Deep inelastic scattering of polarized electrons by polarized He-3 and the study of the neutron spin structure. *Phys. Rev. D*, 54:6620–6650, 1996.

[230] K. Abe et al. Precision determination of the neutron spin structure function g1(n). *Phys. Rev. Lett.*, 79:26–30, 1997.

[231] K. Abe et al. Measurements of the proton and deuteron spin structure functions g(1) and g(2). *Phys. Rev. D*, 58:112003, 1998.

[232] P. L. Anthony et al. Measurement of the deuteron spin structure function g1(d)(x) for 1-(GeV/c)**2 < Q**2 < 40-(GeV/c)**2. *Phys. Lett. B*, 463:339–345, 1999.

[233] P. L Anthony et al. Measurements of the Q**2 dependence of the proton and neutron spin structure functions g(1)**p and g(1)**n. *Phys. Lett. B*, 493:19–28, 2000.

[234] K. Ackerstaff et al. Measurement of the neutron spin structure function g1(n) with a polarized He-3 internal target. *Phys. Lett. B*, 404:383–389, 1997.

[235] A. Airapetian et al. Precise determination of the spin structure function g(1) of the proton, deuteron and neutron. *Phys. Rev. D*, 75:012007, 2007.

[236] B. I. Abelev et al. Longitudinal double-spin asymmetry for inclusive jet production in p+p collisions at s**(1/2) = 200-GeV. *Phys. Rev. Lett.*, 100:232003, 2008.

[237] L. Adamczyk et al. Longitudinal and transverse spin asymmetries for inclusive jet production at mid-rapidity in polarized $p + p$ collisions at $\sqrt{s} = 200$ GeV. *Phys. Rev. D*, 86:032006, 2012.

[238] L. Adamczyk et al. Precision Measurement of the Longitudinal Double-spin Asymmetry for Inclusive Jet Production in Polarized Proton Collisions at $\sqrt{s} = 200$ GeV. *Phys. Rev. Lett.*, 115(9):092002, 2015.

[239] J. Adam et al. Longitudinal double-spin asymmetry for inclusive jet and dijet production in pp collisions at $\sqrt{s} = 510$ GeV. *Phys. Rev. D*, 100(5):052005, 2019.

[240] M. S. Abdallah et al. Longitudinal double-spin asymmetry for inclusive jet and dijet production in polarized proton collisions at $\sqrt{s} = 510$ GeV. *Phys. Rev. D*, 105(9):092011, 2022.

[241] Mohamed Abdallah et al. Longitudinal double-spin asymmetry for inclusive jet and dijet production in polarized proton collisions at $\sqrt{s} = 200$ GeV. *Phys. Rev. D*, 103(9):L091103, 2021.

[242] A. Adare et al. Event Structure and Double Helicity Asymmetry in Jet Production from Polarized $p+p$ Collisions at $\sqrt{s} = 200\tilde{}$GeV. *Phys. Rev. D*, 84:012006, 2011.

[243] Daniel de Florian, Rodolfo Sassot, Marco Stratmann, and Werner Vogelsang. Evidence for polarization of gluons in the proton. *Phys. Rev. Lett.*, 113(1):012001, 2014.

[244] Y. Zhou, N. Sato, and W. Melnitchouk. How well do we know the gluon polarization in the proton? *Phys. Rev. D*, 105(7):074022, 2022.

[245] B. Adeva et al. Polarized quark distributions in the nucleon from semiinclusive spin asymmetries. *Phys. Lett. B*, 420:180–190, 1998.

[246] A. Airapetian et al. Quark helicity distributions in the nucleon for up, down, and strange quarks from semi-inclusive deep-inelastic scattering. *Phys. Rev. D*, 71:012003, 2005.

[247] M. G. Alekseev et al. Quark helicity distributions from longitudinal spin asymmetries in muon-proton and muon-deuteron scattering. *Phys. Lett. B*, 693:227–235, 2010.

[248] M. Alekseev et al. Flavour Separation of Helicity Distributions from Deep Inelastic Muon-Deuteron Scattering. *Phys. Lett. B*, 680:217–224, 2009.

[249] Fu-Guang Cao and A. I. Signal. Nonperturbative structure of the polarized nucleon sea. *Phys. Rev. D*, 68:074002, 2003.

[250] Dmitri Diakonov, V. Yu. Petrov, P. V. Pobylitsa, Maxim V. Polyakov, and C. Weiss. Unpolarized and polarized quark distributions in the large N(c) limit. *Phys. Rev. D*, 56:4069–4083, 1997.

[251] M. Wakamatsu and T. Watabe. Do we expect light flavor sea quark asymmetry also for the spin dependent distribution functions of the nucleon? *Phys. Rev. D*, 62:017506, 2000.

[252] Claude Bourrely and Jacques Soffer. New developments in the statistical approach of parton distributions: tests and predictions up to LHC energies. *Nucl. Phys. A*, 941:307–334, 2015.

[253] M. M. Aggarwal et al. Measurement of the parity-violating longitudinal single-spin asymmetry for $W^{\pm}$ boson production in polarized proton-proton collisions at $\sqrt{s} = 500 - GeV$. *Phys. Rev. Lett.*, 106:062002, 2011.

[254] L. Adamczyk et al. Measurement of longitudinal spin asymmetries for weak boson production in polarized proton-proton collisions at RHIC. *Phys. Rev. Lett.*, 113:072301, 2014.

[255] Jaroslav Adam et al. Measurement of the longitudinal spin asymmetries for weak boson production in proton-proton collisions at $\sqrt{s} = 510$ GeV. *Phys. Rev. D*, 99(5):051102, 2019.

[256] A. Adare et al. Measurement of parity-violating spin asymmetries in $W^{\pm}$ production at midrapidity in longitudinally polarized $p + p$ collisions. *Phys. Rev. D*, 93(5):051103, 2016.

[257] A. Adare et al. Cross section and longitudinal single-spin asymmetry A_L for forward $W^{\pm} \to \mu^{\pm}\nu$ production in polarized $p + p$ collisions at $\sqrt{s} = 510$ GeV. *Phys. Rev. D*, 98(3):032007, 2018.

[258] Daniel de Florian, Rodolfo Sassot, Marco Stratmann, and Werner Vogelsang. Extraction of Spin-Dependent Parton Densities and Their Uncertainties. *Phys. Rev. D*, 80:034030, 2009.

[259] Daniel De Florian, Gonzalo Agustín Lucero, Rodolfo Sassot, Marco Stratmann, and Werner Vogelsang. Monte Carlo sampling variant of the DSSV14 set of helicity parton densities. *Phys. Rev. D*, 100(11):114027, 2019.

[260] 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.

[261] Nobuo Sato, J. F. Owens, and Harrison Prosper. Bayesian Reweighting for Global Fits. *Phys. Rev. D*, 89(11):114020, 2014.

[262] C. Cocuzza, W. Melnitchouk, A. Metz, and N. Sato. In preparation.

[263] Sergey A. Kulagin and W. Melnitchouk. Deuteron spin structure functions in the resonance and DIS regions. *Phys. Rev. C*, 77:015210, 2008.

[264] Sergey A. Kulagin and W. Melnitchouk. Spin structure functions of He-3 at finite Q**2. *Phys. Rev. C*, 78:065203, 2008.

[265] J. J. Ethier and W. Melnitchouk. Comparative study of nuclear effects in polarized electron scattering from 3He. *Phys. Rev. C*, 88(5):054001, 2013.

[266] Daniel de Florian, Rodolfo Sassot, and Marco Stratmann. Global analysis of fragmentation functions for pions and kaons and their uncertainties. *Phys. Rev. D*, 75:114010, 2007.

[267] Alessandro Candido, Stefano Forte, and Felix Hekhorn. Can $\overline{\text{MS}}$ parton distributions be negative? *JHEP*, 11:129, 2020.

[268] John Collins, Ted C. Rogers, and Nobuo Sato. Positivity and renormalization of parton densities. 11 2021.

[269] Richard D. Ball, Stefano Forte, Alberto Guffanti, Emanuele R. Nocera, Giovanni Ridolfi, and Juan Rojo. Unbiased determination of polarized parton distributions and their uncertainties. *Nucl. Phys. B*, 874:36–84, 2013.

[270] Guido Altarelli, Stefano Forte, and Giovanni Ridolfi. On positivity of parton distributions. *Nucl. Phys. B*, 534:277–296, 1998.

[271] Stefano Forte, Guido Altarelli, and Giovanni Ridolfi. Are parton distributions positive? *Nucl. Phys. B Proc. Suppl.*, 74:138–141, 1999.

[272] X. Zheng et al. Precision measurement of the neutron spin asymmetries and spin-dependent structure functions in the valence quark region. *Phys. Rev. C*, 70:065207, 2004.

[273] D. S. Parno et al. Precision Measurements of A_1^n in the Deep Inelastic Regime. *Phys. Lett. B*, 744:309–314, 2015.

[274] Y. Prok et al. Precision measurements of g_1 of the proton and the deuteron with 6 GeV electrons. *Phys. Rev. C*, 90(2):025212, 2014.

[275] Robert Fersch et al. Determination of the Proton Spin Structure Functions for $0.05 < Q^2 < 5GeV^2$ using CLAS. *Phys. Rev. C*, 96(6):065208, 2017.

[276] N. Guler et al. Precise determination of the deuteron spin structure at low to moderate Q^2 with CLAS and extraction of the neutron contribution. *Phys. Rev. C*, 92(5):055201, 2015.

[277] K. Ackerstaff et al. Flavor decomposition of the polarized quark distributions in the nucleon from inclusive and semiinclusive deep inelastic scattering. *Phys. Lett. B*, 464:123–134, 1999.

[278] A. I. Signal and Anthony William Thomas. Possible Strength of the Nonperturbative Strange Sea of the Nucleon. *Phys. Lett. B*, 191:205, 1987.

[279] W. Melnitchouk and M. Malheiro. Strangeness in the nucleon on the light cone. *Phys. Rev. C*, 55:431–440, 1997.

[280] Anthony William Thomas, W. Melnitchouk, and Fernanda Monti Steffens. Dynamical symmetry breaking in the sea of the nucleon. *Phys. Rev. Lett.*, 85:2892–2894, 2000.

[281] Stanley J. Brodsky, Matthias Burkardt, and Ivan Schmidt. Perturbative QCD constraints on the shape of polarized quark and gluon distributions. *Nucl. Phys. B*, 441:197–214, 1995.

