Università degli Studi di Pavia
Dipartimento di Fisica

DOTTORATO DI RICERCA IN FISICA – XXXI CICLO

Phenomenological analysis of unpolarized partonic Transverse Momentum Distributions and Sivers Function

Filippo Delcarro

Submitted to the Graduate School of Physics in partial fulfilment of the requirements for the degree of

DOTTORE DI RICERCA IN FISICA

DOCTOR OF PHILOSOPHY IN PHYSICS

at the

University of Pavia

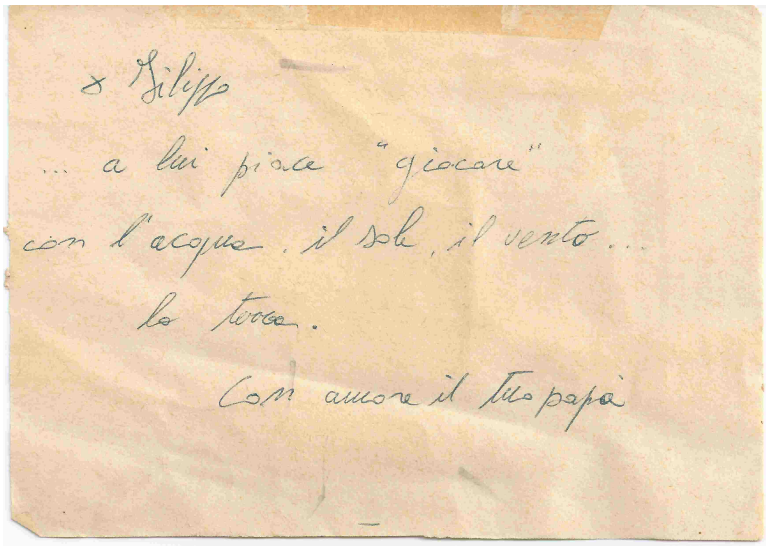
Supervisor: Prof. Alessandro Bacchetta

Cover: *Lucio Fontana, Concetto spaziale, La fine di Dio (1964)*

Phenomenological analysis of unpolarized partonic Transverse Momentum Distributions and Sivers Function

Filippo Delcarro

PhD thesis - University of Pavia
Pavia, Italy, January 2019



A mia madre e mio padre,
che hanno previsto la mia passione per la fisica
quando ero ancora un bambino,
e attraverso il loro sostegno,
hanno permesso che crescesse.

Contents

Chapter 1

Introduction

1.1 Phenomenological study of particles

The phenomenological analysis of the internal structure of nucleons can offer a picture of the composition of matter in terms of its fundamental constituents, quarks and gluons, collectively called partons. Through the combined efforts of experimental and theoretical research we try to achieve a better understanding of the inner dynamics of nucleons and to find a solution to many unanswered questions that separate us from a complete description of their interactions.

It is known that the interactions of partons are governed by the strong force, which is described by Quantum Chromodynamics (QCD). However, we are not able to entirely describe the behavior of the nucleon starting from the dynamics of its constituents, in other words, we do not fully understand QCD in nonperturbative regime and, in particular, we cannot describe its most crucial feature, the confinement of quarks and gluons. Moreover, the properties of nucleons are very well known and high precision measurements of its spin, quark content, charge, magnetic moment and charge radius are available at the moment. Nonetheless, we are not able to explain these quantities starting from first principles, that is, considering simply the properties of partons. In recent years we started gathering the information necessary to reconstruct multi-dimensional maps of the nucleon, which allows the analysis of properties of the nucleons otherwise inaccessible and could help us to understand more accurately QCD and its properties.

We can investigate the internal partonic structure observing the outcome of hard scattering processes. The quantities measured during this collisions are related to the distributions of partons inside the nucleons and can noticeably change with the type of process considered and the energy scale. Through the measurement of scattering processes we cannot obtain the full information of the hadron structure available, but only certain projections on a part of the coordinate and momentum space.