[282] P. Jimenez-Delgado, A. Accardi, and W. Melnitchouk. Impact of hadronic and nuclear corrections on global analysis of spin-dependent parton distributions. *Phys. Rev. D*, 89(3):034025, 2014.

[283] P. Jimenez-Delgado, H. Avakian, and W. Melnitchouk. Constraints on spin-dependent parton distributions at large x from global QCD analysis. *Phys. Lett. B*, 738:263–267, 2014.

[284] C. Alexandrou, S. Bacchio, M. Constantinou, J. Finkenrath, K. Hadjiyiannakou, K. Jansen, G. Koutsou, H. Panagopoulos, and G. Spanoudes. Complete flavor decomposition of the spin and momentum fraction of the proton using

lattice QCD simulations at physical pion mass. *Phys. Rev. D*, 101(9):094513, 2020.

[285] R. M. Whitehill, Yiyu Zhou, N. Sato, and W. Melnitchouk. Accessing gluon polarization with high-PT hadrons in SIDIS. *Phys. Rev. D*, 107(3):034033, 2023.

[286] Kyle Lee, James Mulligan, Mateusz Płoskoń, Felix Ringer, and Feng Yuan. Machine learning-based jet and event classification at the Electron-Ion Collider with applications to hadron structure and spin physics. 10 2022.

[287] Tanjib Khan, Tianbo Liu, and Raza Sabbir Sufian. Gluon helicity distribution in the nucleon from lattice QCD and machine learning. 11 2022.

[288] Colin Egerer et al. Toward the determination of the gluon helicity distribution in the nucleon from lattice quantum chromodynamics. *Phys. Rev. D*, 106(9):094511, 2022.

[289] Martha Constantinou et al. Lattice QCD Calculations of Parton Physics. 2 2022.

[290] L. Adamczyk et al. Measurement of the cross section and longitudinal double-spin asymmetry for di-jet production in polarized pp collisions at $\sqrt{s} = 200$ GeV. *Phys. Rev. D*, 95(7):071103, 2017.

[291] Jaroslav Adam et al. Longitudinal double-spin asymmetries for dijet production at intermediate pseudorapidity in polarized pp collisions at $\sqrt{s} = 200$ GeV. *Phys. Rev. D*, 98(3):032011, 2018.

[292] A. Adare et al. Inclusive cross-section and double helicity asymmetry for π^0 production in $p + p$ collisions at $\sqrt{s} = 200$ GeV: Implications for the polarized gluon distribution in the proton. *Phys. Rev. D*, 76:051106, 2007.

[293] A. Adare et al. The Polarized gluon contribution to the proton spin from the double helicity asymmetry in inclusive pi0 production in polarized p + p collisions at s**(1/2) = 200-GeV. *Phys. Rev. Lett.*, 103:012003, 2009.

[294] A. Adare et al. Inclusive cross section and double helicity asymmetry for piˆ0 production in p^+p collisions at $\sqrt{s} = 62.4$ GeV. *Phys. Rev. D*, 79:012003, 2009.

[295] Rabah Abdul Khalek, Valerio Bertone, Alice Khoudli, and Emanuele R. Nocera. Pion and kaon fragmentation functions at next-to-next-to-leading order. *Phys. Lett. B*, 834:137456, 2022.

[296] John C. Collins, Steve F. Heppelmann, and Glenn A. Ladinsky. Measuring transversity densities in singly polarized hadron hadron and lepton - hadron collisions. *Nucl. Phys. B*, 420:565–582, 1994.

[297] Rajan Gupta, Boram Yoon, Tanmoy Bhattacharya, Vincenzo Cirigliano, Yong-Chull Jang, and Huey-Wen Lin. Flavor diagonal tensor charges of the nucleon from (2+1+1)-flavor lattice QCD. *Phys. Rev. D*, 98(9):091501, 2018.

[298] Zhong-Bo Kang, Alexei Prokudin, Peng Sun, and Feng Yuan. Extraction of Quark Transversity Distribution and Collins Fragmentation Functions with QCD Evolution. *Phys. Rev. D*, 93(1):014009, 2016.

[299] M. Anselmino, M. Boglione, U. D'Alesio, S. Melis, F. Murgia, and A. Prokudin. Simultaneous extraction of transversity and Collins functions from new SIDIS and e+e- data. *Phys. Rev. D*, 87:094019, 2013.

[300] Gary. R. Goldstein, J. Osvaldo Gonzalez Hernandez, and Simonetta Liuti. Flavor dependence of chiral odd generalized parton distributions and the tensor charge from the analysis of combined π^0 and η exclusive electroproduction data. 1 2014.

[301] J. Benel, A. Courtoy, and R. Ferro-Hernandez. A constrained fit of the valence transversity distributions from dihadron production. *Eur. Phys. J. C*, 80(5):465, 2020.

[302] Umberto D'Alesio, Carlo Flore, and Alexei Prokudin. Role of the Soffer bound in determination of transversity and the tensor charge. *Phys. Lett. B*, 803:135347, 2020.

[303] Nodoka Yamanaka, Shoji Hashimoto, Takashi Kaneko, and Hiroshi Ohki. Nucleon charges with dynamical overlap fermions. *Phys. Rev. D*, 98(5):054516, 2018.

[304] C. Alexandrou, S. Bacchio, M. Constantinou, J. Finkenrath, K. Hadjiyiannakou, K. Jansen, G. Koutsou, and A. Vaquero Aviles-Casco. Nucleon axial, tensor, and scalar charges and σ-terms in lattice QCD. *Phys. Rev. D*, 102(5):054517, 2020.

[305] Aurore Courtoy, Stefan Baeßler, Martín González-Alonso, and Simonetta Liuti. Beyond-Standard-Model Tensor Interaction and Hadron Phenomenology. *Phys. Rev. Lett.*, 115:162001, 2015.

[306] N. Yamanaka, B. K. Sahoo, N. Yoshinaga, T. Sato, K. Asahi, and B. P. Das. Probing exotic phenomena at the interface of nuclear and particle physics with the electric dipole moments of diamagnetic atoms: A unique window to hadronic and semi-leptonic CP violation. *Eur. Phys. J. A*, 53(3):54, 2017.

[307] Tianbo Liu, Zhiwen Zhao, and Haiyan Gao. Experimental constraint on quark electric dipole moments. *Phys. Rev. D*, 97(7):074018, 2018.

[308] Martin González-Alonso, Oscar Naviliat-Cuncic, and Nathal Severijns. New physics searches in nuclear and neutron β decay. *Prog. Part. Nucl. Phys.*, 104:165–223, 2019.

[309] A. Bianconi, S. Boffi, R. Jakob, and M. Radici. Two hadron interference fragmentation functions. Part 1. General framework. *Phys. Rev. D*, 62:034008, 2000.

[310] Federico A. Ceccopieri, Marco Radici, and Alessandro Bacchetta. Evolution equations for extended dihadron fragmentation functions. *Phys. Lett. B*, 650:81–89, 2007.

[311] K. Konishi, A. Ukawa, and G. Veneziano. Jet Calculus: A Simple Algorithm for Resolving QCD Jets. *Nucl. Phys. B*, 157:45–107, 1979.

[312] D. de Florian and L. Vanni. Two hadron production in e+ e- annihilation to next-to-leading order accuracy. *Phys. Lett. B*, 578:139–149, 2004.

[313] U. P. Sukhatme and K. E. Lassila. Q^2 Evolution of Multi - Hadron Fragmentation Functions. *Phys. Rev. D*, 22:1184, 1980.

[314] Xavier Artru and John C. Collins. Measuring transverse spin correlations by 4 particle correlations in e+ e- —> 2 jets. *Z. Phys. C*, 69:277–286, 1996.

[315] Daniel Boer, Rainer Jakob, and Marco Radici. Interference fragmentation functions in electron positron annihilation. *Phys. Rev. D*, 67:094003, 2003. [Erratum: Phys.Rev.D 98, 039902 (2018)].

[316] Hrayr H. Matevosyan, Aram Kotzinian, and Anthony W. Thomas. Dihadron fragmentation functions in the quark-jet model: Transversely polarized quarks. *Phys. Rev. D*, 97(1):014019, 2018.

[317] C. Adolph et al. Transverse spin effects in hadron-pair production from semi-inclusive deep inelastic scattering. *Phys. Lett. B*, 713:10–16, 2012.

[318] Transverse-spin-dependent azimuthal asymmetries of pion and kaon pairs produced in muon-proton and muon-deuteron semi-inclusive deep inelastic scattering. 1 2023.

[319] A. Airapetian et al. Evidence for a Transverse Single-Spin Asymmetry in Leptoproduction of pi+pi- Pairs. *JHEP*, 06:017, 2008.

[320] Alessandro Bacchetta and Marco Radici. Partial wave analysis of two hadron fragmentation functions. *Phys. Rev. D*, 67:094002, 2003.

[321] Alessandro Bacchetta and Marco Radici. Two hadron semiinclusive production including subleading twist. *Phys. Rev. D*, 69:074026, 2004.

[322] Alessandro Bacchetta and Marco Radici. Dihadron interference fragmentation functions in proton-proton collisions. *Phys. Rev. D*, 70:094032, 2004.

[323] Aurore Courtoy, Alessandro Bacchetta, Marco Radici, and Andrea Bianconi. First extraction of Interference Fragmentation Functions from e^+e^- data. *Phys. Rev. D*, 85:114023, 2012.

[324] P. V. Pobylitsa. Transverse momentum dependent parton distributions in large N(c) QCD. 1 2003.

[325] Miguel G. Echevarria, Ahmad Idilbi, Zhong-Bo Kang, and Ivan Vitev. QCD Evolution of the Sivers Asymmetry. *Phys. Rev. D*, 89:074013, 2014.

[326] Miguel G. Echevarria, Zhong-Bo Kang, and John Terry. Global analysis of the Sivers functions at NLO+NNLL in QCD. *JHEP*, 01:126, 2021.

[327] Alessandro Bacchetta, Filippo Delcarro, Cristian Pisano, and Marco Radici. The 3-dimensional distribution of quarks in momentum space. *Phys. Lett. B*, 827:136961, 2022.

[328] M. Anselmino, M. Boglione, U. D'Alesio, F. Murgia, and A. Prokudin. Study of the sign change of the Sivers function from STAR Collaboration W/Z production data. *JHEP*, 04:046, 2017.

[329] Yuri V. Kovchegov and Matthew D. Sievert. Small-x Helicity Evolution: an Operator Treatment. *Phys. Rev. D*, 99(5):054032, 2019.

[330] Yuri V. Kovchegov and Matthew D. Sievert. Valence Quark Transversity at Small x. *Phys. Rev. D*, 99(5):054033, 2019.

[331] Yuri V. Kovchegov and Daniel Pitonyak. Private communication, 2023.

[332] Jacques Soffer. Positivity constraints for spin dependent parton distributions. *Phys. Rev. Lett.*, 74:1292–1294, 1995.

[333] R. Seidl et al. Invariant-mass and fractional-energy dependence of inclusive production of di-hadrons in e^+e^- annihilation at $\sqrt{s} = 10.58$ GeV. *Phys. Rev. D*, 96(3):032005, 2017.

[334] A. Vossen et al. Observation of transverse polarization asymmetries of charged pion pairs in e^+e^- annihilation near $\sqrt{s} = 10.58$ GeV. *Phys. Rev. Lett.*, 107:072004, 2011.

[335] S. Brandt, C. Peyrou, R. Sosnowski, and A. Wroblewski. The Principal axis of jets. An Attempt to analyze high-energy collisions as two-body processes. *Phys. Lett.*, 12:57–61, 1964.

[336] 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.

[337] L. Adamczyk et al. Observation of Transverse Spin-Dependent Azimuthal Correlations of Charged Pion Pairs in $p^{\uparrow} + p$ at $\sqrt{s} = 200$ GeV. *Phys. Rev. Lett.*, 115:242501, 2015.

[338] B. Pokhrel and B. Surrow. Private communication, 2023.

[339] L. Adamczyk et al. Transverse spin-dependent azimuthal correlations of charged pion pairs measured in $p^{\uparrow}$+p collisions at $\sqrt{s} = 500$ GeV. *Phys. Lett. B*, 780:332–339, 2018.

[340] Mario Pitschmann, Chien-Yeah Seng, Craig D. Roberts, and Sebastian M. Schmidt. Nucleon tensor charges and electric dipole moments. *Phys. Rev. D*, 91:074004, 2015.

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

[342] 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.

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

[344] 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.

[345] V. Braun, P. Gornicki, and L. Mankiewicz. Ioffe - time distributions instead of parton momentum distributions in description of deep inelastic scattering. *Phys. Rev. D*, 51:6036–6051, 1995.

[346] William Detmold and C. J. David Lin. Deep-inelastic scattering and the operator product expansion in lattice QCD. *Phys. Rev. D*, 73:014501, 2006.

[347] V. Braun and Dieter Müller. Exclusive processes in position space and the pion distribution amplitude. *Eur. Phys. J. C*, 55:349–361, 2008.

[348] A. J. Chambers, R. Horsley, Y. Nakamura, H. Perlt, P. E. L. Rakow, G. Schierholz, A. Schiller, K. Somfleth, R. D. Young, and J. M. Zanotti. Nucleon Structure Functions from Operator Product Expansion on the Lattice. *Phys. Rev. Lett.*, 118(24):242001, 2017.

[349] Maxwell T. Hansen, Harvey B. Meyer, and Daniel Robaina. From deep inelastic scattering to heavy-flavor semileptonic decays: Total rates into multihadron final states from lattice QCD. *Phys. Rev. D*, 96(9):094513, 2017.

[350] A. V. Radyushkin. Quasi-parton distribution functions, momentum distributions, and pseudo-parton distribution functions. *Phys. Rev. D*, 96(3):034025, 2017.

[351] Yan-Qing Ma and Jian-Wei Qiu. Exploring Partonic Structure of Hadrons Using ab initio Lattice QCD Calculations. *Phys. Rev. Lett.*, 120(2):022003, 2018.

[352] A. V. Radyushkin. Quark pseudodistributions at short distances. *Phys. Lett. B*, 781:433–442, 2018.

[353] Jian Liang, Keh-Fei Liu, and Yi-Bo Yang. Lattice calculation of hadronic tensor of the nucleon. *EPJ Web Conf.*, 175:14014, 2018.

[354] William Detmold, Issaku Kanamori, C. J. David Lin, Santanu Mondal, and Yong Zhao. Moments of pion distribution amplitude using operator product expansion on the lattice. *PoS*, LATTICE2018:106, 2018.

[355] Xiangdong Ji and Jian-Hui Zhang. Renormalization of quasiparton distribution. *Phys. Rev. D*, 92:034006, 2015.

[356] Tomomi Ishikawa, Yan-Qing Ma, Jian-Wei Qiu, and Shinsuke Yoshida. Practical quasi parton distribution functions. 9 2016.

[357] Jiunn-Wei Chen, Xiangdong Ji, and Jian-Hui Zhang. Improved quasi parton distribution through Wilson line renormalization. *Nucl. Phys. B*, 915:1–9, 2017.

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

[359] 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.

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

[361] Tomomi Ishikawa, Yan-Qing Ma, Jian-Wei Qiu, and Shinsuke Yoshida. Renormalizability of quasiparton distribution functions. *Phys. Rev. D*, 96(9):094019, 2017.

[362] 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.

[363] Gregoris Spanoudes and Haralambos Panagopoulos. Renormalization of Wilson-line operators in the presence of nonzero quark masses. *Phys. Rev. D*, 98(1):014509, 2018.

[364] 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.

[365] Zheng-Yang Li, Yan-Qing Ma, and Jian-Wei Qiu. Multiplicative Renormalizability of Operators defining Quasiparton Distributions. *Phys. Rev. Lett.*, 122(6):062002, 2019.

[366] 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.

[367] 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.

[368] Hsiang-nan Li. Nondipolar Wilson links for quasiparton distribution functions. *Phys. Rev. D*, 94(7):074036, 2016.

[369] Chirstopher Monahan and Kostas Orginos. Quasi parton distributions and the gradient flow. *JHEP*, 03:116, 2017.

[370] Anatoly Radyushkin. Nonperturbative Evolution of Parton Quasi-Distributions. *Phys. Lett. B*, 767:314–320, 2017.

[371] Anatoly Radyushkin. Target Mass Effects in Parton Quasi-Distributions. *Phys. Lett. B*, 770:514–522, 2017.

[372] Carl E. Carlson and Michael Freid. Lattice corrections to the quark quasidistribution at one-loop. *Phys. Rev. D*, 95(9):094504, 2017.

[373] Raúl A. Briceño, Maxwell T. Hansen, and Christopher J. Monahan. Role of the Euclidean signature in lattice calculations of quasidistributions and other nonlocal matrix elements. *Phys. Rev. D*, 96(1):014502, 2017.

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

[375] G. C. Rossi and M. Testa. Note on lattice regularization and equal-time correlators for parton distribution functions. *Phys. Rev. D*, 96(1):014507, 2017.

[376] 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.

[377] Wei Wang, Shuai Zhao, and Ruilin Zhu. Gluon quasidistribution function at one loop. *Eur. Phys. J. C*, 78(2):147, 2018.

[378] 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.

[379] Christopher Monahan. Smeared quasidistributions in perturbation theory. *Phys. Rev. D*, 97(5):054507, 2018.

[380] Anatoly Radyushkin. One-loop evolution of parton pseudo-distribution functions on the lattice. *Phys. Rev. D*, 98(1):014019, 2018.

[381] Jian-Hui Zhang, Jiunn-Wei Chen, and Christopher Monahan. Parton distribution functions from reduced Ioffe-time distributions. *Phys. Rev. D*, 97(7):074508, 2018.

[382] 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.

[383] Ji Xu, Qi-An Zhang, and Shuai Zhao. Light-cone distribution amplitudes of vector meson in a large momentum effective theory. *Phys. Rev. D*, 97(11):114026, 2018.

[384] Yu Jia, Shuangran Liang, Xiaonu Xiong, and Rui Yu. Partonic quasidistributions in two-dimensional QCD. *Phys. Rev. D*, 98(5):054011, 2018.

[385] Raúl A. Briceño, Juan V. Guerrero, Maxwell T. Hansen, and Christopher J. Monahan. Finite-volume effects due to spatially nonlocal operators. *Phys. Rev. D*, 98(1):014511, 2018.

[386] Giancarlo Rossi and Massimo Testa. Euclidean versus Minkowski short distance. *Phys. Rev. D*, 98(5):054028, 2018.

[387] A. V. Radyushkin. Structure of parton quasi-distributions and their moments. *Phys. Lett. B*, 788:380–387, 2019.

[388] Xiangdong Ji, Yizhuang Liu, and Ismail Zahed. Quasiparton distribution functions: Two-dimensional scalar and spinor QCD. *Phys. Rev. D*, 99(5):054008, 2019.

[389] Joseph Karpie, Kostas Orginos, and Savvas Zafeiropoulos. Moments of Ioffe time parton distribution functions from non-local matrix elements. *JHEP*, 11:178, 2018.

[390] Vladimir M. Braun, Alexey Vladimirov, and Jian-Hui Zhang. Power corrections and renormalons in parton quasidistributions. *Phys. Rev. D*, 99(1):014013, 2019.

[391] 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.

[392] 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.

[393] Raúl A. Briceño, Juan V. Guerrero, Maxwell T. Hansen, and Christopher J. Monahan. PDFs in small boxes. *PoS*, LATTICE2018:111, 2018.

[394] Markus A. Ebert, Iain W. Stewart, and Yong Zhao. Towards Quasi-Transverse Momentum Dependent PDFs Computable on the Lattice. *JHEP*, 09:037, 2019.

[395] 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.

[396] Constantia Alexandrou, Krzysztof Cichy, Vincent Drach, Elena Garcia-Ramos, Kyriakos Hadjiyiannakou, Karl Jansen, Fernanda Steffens, and Christian Wiese. Lattice calculation of parton distributions. *Phys. Rev. D*, 92:014502, 2015.

[397] 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.

[398] Constantia Alexandrou, Krzysztof Cichy, Martha Constantinou, Kyriakos Hadjiyiannakou, Karl Jansen, Fernanda Steffens, and Christian Wiese. Updated Lattice Results for Parton Distributions. *Phys. Rev. D*, 96(1):014513, 2017.

[399] 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.

[400] 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.

[401] Gunnar S. Bali et al. Pion distribution amplitude from Euclidean correlation functions. *Eur. Phys. J. C*, 78(3):217, 2018.

[402] 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.

[403] 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.

[404] 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.

[405] 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.

[406] 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.

[407] 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.

[408] 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.