The study of the fundamental constituents of matter has progressed in

parallel with the discovery of new methods to observe them with increasing details. The use of particle scattering as a fundamental tool to investigate the internal structure of the constituents of matter became first apparent with the gold foil scattering experiment performed by Geiger, Marsden and Rutherford in 1909 [1]. The experiment was originally conceived as a test to verify Thomson's model of the atom. However its outcome was not compatible with predictions and gave a contrasting picture of the atomic inner composition, which eventually led to the proposal of a planetary model for the atom, where negatively charged electrons orbit around a positively charged nucleus. The gold foil experiment opened the way to the investigation of the structure of fundamental constituents through the phenomenology of scattering processes.

The difference between theoretical predictions based on an atomic model with a point-like nucleus and experimental observations, which suggested a more complex structure, has been since used to explore the internal configuration of the atomic nucleus and, subsequently, of the nucleon. About fifty years after the original scattering experiments, Hofstadter proposed the theory of nuclear form factors [2], related to the internal charge distribution, to explain the ratio between the theoretical estimate of the Rutherford model and the experimental measurements. This concept was later extended also to protons and neutrons [3].

A step further was taken in 1967 at the Stanford Linear Accelerator Center (SLAC) with deep inelastic scattering (DIS) experiments [4, 5]. In this process electrons are scattered off a proton target, in a way similar to the original Rutherford experiment. There is however an important distinction, as the energy of the probe is greatly increased, enabling to explore the internal structure of the target to a much higher resolution.

Analysis of the spectrum of baryons and mesons, through high-energy scattering, led to the discovery of quarks, the fundamental constituents of strongly interacting particles [6, 7]. The observation coming from DIS experiments led to an image of the proton as a cloud of free point-like fermions, which was named *parton model* [8].

The study of the relation between the transverse motion of partons and the corresponding effects on hard processes involving hadrons started in the 70's with the papers by Feynman, Fox and Field [9, 10]. Their work can be considered as the starting point for the generalization of the parton model to include also contributions to the dynamics of partons coming from their transverse momentum,which is the momentum component perpendicular to the longitudinal direction, in turn defined as the direction along the proton momentum in the center of mass of proton and probe. They proposed that the transverse momentum measured in Drell–Yan processes could originate either from non-perturbative effects, such as intrinsic momentum of partons confined in the nucleons, or from perturbative effect, as radiative corrections for gluons emitted by active quarks, thus connecting for the first time the outcome of hadronic high-energy processes with the basic properties of partons.

In recent years, the unpolarized structure of the proton was explored in deep inelastic electron-proton scattering experiments, such as H1 and ZEUS at DESY. At the same time, collaborations at CERN, SLAC, Jefferson Lab and HERMES mapped the polarized nucleon structure. Along with the experimental breakthroughs, a theoretical framework was developed, which describes the three-dimensional picture of nucleons in terms of a wide variety of partonic distributions.

The investigation of the three-dimensional hadron structure is motivated not only by the pursuit of a more complete description of matter, but also by several open questions in QCD physics. One of the fundamental problems that still have no answer is the proton spin puzzle. While initially it was expected that the well known total value 1/2 of the nucleon spin should come from the spin of its quarks, this was disputed by the experimental measurements taken by the EMC collaboration and by the subsequent ones [11]. The missing contribution could come from the spin of gluons or the orbital angular momentum of the single partons, which cannot be studied through a simple one-dimensional description of their distribution. In addition, the experimental investigation of hard processes and the search for physics beyond the standard model is moving towards the observation of effects that require very high precision, where it becomes necessary to exclude contributions that could come from the internal composition and momentum of the hadronic constituents.

As an example, a recent study [12] investigated the uncertainties on the determination of the W boson mass extracted at the LHC, which could be induced by a possible flavor dependence of partonic intrinsic transverse momentum. The results suggests that a data analysis which ignores effects of flavor dependence may be not sufficiently accurate and that a complete flavor decomposition of the unpolarized TMDs could be necessary for future precision measurements at high energies.