[409] Gunnar S. Bali, Vladimir M. Braun, Benjamin Gläßle, Meinulf Göckeler, Michael Gruber, Fabian Hutzler, Piotr Korcyl, Andreas Schäfer, Philipp Wein, and Jian-Hui Zhang. Pion distribution amplitude from Euclidean correlation functions: Exploring universality and higher-twist effects. *Phys. Rev. D*, 98(9):094507, 2018.

[410] 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.

[411] 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.

[412] 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.

[413] Gunnar S. Bali, Vladimir M. Braun, Meinulf Göckeler, Michael Gruber, Fabian Hutzler, Piotr Korcyl, Andreas Schäfer, and Philipp Wein. Pion and Kaon Distribution Amplitudes from lattice QCD: towards the continuum limit. *PoS*, LATTICE2018:107, 2018.

[414] Charles Shugert, Taku Izubichi, Luchang Jin, Christos Kallidonis, Nikhil Karthik, Swagato Mukherjee, Peter Petreczky, and Sergey Syritsyn. Pion quasi parton distribution function on a fine lattice. *PoS*, LATTICE2018:110, 2018.

[415] Raza Sabbir Sufian, Joseph Karpie, Colin Egerer, Kostas Orginos, Jian-Wei Qiu, and David G. Richards. Pion Valence Quark Distribution from Matrix Element Calculated in Lattice QCD. *Phys. Rev. D*, 99(7):074507, 2019.

[416] 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.

[417] 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.

[418] Bálint Joó, Joseph Karpie, Kostas Orginos, Anatoly Radyushkin, David Richards, and Savvas Zafeiropoulos. Parton Distribution Functions from Ioffe time pseudo-distributions. *JHEP*, 12:081, 2019.

[419] Bálint Joó, Joseph Karpie, Kostas Orginos, Anatoly V. Radyushkin, David G. Richards, Raza Sabbir Sufian, and Savvas Zafeiropoulos. Pion valence structure from Ioffe-time parton pseudodistribution functions. *Phys. Rev. D*, 100(11):114512, 2019.

[420] Yahui Chai et al. Parton distribution functions of Δ^+ on the lattice. *Phys. Rev. D*, 102(1):014508, 2020.

[421] Bálint Joó, Joseph Karpie, Kostas Orginos, Anatoly V. Radyushkin, David G. Richards, and Savvas Zafeiropoulos. Parton Distribution Functions from Ioffe Time Pseudodistributions from Lattice Calculations: Approaching the Physical Point. *Phys. Rev. Lett.*, 125(23):232003, 2020.

[422] 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.

[423] 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.

[424] 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.

[425] 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.

[426] 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.

[427] Huey-Wen Lin, Jiunn-Wei Chen, and Rui Zhang. Lattice Nucleon Isovector Unpolarized Parton Distribution in the Physical-Continuum Limit. 11 2020.

[428] Joseph Karpie, Kostas Orginos, Anatoly Radyushkin, and Savvas Zafeiropoulos. The continuum and leading twist limits of parton distribution functions in lattice QCD. *JHEP*, 11:024, 2021.

[429] Colin Egerer, Robert G. Edwards, Christos Kallidonis, Kostas Orginos, Anatoly V. Radyushkin, David G. Richards, Eloy Romero, and Savvas Zafeiropoulos. Towards high-precision parton distributions from lattice QCD via distillation. *JHEP*, 11:148, 2021.

[430] Colin Egerer et al. Transversity parton distribution function of the nucleon using the pseudodistribution approach. *Phys. Rev. D*, 105(3):034507, 2022.

[431] 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.

[432] 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.

[433] Fei Yao et al. Nucleon Transversity Distribution in the Continuum and Physical Mass Limit from Lattice QCD. 8 2022.

[434] Shohini Bhattacharya, Krzysztof Cichy, Martha Constantinou, Andreas Metz, Aurora Scapellato, and Fernanda Steffens. One-loop matching for the twist-3 parton distribution $g_T(x)$. *Phys. Rev. D*, 102(3):034005, 2020.

[435] 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.

[436] Shohini Bhattacharya and Andreas Metz. Burkhardt-Cottingham-type sum rules for light-cone and quasi-PDFs. *Phys. Rev. D*, 105(5):054027, 2022.

[437] 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.

[438] 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.

[439] 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.

[440] William Detmold, Anthony V. Grebe, Issaku Kanamori, C. J. David Lin, Santanu Mondal, Robert J. Perry, and Yong Zhao. Parton physics from a heavy-quark operator product expansion: Lattice QCD calculation of the second moment of the pion distribution amplitude. *Phys. Rev. D*, 105(3):034506, 2022.

[441] Jun Hua et al. Pion and Kaon Distribution Amplitudes from Lattice QCD. *Phys. Rev. Lett.*, 129(13):132001, 2022.

[442] Xiang Gao, Andrew D. Hanlon, Nikhil Karthik, Swagato Mukherjee, Peter Petreczky, Philipp Scior, Sergey Syritsyn, and Yong Zhao. Pion distribution amplitude at the physical point using the leading-twist expansion of the quasi-distribution-amplitude matrix element. *Phys. Rev. D*, 106(7):074505, 2022.

[443] Phiala Shanahan, Michael Wagman, and Yong Zhao. Collins-Soper kernel for TMD evolution from lattice QCD. *Phys. Rev. D*, 102(1):014511, 2020.

[444] Qi-An Zhang et al. Lattice-QCD Calculations of TMD Soft Function Through Large-Momentum Effective Theory. *Phys. Rev. Lett.*, 125(19):192001, 2020.

[445] Yuan Li et al. Lattice QCD Study of Transverse-Momentum Dependent Soft Function. *Phys. Rev. Lett.*, 128(6):062002, 2022.

[446] 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.

[447] Maximilian Schlemmer, Alexey Vladimirov, Christian Zimmermann, Michael Engelhardt, and Andreas Schäfer. Determination of the Collins-Soper Kernel from Lattice QCD. *JHEP*, 08:004, 2021.

[448] 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.

[449] Leonard Gamberg, Zhong-Bo Kang, Ivan Vitev, and Hongxi Xing. Quasi-parton distribution functions: a study in the diquark spectator model. *Phys. Lett. B*, 743:112–120, 2015.

[450] Alessandro Bacchetta, Marco Radici, Barbara Pasquini, and Xiaonu Xiong. Reconstructing parton densities at large fractional momenta. *Phys. Rev. D*, 95(1):014036, 2017.

[451] Seung-il Nam. Quasi-distribution amplitudes for pion and kaon via the nonlocal chiral-quark model. *Mod. Phys. Lett. A*, 32(39):1750218, 2017.

[452] Wojciech Broniowski and Enrique Ruiz Arriola. Nonperturbative partonic quasidistributions of the pion from chiral quark models. *Phys. Lett. B*, 773:385–390, 2017.

[453] T. J. Hobbs. Quantifying finite-momentum effects in the quark quasidistribution functions of mesons. *Phys. Rev. D*, 97(5):054028, 2018.

[454] Wojciech Broniowski and Enrique Ruiz Arriola. Partonic quasidistributions of the proton and pion from transverse-momentum distributions. *Phys. Rev. D*, 97(3):034031, 2018.

[455] Shu-Sheng Xu, Lei Chang, Craig D. Roberts, and Hong-Shi Zong. Pion and kaon valence-quark parton quasidistributions. *Phys. Rev. D*, 97(9):094014, 2018.

[456] Krzysztof Cichy and Martha Constantinou. A guide to light-cone PDFs from Lattice QCD: an overview of approaches, techniques and results. *Adv. High Energy Phys.*, 2019:3036904, 2019.

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

[458] Martha Constantinou. The x-dependence of hadronic parton distributions: A review on the progress of lattice QCD. *Eur. Phys. J. A*, 57(2):77, 2021.

[459] Krzysztof Cichy. Progress in x-dependent partonic distributions from lattice QCD. *PoS*, LATTICE2021:017, 2022.

[460] Krzysztof Cichy. Overview of lattice calculations of the x-dependence of PDFs, GPDs and TMDs. *EPJ Web Conf.*, 258:01005, 2022.

[461] Dieter Müller, D. Robaschik, B. Geyer, F. M. Dittes, and J. Hořejši. Wave functions, evolution equations and evolution kernels from light ray operators of QCD. *Fortsch. Phys.*, 42:101–141, 1994.

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

[463] A. V. Radyushkin. Scaling limit of deeply virtual Compton scattering. *Phys. Lett. B*, 380:417–425, 1996.

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

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

[466] K. Goeke, Maxim V. Polyakov, and M. Vanderhaeghen. Hard exclusive reactions and the structure of hadrons. *Prog. Part. Nucl. Phys.*, 47:401–515, 2001.

[467] Sigfrido Boffi and Barbara Pasquini. Generalized parton distributions and the structure of the nucleon. *Riv. Nuovo Cim.*, 30(9):387–448, 2007.

[468] Michel Guidal, Hervé Moutarde, and Marc Vanderhaeghen. Generalized Parton Distributions in the valence region from Deeply Virtual Compton Scattering. *Rept. Prog. Phys.*, 76:066202, 2013.

[469] Dieter Mueller. Generalized Parton Distributions – visions, basics, and realities –. *Few Body Syst.*, 55:317–337, 2014.

[470] Matthias Burkardt. Impact parameter dependent parton distributions and off forward parton distributions for zeta —> 0. *Phys. Rev. D*, 62:071503, 2000. [Erratum: Phys.Rev.D 66, 119903 (2002)].

[471] John P. Ralston and Bernard Pire. Femtophotography of protons to nuclei with deeply virtual Compton scattering. *Phys. Rev. D*, 66:111501, 2002.

[472] M. Diehl. Generalized parton distributions in impact parameter space. *Eur. Phys. J. C*, 25:223–232, 2002. [Erratum: Eur.Phys.J.C 31, 277–278 (2003)].

[473] Matthias Burkardt. Impact parameter space interpretation for generalized parton distributions. *Int. J. Mod. Phys. A*, 18:173–208, 2003.

[474] M. V. Polyakov and A. G. Shuvaev. On'dual' parametrizations of generalized parton distributions. 7 2002.

[475] M. V. Polyakov. Generalized parton distributions and strong forces inside nucleons and nuclei. *Phys. Lett. B*, 555:57–62, 2003.

[476] Maxim V. Polyakov and Peter Schweitzer. Forces inside hadrons: pressure, surface tension, mechanical radius, and all that. *Int. J. Mod. Phys. A*, 33(26):1830025, 2018.

[477] John C. Collins and Andreas Freund. Proof of factorization for deeply virtual Compton scattering in QCD. *Phys. Rev. D*, 59:074009, 1999.

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

[479] L. Mankiewicz, G. Piller, and T. Weigl. Hard exclusive meson production and nonforward parton distributions. *Eur. Phys. J. C*, 5:119–128, 1998.

[480] V. Bertone, H. Dutrieux, C. Mezrag, H. Moutarde, and P. Sznajder. Deconvolution problem of deeply virtual Compton scattering. *Phys. Rev. D*, 103(11):114019, 2021.

[481] Philipp Hagler, John W. Negele, Dru Bryant Renner, W. Schroers, T. Lippert, and K. Schilling. Moments of nucleon generalized parton distributions in lattice QCD. *Phys. Rev. D*, 68:034505, 2003.

[482] Dirk Brommel et al. Moments of generalized parton distributions and quark angular momentum of the nucleon. *PoS*, LATTICE2007:158, 2007.

[483] C. Alexandrou, J. Carbonell, M. Constantinou, P. A. Harraud, P. Guichon, K. Jansen, C. Kallidonis, T. Korzec, and M. Papinutto. Moments of nucleon generalized parton distributions from lattice QCD. *Phys. Rev. D*, 83:114513, 2011.

[484] C. Alexandrou, M. Constantinou, S. Dinter, V. Drach, K. Jansen, C. Kallidonis, and G. Koutsou. Nucleon form factors and moments of generalized parton distributions using $N_f = 2 + 1 + 1$ twisted mass fermions. *Phys. Rev. D*, 88(1):014509, 2013.

[485] Martha Constantinou. Hadron Structure. *PoS*, LATTICE2014:001, 2015.

[486] J. R. Green, J. W. Negele, A. V. Pochinsky, S. N. Syritsyn, M. Engelhardt, and S. Krieg. Nucleon electromagnetic form factors from lattice QCD using a nearly physical pion mass. *Phys. Rev. D*, 90:074507, 2014.

[487] Constantia Alexandrou, Martha Constantinou, Kyriakos Hadjiyiannakou, Karl Jansen, Christos Kallidonis, Giannis Koutsou, and Alejandro Vaquero Aviles-Casco. Nucleon electromagnetic form factors using lattice simulations at the physical point. *Phys. Rev. D*, 96(3):034503, 2017.

[488] Constantia Alexandrou, Martha Constantinou, Kyriakos Hadjiyiannakou, Karl Jansen, Christos Kallidonis, Giannis Koutsou, and Alejandro Vaquero Aviles-Casco. Nucleon axial form factors using $N_f = 2$ twisted mass fermions with a physical value of the pion mass. *Phys. Rev. D*, 96(5):054507, 2017.

[489] Nesreen Hasan, Jeremy Green, Stefan Meinel, Michael Engelhardt, Stefan Krieg, John Negele, Andrew Pochinsky, and Sergey Syritsyn. Computing the nucleon charge and axial radii directly at $Q^2 = 0$ in lattice QCD. *Phys. Rev. D*, 97(3):034504, 2018.

[490] Rajan Gupta, Yong-Chull Jang, Huey-Wen Lin, Boram Yoon, and Tanmoy Bhattacharya. Axial Vector Form Factors of the Nucleon from Lattice QCD. *Phys. Rev. D*, 96(11):114503, 2017.

[491] Stefano Capitani, Michele Della Morte, Dalibor Djukanovic, Georg M. von Hippel, Jiayu Hua, Benjamin Jäger, Parikshit M. Junnarkar, Harvey B. Meyer, Thomas D. Rae, and Hartmut Wittig. Isovector axial form factors of the nucleon in two-flavor lattice QCD. *Int. J. Mod. Phys. A*, 34(02):1950009, 2019.

[492] C. Alexandrou, S. Bacchio, M. Constantinou, J. Finkenrath, K. Hadjiyiannakou, K. Jansen, G. Koutsou, and A. Vaquero Aviles-Casco. Proton and neutron electromagnetic form factors from lattice QCD. *Phys. Rev. D*, 100(1):014509, 2019.

[493] Eigo Shintani, Ken-Ichi Ishikawa, Yoshinobu Kuramashi, Shoichi Sasaki, and Takeshi Yamazaki. Nucleon form factors and root-mean-square radii on a (10.8 fm)^4 lattice at the physical point. *Phys. Rev. D*, 99(1):014510, 2019. [Erratum: Phys.Rev.D 102, 019902 (2020)].

[494] Yong-Chull Jang, Tanmoy Bhattacharya, Rajan Gupta, Huey-Wen Lin, and Boram Yoon. Updates on Nucleon Form Factors from Clover-on-HISQ Lattice Formulation. *PoS*, LATTICE2018:123, 2018.

[495] G. S. Bali, S. Collins, M. Gruber, A. Schäfer, P. Wein, and T. Wurm. Solving the PCAC puzzle for nucleon axial and pseudoscalar form factors. *Phys. Lett. B*, 789:666–674, 2019.

[496] Gunnar S. Bali, Sara Collins, Meinulf Göckeler, Rudolf Rödl, Andreas Schäfer, and André Sternbeck. Nucleon generalized form factors from two-flavor lattice QCD. *Phys. Rev. D*, 100(1):014507, 2019.

[497] C. Alexandrou et al. Moments of nucleon generalized parton distributions from lattice QCD simulations at physical pion mass. *Phys. Rev. D*, 101(3):034519, 2020.

[498] Martha Constantinou et al. Parton distributions and lattice-QCD calculations: Toward 3D structure. *Prog. Part. Nucl. Phys.*, 121:103908, 2021.

[499] C. Alexandrou et al. First moments of the nucleon transverse quark spin densities using lattice QCD. 2 2022.

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

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

[502] 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.

[503] Anatoly V. Radyushkin. Generalized parton distributions and pseudodistributions. *Phys. Rev. D*, 100(11):116011, 2019.

[504] J. P. Ma, Z. Y. Pang, and G. P. Zhang. QCD factorization of quasi generalized quark distributions. *JHEP*, 08:130, 2022.

[505] J. P. Ma, Z. Y. Pang, C. P. Zhang, and G. P. Zhang. QCD Factorization of Quasi Generalized Gluon Distributions. 12 2022.

[506] Yao Ji, Fei Yao, and Jian-Hui Zhang. Connecting Euclidean to light-cone correlations: From flavor nonsinglet in forward kinematics to flavor singlet in nonforward kinematics. 12 2022.

[507] Zhi-Lei Ma, Jia-Qing Zhu, and Zhun Lu. Quasiparton distribution function and quasigeneralized parton distribution of the pion in a spectator model. *Phys. Rev. D*, 101(11):114005, 2020.

[508] Xuan Luo and Hao Sun. T-odd generalized and quasi transverse momentum dependent parton distribution in a scalar spectator model. *Eur. Phys. J. C*, 80(9):828, 2020.

[509] Jiunn-Wei Chen, Huey-Wen Lin, and Jian-Hui Zhang. Pion generalized parton distribution from lattice QCD. *Nucl. Phys. B*, 952:114940, 2020.

[510] 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.

[511] Huey-Wen Lin. Nucleon Tomography and Generalized Parton Distribution at Physical Pion Mass from Lattice QCD. *Phys. Rev. Lett.*, 127(18):182001, 2021.

[512] 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.

[513] Alec Hannaford-Gunn, Kadir Utku Can, Roger Horsley, Yoshifumi Nakamura, Holger Perlt, Paul E. L. Rakow, Hinnerk Stüben, Gerrit Schierholz, Ross D. Young, and James M. Zanotti. Generalized parton distributions from the off-forward Compton amplitude in lattice QCD. *Phys. Rev. D*, 105(1):014502, 2022.

[514] 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.

[515] Huey-Wen Lin. Nucleon helicity generalized parton distribution at physical pion mass from lattice QCD. *Phys. Lett. B*, 824:136821, 2022.

[516] 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.

[517] S. Meissner, A. Metz, and K. Goeke. Relations between generalized and transverse momentum dependent parton distributions. *Phys. Rev. D*, 76:034002, 2007.

[518] R. Jakob, P. J. Mulders, and J. Rodrigues. Modeling quark distribution and fragmentation functions. *Nucl. Phys. A*, 626:937–965, 1997.

[519] Stanley J. Brodsky, Dae Sung Hwang, and Ivan Schmidt. Final state interactions and single spin asymmetries in semiinclusive deep inelastic scattering. *Phys. Lett. B*, 530:99–107, 2002.

[520] Leonard P. Gamberg, Gary R. Goldstein, and Marc Schlegel. Transverse Quark Spin Effects and the Flavor Dependence of the Boer-Mulders Function. *Phys. Rev. D*, 77:094016, 2008.

[521] Alessandro Bacchetta, Francesco Conti, and Marco Radici. Transverse-momentum distributions in a diquark spectator model. *Phys. Rev. D*, 78:074010, 2008.

[522] Fatma Aslan, Matthias Burkardt, Cédric Lorcé, Andreas Metz, and Barbara Pasquini. Twist-3 generalized parton distributions in deeply-virtual Compton scattering. *Phys. Rev. D*, 98(1):014038, 2018.

[523] Fatma Aslan and Matthias Burkardt. Singularities in Twist-3 Quark Distributions. *Phys. Rev. D*, 101(1):016010, 2020.

[524] S. Chekanov et al. A ZEUS next-to-leading-order QCD analysis of data on deep inelastic scattering. *Phys. Rev. D*, 67:012007, 2003.

[525] M. Gluck, E. Reya, M. Stratmann, and W. Vogelsang. Models for the polarized parton distributions of the nucleon. *Phys. Rev. D*, 63:094005, 2001.

[526] Cédric Mezrag, Jorge Segovia, Lei Chang, and Craig D. Roberts. Parton distribution amplitudes: Revealing correlations within the proton and Roper. *Phys. Lett. B*, 783:263–267, 2018.

[527] Narinder Kumar, Chandan Mondal, and Neetika Sharma. Gravitational form factors and angular momentum densities in light-front quark-diquark model. *Eur. Phys. J. A*, 53(12):237, 2017.

[528] Kyle D. Bednar, Ian C. Cloët, and Peter C. Tandy. Nucleon quark distribution functions from the Dyson–Schwinger equations. *Phys. Lett. B*, 782:675–681, 2018.