1.2 3D mapping of the nucleon

The description of the internal structure of a hadron can be achieved by determining the specific momenta and positions of its partons [13]. Their state can be described by Wigner distributions in five dimensions. They are close to a classical probability density in phase space, but they cannot be consistently considered as probability densities and are not positive definite. Nevertheless, they can be used to compute the expectation value of any physical observable. From this point of view, they represent the complete knowledge of the partonic structure. We cannot access directly Wigner distributions, but we are able to observe their projections on some of the available dimensions in coordinate and momentum space.

For example, integrating over all coordinates and the two transverse components of momentum, we obtain one of the fundamental ways to describe the nucleons: the collinear parton distribution functions (PDF), which give the

quark pol.

nucleon pol.	U	L	T
U	$\boldsymbol{f_1}$		$h_1^\perp$
L		$\boldsymbol{g_{1L}}$	$h_{1L}^\perp$
T	$f_{1T}^\perp$	g_{1T}	$\boldsymbol{h_1}, h_{1T}^\perp$

(a) TMD PDFs

quark pol.

hadron pol.	U	L	T
U	$\boldsymbol{D_1}$		$H_1^\perp$
L		$\boldsymbol{G_{1L}}$	$H_{1L}^\perp$
T	$D_{1T}^\perp$	G_{1T}	$\boldsymbol{H_1}, H_{1T}^\perp$

(b) TMD FFs

Table 1.1: Twist-2 quark TMD distribution functions (a) and fragmentation functions (b). U, L, T correspond to unpolarized, longitudinally polarized and transversely polarized nucleons or hadrons, and quarks. Functions in black are T-even, while functions in red are T-odd. Functions in boldface survive transverse momentum integration.

probability of finding a quark or gluon inside a proton with a certain fraction x of the collinear momentum. They represent one-dimensional maps of the hadron structure in momentum space. They are essential objects for the description of hadronic processes and have been studied in depth, however they are sensitive only to the momentum in the collinear direction.

Considering also the dependence of partonic distributions on the transverse momentum $\boldsymbol{k}_\perp$ we can obtain more complete three-dimensional maps of the nucleons in momentum space. This 3D picture is defined through Transverse Momentum Dependent (TMD) parton distribution functions, which can be considered as generalizations of collinear distributions. There are many non-trivial questions related to TMDs which are still unanswered. For instance, we do not have enough information on the partonic density to distinguish if it is higher in the center of the map or if it increases towards the border. Moreover, the dependence of these distributions on the energy scale and on the parton flavor is not completely understood.

To have a complete description of high energy processes we need to introduce also TMD fragmentation functions (FF), which give the probability that a quark fragments into a certain hadron in the final state, as a function of the hadron collinear and transverse momentum.

The foundation of the present formalism for TMDs can be traced back to the work of Collins and Soper [14], where they combined for the first time the nonperturbative and perturbative components of TMDs in a formally coherent way.

It is possible to introduce different TMDs depending on the polarization of the nucleon (or the final-state hadron in the case of FFs) and of the partons

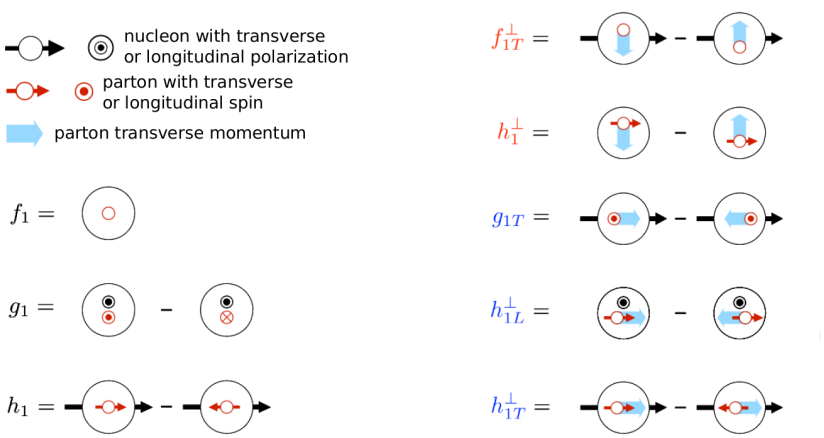


Figure 1.1: Probabilistic interpretation of twist-2 TMDs. To avoid ambiguities we specify that the proton is moving out of the page.