[529] Stanley J. Brodsky, Dae Sung Hwang, Bo-Qiang Ma, and Ivan Schmidt. Light cone representation of the spin and orbital angular momentum of relativistic composite systems. *Nucl. Phys. B*, 593:311–335, 2001.

[530] Alessandro Bacchetta, Andreas Schaefer, and Jian-Jun Yang. Sivers function in a spectator model with axial vector diquarks. *Phys. Lett. B*, 578:109–118, 2004.

[531] Stephen L. Adler, E. W. Colglazier, Jr., J. B. Healy, Inga Karliner, Judy Lieberman, Yee Jack Ng, and Hung-Sheng Tsao. Renormalization Constants for Scalar, Pseudoscalar, and Tensor Currents. *Phys. Rev. D*, 11:3309, 1975.

[532] Cédric Lorcé, Hervé Moutarde, and Arkadiusz P. Trawiński. Revisiting the mechanical properties of the nucleon. *Eur. Phys. J. C*, 79(1):89, 2019.

[533] J. Ashman et al. A Measurement of the Spin Asymmetry and Determination of the Structure Function g(1) in Deep Inelastic Muon-Proton Scattering. *Phys. Lett. B*, 206:364, 1988.

[534] Elke C. Aschenauer, Ignacio Borsa, Rodolfo Sassot, and Charlotte Van Hulse. Semi-inclusive Deep-Inelastic Scattering, Parton Distributions and Fragmentation Functions at a Future Electron-Ion Collider. *Phys. Rev. D*, 99(9):094004, 2019.

[535] T. Stavreva, F. I. Olness, I. Schienbein, T. Jezo, A. Kusina, K. Kovarik, and J. Y. Yu. Heavy Quark Production in the ACOT Scheme at NNLO and N3LO. *Phys. Rev. D*, 85:114014, 2012.

[536] R. S. Thorne. A Variable-flavor number scheme for NNLO. *Phys. Rev. D*, 73:054019, 2006.

[537] G. Onengut et al. Measurement of nucleon structure functions in neutrino scattering. *Phys. Lett. B*, 632:65–75, 2006.

[538] M. Goncharov et al. Precise Measurement of Dimuon Production Cross-Sections in ν_μ Fe and $\bar{\nu}_\mu$ Fe Deep Inelastic Scattering at the Tevatron. *Phys. Rev. D*, 64:112006, 2001.

[539] A. Accardi, F. Arleo, W. K. Brooks, David D'Enterria, and V. Muccifora. Parton Propagation and Fragmentation in QCD Matter. *Riv. Nuovo Cim.*, 32(9-10):439–554, 2009.

[540] A. Majumder and M. Van Leeuwen. The Theory and Phenomenology of Perturbative QCD Based Jet Quenching. *Prog. Part. Nucl. Phys.*, 66:41–92, 2011.

[541] M. Dittmar et al. Working Group I: Parton distributions: Summary report for the HERA LHC Workshop Proceedings. 11 2005.

[542] K. Hafidi et al. Studies of partonic distributions using semi-inclusive production of Kaons.

[543] 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.

[544] Elke C. Aschenauer, R. Sassot, and M. Stratmann. Unveiling the Proton Spin Decomposition at a Future Electron-Ion Collider. *Phys. Rev. D*, 92(9):094030, 2015.

[545] Ignacio Borsa, Gonzalo Lucero, Rodolfo Sassot, Elke C. Aschenauer, and Ana S. Nunes. Revisiting helicity parton distributions at a future electron-ion collider. *Phys. Rev. D*, 102(9):094018, 2020.

[546] Mauro Anselmino, Paolo Gambino, and Jan Kalinowski. Polarized deep inelastic scattering at high-energies and parity violating structure functions. *Z. Phys. C*, 64:267–274, 1994.

[547] T. Hobbs and W. Melnitchouk. Finite-Q**2 corrections to parity-violating DIS. *Phys. Rev. D*, 77:114023, 2008.

[548] Y. X. Zhao, A. Deshpande, J. Huang, K. S. Kumar, and S. Riordan. Neutral-Current Weak Interactions at an EIC. *Eur. Phys. J. A*, 53(3):55, 2017.

[549] Radja Boughezal, Frank Petriello, and Hongxi Xing. Inclusive jet production as a probe of polarized parton distribution functions at a future EIC. *Phys. Rev. D*, 98(5):054031, 2018.

[550] Ignacio Borsa, Daniel de Florian, and Iván Pedron. Inclusive-jet and dijet production in polarized deep inelastic scattering. *Phys. Rev. D*, 103(1):014008, 2021.

[551] Felix Hekhorn and Marco Stratmann. Differential heavy quark distributions and correlations in longitudinally polarized deep-inelastic scattering. *Phys. Rev. D*, 104(1):016033, 2021.

[552] W. Cosyn and C. Weiss. Polarized electron-deuteron deep-inelastic scattering with spectator nucleon tagging. *Phys. Rev. C*, 102:065204, 2020.

[553] Daniel de Florian, Patriz Hinderer, Asmita Mukherjee, Felix Ringer, and Werner Vogelsang. Approximate next-to-next-to-leading order corrections to hadronic jet production. *Phys. Rev. Lett.*, 112:082001, 2014.

[554] Radja Boughezal, Hai Tao Li, and Frank Petriello. W-boson production in polarized proton-proton collisions at RHIC through next-to-next-to-leading order in perturbative QCD. *Phys. Lett. B*, 817:136333, 2021.

[555] Daniel Adamiak, Yuri V. Kovchegov, W. Melnitchouk, Daniel Pitonyak, Nobuo Sato, and Matthew D. Sievert. First analysis of world polarized DIS data with small-x helicity evolution. *Phys. Rev. D*, 104(3):L031501, 2021.

[556] Christopher Braun. *Hadron-pair production on transversely polarized targets in semi-inclusive deep inelastic scattering*. PhD thesis, Erlangen - Nuremberg U., 2014.

[557] B. Pokhrel and B. Surrow. Private communication, 2023.

[558] Leonard Gamberg, Zhong-Bo Kang, Daniel Pitonyak, Alexei Prokudin, Nobuo Sato, and Ralf Seidl. Electron-Ion Collider impact study on the tensor charge of the nucleon. *Phys. Lett. B*, 816:136255, 2021.

[559] Shohini Bhattacharya, Zhong-Bo Kang, Andreas Metz, Gregory Penn, and Daniel Pitonyak. First global QCD analysis of the TMD g1T from semi-inclusive DIS data. *Phys. Rev. D*, 105(3):034007, 2022.

APPENDIX A

COMPARISON TO PYTHIA-GENERATED DATA

Here we show the PYTHIA generated data used in the analysis of the DiFF D_1 (see Subsection 5.4.1). The plots below show the results for the five different energies used.

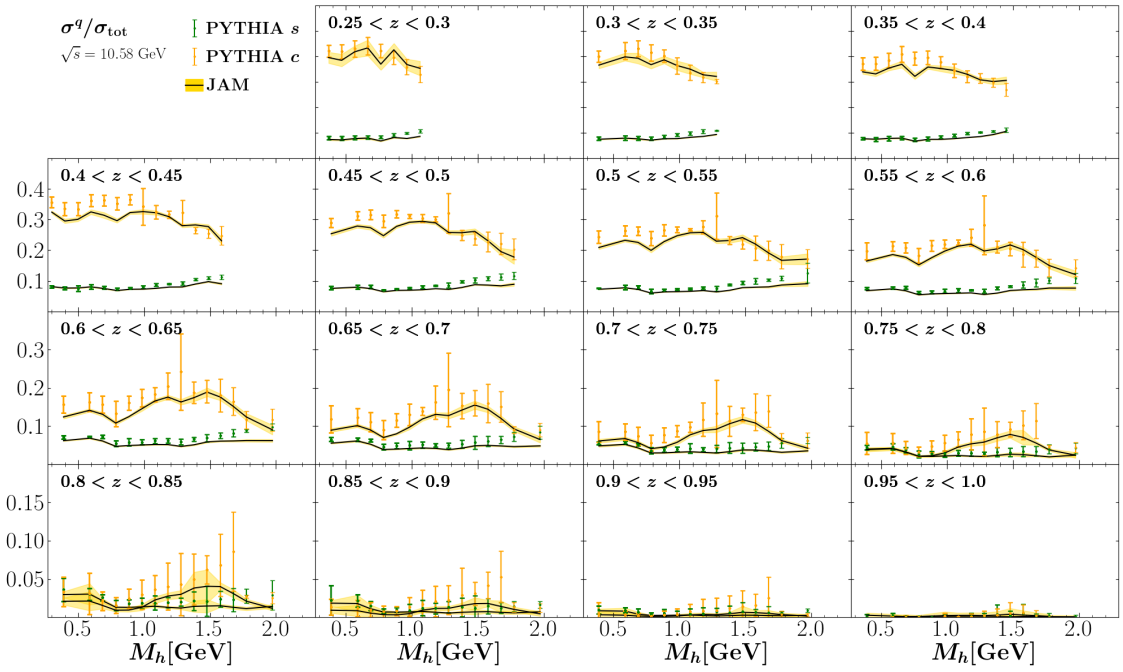

Figure A.1: PYTHIA data vs. Theory ($\sqrt{s} = 10.58$ GeV). Plotted are the strange (green points) and charm (orange points) cross section ratios plotted as a function of M_h against the mean JAM result (black line) with 1σ uncertainty bands in gold. The different panels show different bins of z.

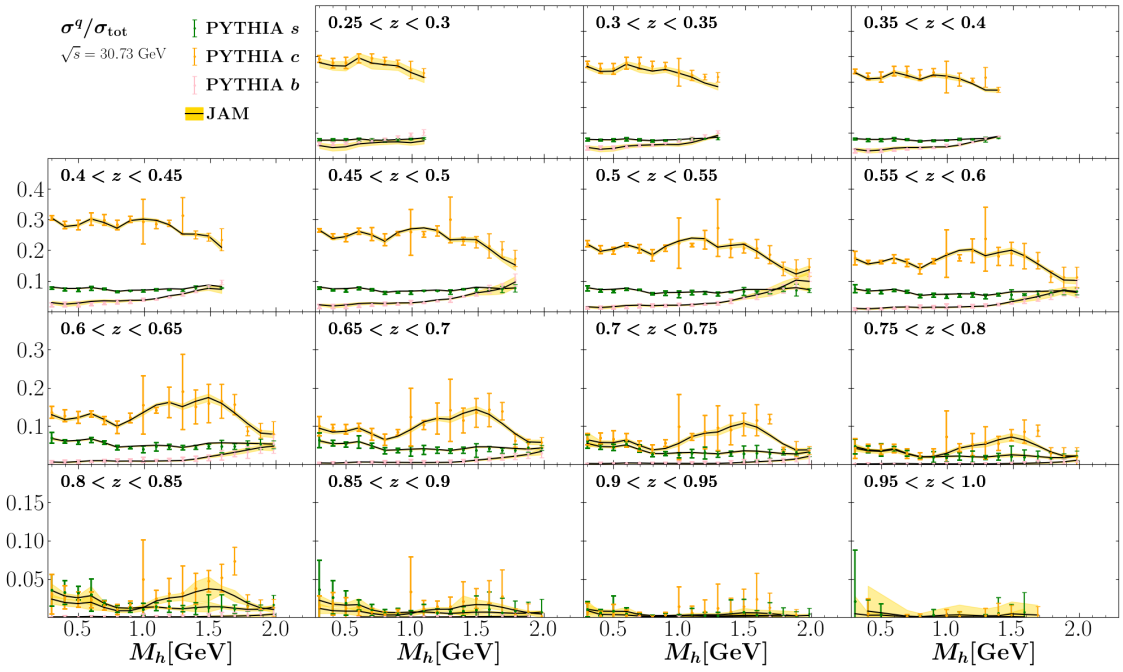

Figure A.2: PYTHIA data vs. Theory ($\sqrt{s} = 30.73$ GeV). Plotted are the strange (green points), charm (orange points), and bottom (pink points) cross section ratios plotted as a function of M_h against the mean JAM result (black line) with 1σ uncertainty bands in gold. The different panels show different bins of z.

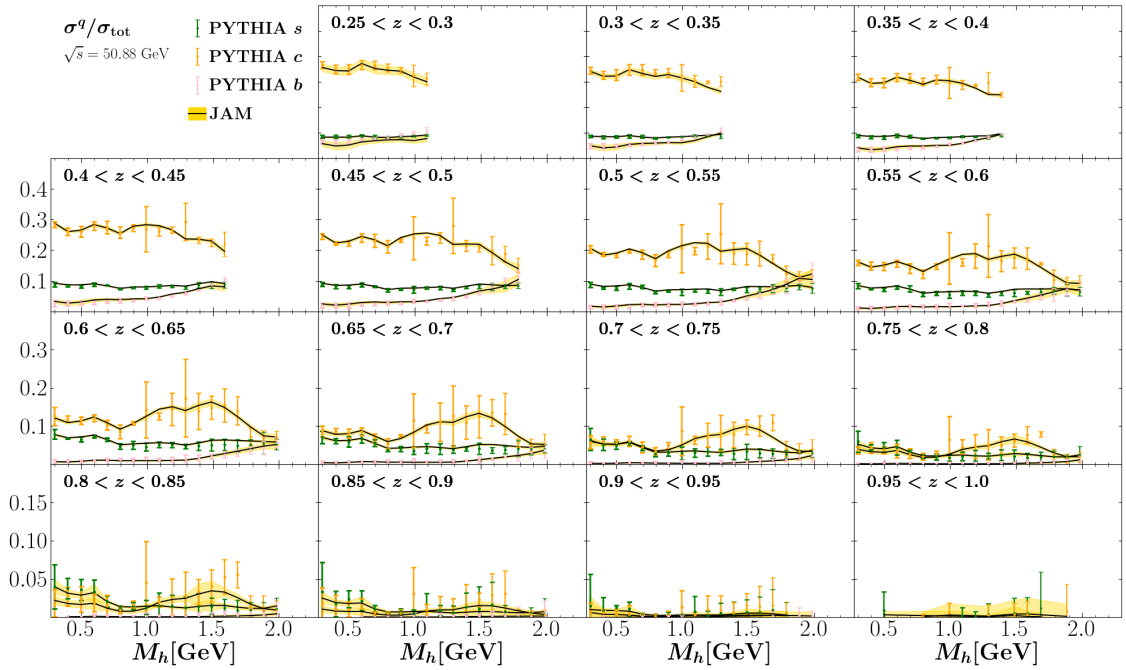

Figure A.3: PYTHIA data vs. Theory ($\sqrt{s} = 50.88$ GeV). Same as Fig. A.2 but for $\sqrt{s} = 50.88$ GeV.

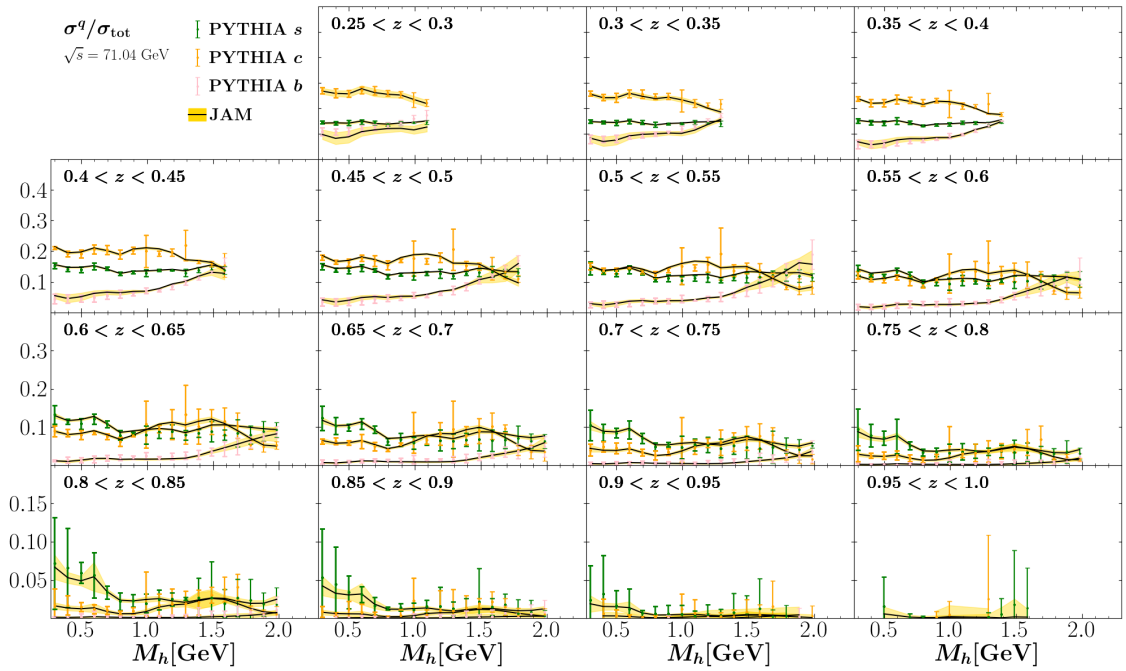

Figure A.4: PYTHIA data vs. Theory ($\sqrt{s} = 71.04$ GeV). Same as Fig. A.2 but for $\sqrt{s} = 71.04$ GeV.

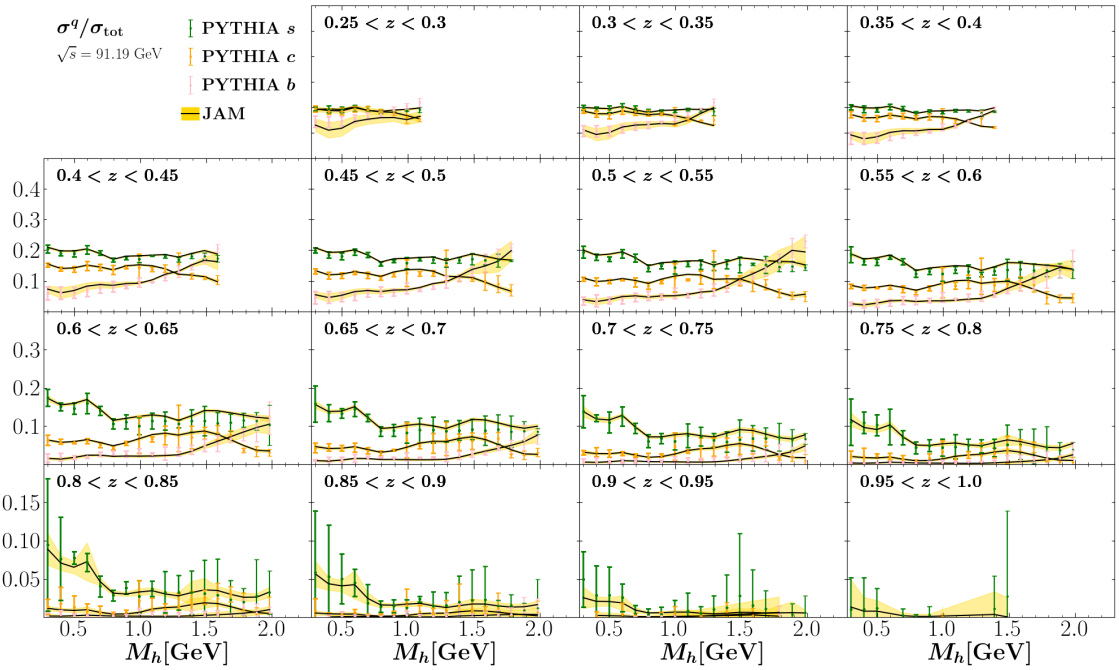

Figure A.5: PYTHIA data vs. Theory ($\sqrt{s} = 91.19$ GeV). Same as Fig. A.2 but for $\sqrt{s} = 91.19$ GeV.