inside [15, 16, 17, 18, 19, 20]. For the moment, we consider twist-2 TMDs, i.e., we include only leading-power contributions of the (M/Q) expansion used for factorization theorems. At this level of accuracy we can distinguish 8 different TMD PDFs and an equal number of TMD FFs. They can be observed in Table 1.1, where U, L, T denotes unpolarized, longitudinally and transversely polarized particles, respectively. For completeness we included also the polarized fragmentation functions, even if they will not be necessary in our work. Their probabilistic interpretations in terms of quark transverse momentum, and of the spin of target and quark are presented in Fig. 1.1. They can be seen as the interference of the amplitudes given by all the possible combinations of nucleons and partons with a specific polarization.

In the last years there has been a great research effort to determine the dependence of TMDs in relation to collinear and transverse momentum, energy and flavor of partons. At present, they cannot be easily computed from first principles, because they require the ability to carry out QCD calculations in its nonperturbative regime. Nevertheless, we can access them through their relation with many different experimental observables in hard scattering experiments involving hadrons, in a way that is specified by factorization theorems (see, e.g., Refs. [21, 22]). These theorems also elucidate the universality properties of PDFs and FFs (i.e., the fact that they are the same in different processes) and their evolution equations (i.e., how they get modified by the change in the hard scale of the process). Availability of measurements of different processes in different experiments makes it possible to test factorization theorems and extract PDFs and FFs through so-called global fits. Attempts to extract some of the TMDs discussed have already been presented in the literature [23, 24, 25, 26, 27, 28, 29, 30, 31].

In this thesis, we will initially focus on the unpolarized TMD PDF and the unpolarized TMD FF. Even if they do not take into account any polarization,

they are fundamental for the description of the nucleons. Moreover an accurate determination of these unpolarized distributions is crucial also for the determination of the polarized TMDs, because the observables used to investigate them require a good knowledge of the unpolarized structure, as we will observe in the case of the Sivers function.

1.3 Polarized processes and the Sivers function

Transverse spin phenomena have gained increased attention in the investigation of the 3D picture of the nucleon in momentum space. As we observed in Table 1.1, it is possible to define different TMDs considering also polarization. In particular, in this thesis we will focus on the extraction of the function $f_{1T}^{\perp}$, proposed by Sivers [32], which represents the number density of unpolarized partons inside a transversely polarized nucleon.

In addition to giving informations on the internal structure of hadrons, the study of polarized processes in a three-dimensional framework allows to observe the correlations between transverse momentum and spin, that is spin-orbit correlations at the partonic level. These properties cannot be studied in a collinear approximation, where orbital motion is not considered.

Our current knowledge of QCD in the perturbative regime still presents challenging question regarding spin and its correlation with the motion of partons inside the nucleon. We already cited the missing spin budget of the hadron. Another phenomenon that leaves many open questions is the behavior of transverse Single-Spin Asymmetries (SSA) [33]. These are commonly observed in inclusive hadron production in high-energy collisions of protons with large p_T, and it was expected that large transverse SSAs could not be observed in the hard elementary process. In fact, in QED and pQCD interactions, due to their intrinsic property of helicity conservation in the massless limit, these asymmetries should vanish with the increase of energy scale. Their persistence in many experimental measurements has to be attributed to contributions outside the perturbative regime of QCD [34]. These nonperturbative contributions are also indirectly connected to parton orbital motion. Therefore, being able to relate the origin of these transverse SSAs to intrinsic property of nucleons and their components can be useful to achieve a deeper knowledge of the inner dynamics of nucleons.