APPENDIX B

RECOVERING LIGHT-CONE GPDS FROM QUASI-GPDS

Here we provide the most important steps involved for checking that the quasi-GPDs reduce analytically to the respective light-cone GPDs in the $P^3 \to \infty$ limit (see Subsection 6.2.2). We start with the poles of the propagators, which can be expanded as

$$k^0_{1+} = \begin{cases} xP^3 + \dfrac{1}{2(x+\xi)P^3} \\ \times\left[\vec{k}_\perp^2 - \vec{k}_\perp\cdot\vec{\Delta}_\perp - \dfrac{t}{4}(x\,\xi+1) + m_q^2 + x\,\xi M^2 - i\varepsilon\right] + \mathcal{O}\left(\dfrac{1}{(P^3)^2}\right) & x \geq -\xi\,, \\ -(x+2\xi)P^3 - \dfrac{1}{2(x+\xi)P^3} \\ \times\left[\vec{k}_\perp^2 - \vec{k}_\perp\cdot\vec{\Delta}_\perp - \dfrac{t}{4}(x\,\xi+1) + m_q^2 + x\,\xi M^2 - i\varepsilon\right] + \mathcal{O}\left(\dfrac{1}{(P^3)^2}\right) & x \leq -\xi\,, \end{cases}$$

$$k^0_{1-} = \begin{cases} -(x+2\xi)P^3 - \dfrac{1}{2(x+\xi)P^3} \\ \times\left[\vec{k}_\perp^2 - \vec{k}_\perp\cdot\vec{\Delta}_\perp - \dfrac{t}{4}(x\,\xi+1) + m_q^2 + x\,\xi M^2 - i\varepsilon\right] + \mathcal{O}\left(\dfrac{1}{(P^3)^2}\right) & x \geq -\xi\,, \\ xP^3 + \dfrac{1}{2(x+\xi)P^3} \\ \times\left[\vec{k}_\perp^2 - \vec{k}_\perp\cdot\vec{\Delta}_\perp - \dfrac{t}{4}(x\,\xi+1) + m_q^2 + x\,\xi M^2 - i\varepsilon\right] + \mathcal{O}\left(\dfrac{1}{(P^3)^2}\right) & x \leq -\xi\,, \end{cases}$$

$$k^0_{2+} = \begin{cases} xP^3 + \dfrac{1}{2(x-\xi)P^3} & \\ \times\left[\vec{k}_\perp^2 + \vec{k}_\perp\cdot\vec{\Delta}_\perp + \dfrac{t}{4}(x\,\xi - 1) + m_q^2 - x\,\xi M^2 - i\varepsilon\right] + \mathcal{O}\left(\dfrac{1}{(P^3)^2}\right) & x \geq \quad \xi\,, \\ -(x-2\xi)P^3 + \dfrac{1}{2(\xi-x)P^3} & \\ \times\left[\vec{k}_\perp^2 + \vec{k}_\perp\cdot\vec{\Delta}_\perp + \dfrac{t}{4}(x\,\xi - 1) + m_q^2 - x\,\xi M^2 - i\varepsilon\right] + \mathcal{O}\left(\dfrac{1}{(P^3)^2}\right) & x \leq \quad \xi\,, \end{cases}$$

$$k^0_{2-} = \begin{cases} -(x-2\xi)P^3 - \dfrac{1}{2(x-\xi)P^3} & \\ \times\left[\vec{k}_\perp^2 + \vec{k}_\perp\cdot\vec{\Delta}_\perp + \dfrac{t}{4}(x\,\xi - 1) + m_q^2 - x\,\xi M^2 - i\varepsilon\right] + \mathcal{O}\left(\dfrac{1}{(P^3)^2}\right) & x \geq \quad \xi\,, \\ xP^3 - \dfrac{1}{2(\xi-x)P^3} & \\ \times\left[\vec{k}_\perp^2 + \vec{k}_\perp\cdot\vec{\Delta}_\perp + \dfrac{t}{4}(x\,\xi - 1) + m_q^2 - x\,\xi M^2 - i\varepsilon\right] + \mathcal{O}\left(\dfrac{1}{(P^3)^2}\right) & x \leq \quad \xi\,, \end{cases}$$

$$k^0_{3+} = \begin{cases} xP^3 + \dfrac{1}{2(x-1)P^3} & \\ \times\left[\vec{k}_\perp^2 + \dfrac{t}{4}(1-x) - (1-x)M^2 + m_s^2 - i\varepsilon\right] + \mathcal{O}\left(\dfrac{1}{(P^3)^2}\right) & x \geq \quad 1\,, \\ -(x-2)P^3 + \dfrac{1}{2(1-x)P^3} & \\ \times\left[\vec{k}_\perp^2 - \dfrac{t}{4}(1-x) + (1-x)M^2 + m_s^2 - i\varepsilon\right] + \mathcal{O}\left(\dfrac{1}{(P^3)^2}\right) & x \leq \quad 1\,, \end{cases}$$

$$k^0_{3-} = \begin{cases} -(x-2)P^3 + \dfrac{1}{2(x-1)P^3} & \\ \times\left[-\vec{k}_\perp^2 + \dfrac{t}{4}(1-x) - (1-x)M^2 - m_s^2 + i\varepsilon\right] + \mathcal{O}\left(\dfrac{1}{(P^3)^2}\right) & x \geq \quad 1 \\ xP^3 + \dfrac{1}{2(1-x)P^3} & \\ \times\left[-\vec{k}_\perp^2 - \dfrac{t}{4}(1-x) + (1-x)M^2 - m_s^2 + i\varepsilon\right] + \mathcal{O}\left(\dfrac{1}{(P^3)^2}\right) & x \leq \quad 1\,, \end{cases}$$

It is evident from these equations that the analytical expressions of the expansions of the poles depend on x, but the poles always lie in the same half plane, as discussed in Chapter 6.

In the following we focus on the quasi-GPD $H_{\mathrm{Q}(0)}$. We first note that the dominant contribution is from those residues for which the leading order term is $x\,P^3$. Specifically, for the other residues the numerator of $H_{\mathrm{Q}(0)}$ has a leading contribution of order $(P^3)^3$, while the leading contribution of the denominator is of order $(P^3)^5$, resulting in an overall suppression like $1/(P^3)^2$. For $x \leq -\xi$ we close the integration contour in the lower half plane. Then none of the poles have $x\,P^3$ as the leading term, leading to a power-suppressed contribution. A corresponding discussion applies for $x \geq 1$ if one closes the integration contour in the upper half plane.

For the DGLAP region ($x \geq \xi$), we close the integration contour in the upper half plane. Then the dominant contribution comes from the residue at the pole k^0_{3-}. Therefore in that region

$$\lim_{P^3\to\infty} H_{\mathrm{Q}(0)} = -\lim_{P^3\to\infty} \frac{g^2\,P^3}{(2\pi)^3}\int d^2\vec{k}_\perp \times \frac{N_{H(0)}(k^0_{3-})}{(k^0_{3-}-k^0_{1+})(k^0_{3-}-k^0_{1-})(k^0_{3-}-k^0_{2+})(k^0_{3-}-k^0_{2-})(k^0_{3-}-k^0_{3+})}\,. \tag{B.1}$$

The leading order term of the numerator is given by

$$\begin{aligned} N_{H(0)}(k^0_{3-}) &= \delta\,(k^0_{3-})^2 - 2\frac{k^0_{3-}}{P^3}\left[x(P^3)^2 - m_q M - x\frac{t}{4} - \frac{\delta\xi t}{2}\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}^2_\perp}\right] \\ &\quad +\delta\left[x^2(P^3)^2 + \vec{k}^2_\perp + m_q^2 + (1-2x)\frac{t}{4} - \delta\xi t\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}^2_\perp}\right]. \end{aligned} \tag{B.2}$$

Then using

$$\begin{aligned} \delta(k^0_{3-})^2 &= \left(1+\frac{M^2-\frac{t}{4}}{2(P^3)^2}\right)\times x^2(P^3)^2\left(1-\frac{\vec{k}^2_\perp+\frac{t}{4}(1-x)-(1-x)M^2+m_s^2}{2x(1-x)(P^3)^2}\right)^2+\dots \\ &= x^2(P^3)^2+\frac{1}{2}x^2\left(M^2-\frac{t}{4}\right) \\ &\quad -\frac{x}{(1-x)}\left(\vec{k}^2_\perp+\frac{t}{4}(1-x)-(1-x)M^2+m_s^2\right)+\dots\,, \end{aligned} \tag{B.3}$$

where $\dots$ indicates suppressed terms, and

$$2\frac{k^0_{3-}}{P^3} \approx 2x + \frac{1}{(1-x)(P^3)^2}\left(\vec{k}^2_\perp+\frac{t}{4}(1-x)-(1-x)M^2+m_s^2\right)+\dots\,, \tag{B.4}$$

provides

$$
\begin{aligned}
N_{H(0)}(k^0_{3-}) &= x^2(P^3)^2 + \frac{1}{2}x^2\left(M^2 - \frac{t}{4}\right) - \frac{x}{(1-x)}\left(\vec{k}_\perp^2 + \frac{t}{4}(1-x) - (1-x)M^2 + m_s^2\right) \\
&\quad -2x^2(P^3)^2 + \frac{x}{(1-x)}\left(\vec{k}_\perp^2 + \frac{t}{4}(1-x) - (1-x)M^2 + m_s^2\right) \\
&\quad +2x\big(m_q M + x\frac{t}{4}\big) + x\xi t\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2} + x^2(P^3)^2 \\
&\quad +\frac{1}{2}x^2\left(M^2 - \frac{t}{4}\right) + \vec{k}_\perp^2 + m_q^2 + (1-2x)\frac{t}{4} - \xi t\frac{\vec{k}_\perp\cdot\vec{\Delta}_\perp}{\vec{\Delta}_\perp^2} + \ldots \\
&\quad = N_H + \ldots\,. \qquad \text{(B.5)}
\end{aligned}
$$

Additionally, the denominator in Eq. (B.1) simplifies as

$$
\begin{aligned}
&(k^0_{3-} - k^0_{1+})(k^0_{3-} - k^0_{1-})(k^0_{3-} - k^0_{2+})(k^0_{3-} - k^0_{2-})(k^0_{3-} - k^0_{3+}) \\
&= -8(P^3)^3(x^2 - \xi^2)(1-x)(k^0_{3-} - k^0_{1+})(k^0_{3-} - k^0_{2+}) + \ldots \\
&= -\frac{2P^3}{(1-x)(1-\xi^2)}D_1 D_2^{x\geq\xi} + \ldots\,. \qquad \text{(B.6)}
\end{aligned}
$$

Using Eqs. (B.5) and (B.6) in Eq. (B.1), one readily confirms

$$
\lim_{P^3\to\infty} H_{\mathrm{Q}(0)} = \frac{g^2(1-x)(1-\xi^2)}{2(2\pi)^3}\int d^2\vec{k}_\perp \frac{N_H}{D_1\, D_2^{x\geq\xi}} = H\,. \qquad \text{(B.7)}
$$

The overall logic to analytically recover H in the ERBL region ($-\xi \leq x \leq \xi$) remains the same as discussed above. In this case it is convenient to close the integration contour in the lower half plane, so that the dominant contribution comes from the residue at k^0_{1+} only. With a very similar analysis we have shown that all the quasi-GPDs reduce to the corresponding light-cone GPDs in the large-P^3 limit.