Two main methods have been used to investigate the transverse single-spin asymmetries. One is called *twist-3 approach* and it is based on the collinear QCD factorization and describe the process in terms of higher-twist quark-gluon correlations. Another possible approach is generalized parton model (GPM) that expands the parton model focusing on its phenomenology, by including TMD parton distributions and fragmentation functions. There is a proven connection between twist-3 correlations and the moments of the cor-

responding TMDs [35]. In this thesis, however, we will not consider twist-3 asymmetries but rather focus on twist-2 asymmetries that can be interpreted without ambiguity in terms of TMDs, in particular of the Sivers asymmetry.

The description of the asymmetries in terms of TMDs was firstly theorized at the beginning of the 90's, when two fundamental correlations between partonic transverse motion and nucleon spin were proposed by Sivers [32, 36] and Collins [37]. The function $H_1^\perp$ proposed by Collins describes the correlation between the transverse spin of a fragmenting quark in the initial state and the transverse momentum of the final-state hadron. The Collins fragmentation function, convoluted with the quark transversity distribution h_1, can be accessed in SIDIS experiments, studying the azimuthal modulation that it induces due to its dependence on the hadronic transverse momentum.

While Collins described the dynamics of final-state hadrons, Sivers proposed a new TMD distribution functions, which now bears his name, describing the position of unpolarized quarks inside a transversely polarized hadron. The Sivers distribution could be used to describe the large single-spin asymmetries in pion production off hadron-hadron scattering. The seminal result of Sivers was to suggest that the transverse momentum distribution of a quark inside an hadron could include an azimuthal asymmetry if the initial hadrons have transverse polarization. However, this claim was dismissed by Collins [37] who observed that the existence of a nonzero asymmetry would violate time-reversal invariance for strong interactions. He showed that to satisfy these symmetry properties, the number density of partons should not be dependent on the spin of the hadron in the initial state.

The study of the Sivers distribution was then abandoned for about a decade, until Brodsky, Hwang and Schmidt [38] showed that taking into account final-state interactions in deep inelastic lepton-proton scattering, where the produced hadron and the target spectator system exchange gluons, a nonzero single-spin asymmetry at leading twist is obtained. Shortly after, the same authors [39] showed that a similar mechanism exists also for initial state interaction for polarized Drell–Yan process and generates a leading twist single spin asymmetry. In this case the required phase interference is produced by the initial-state interaction between the annihilating quark and the spectator target. The single spin asymmetry derived for the Drell–Yan process with transverse polarization $\pi p_\uparrow \to \ell^+\ell^- X$ is similar to the one which follows from final-state interactions in deep inelastic semi-inclusive leptoproduction $\ell p_\uparrow \to \ell' \pi X$. The resulting SSA calculated for Drell–Yan and for SIDIS processes present however a fundamental difference: the Sivers function appears with an opposite sign.

As a consequence, Collins [40] recognized that his proof for a vanishing Sivers distribution was not valid. In fact, it did not correctly take into account the presence of Wilson lines in the operators, necessary to satisfy gauge invariance. The future-pointing Wilson lines under time-reversal transform in past pointing Wilson lines, leading to the conclusion that the Sivers asymme-

try does not vanish, but instead has an opposite sign for DIS and Drell–Yan processes.

This peculiarity constitutes a crucial experimental check for the polarized TMD framework, which however still awaits verification. A precise determination of the Sivers function, as the one presented in the thesis, is an essential step towards this important goal.

1.4 Outline of the thesis

In this thesis we present an extraction of the unpolarized TMD PDFs and FFs from semi-inclusive deep inelastic scattering, Drell–Yan process and Z boson production, taking into account the contribution of evolution at different energy scales. Afterwards, we present a determination of the Sivers distribution form polarized semi-inclusive DIS, with the inclusion of TMD evolution and consistent unpolarized TMDs.

The work is organized in the following way. In Ch. 2, the theory for transverse momentum dependent parton distribution and fragmentation function is shown, focusing on the unpolarized ones and on the Sivers distribution. We discuss the relation between the experimental observables and the TMDs and how they evolve with respect to their characteristic scales.

In Ch. 3 the parametrization for the description of TMDs in SIDIS, DY processes, and Z production is briefly outlined, including a description of the assumptions and approximations in the phenomenological implementation of TMD evolution equations. We present the data included in the global fit and the criteria for selecting the data analyzed. In the last part we show the results of our analysis and their stability.

In Ch. 4 we discuss our framework for the extraction of the Sivers function from SIDIS data, coming from Hermes , Compass , and JLab experiments. In the first part the parametrization and choices for the evolution of the Sivers are discussed. Finally, we examine the agreement between our theoretical prediction and the experimental measurements.

Chapter 2

Theory of polarized and unpolarized Transverse Momentum Dependent distributions

In this chapter we outline the formalism needed to describe polarized and unpolarized scattering processes in terms of Transverse Momentum Dependent PDFs and FFs. Starting from the case of deep inelastic scattering, we proceed to discuss the semi-inclusive case and the Drell–Yan process. Then we present the framework necessary to the investigation of the Sivers function in the case of polarized SIDIS. In addition, the evolution in relation to the hard energy scale of the TMDs discussed is shown.

2.1 Inclusive DIS

In order to understand the relation between the partonic internal structure of nucleons and scattering processes, it is interesting to observe how the experimental observables of inclusive DIS are connected with hadronic structure functions in parton model. Deep inelastic scattering can be represented schematically as

$$\ell(l) + N(P) \to \ell(l') + X \tag{2.1}$$

where ℓ denotes a lepton in the initial beam with a four-momentum l that scatters of a nucleon N, changing its momentum and producing a remnant X. The invariants usually used to study the kinematics of the process are[1]

$$x_B = \frac{Q^2}{2P \cdot q} \qquad\qquad y = \frac{P \cdot q}{P \cdot l} = \frac{Q^2}{x_B s} \tag{2.2}$$

[1]To avoid confusion, we point out that P refers to the momentum of the proton, while P_h is related to the momentum of the final state hadron.

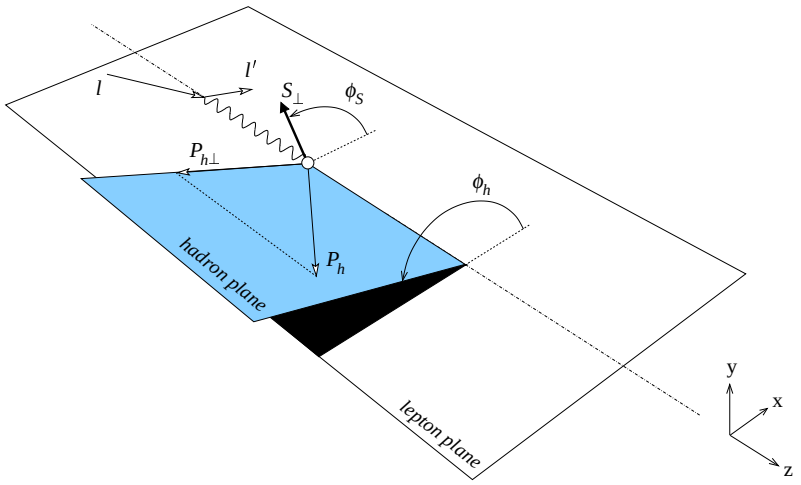


Figure 2.1: Kinematics of a lepton scattering off a hadron, in a frame with parallel hadron and virtual photon momenta.

where the virtuality of the photon is $q = l - l'$ and the hard scale is $Q^2 = -q^2$; we will denote with λ the helicity of the leptons and with S the spin of the hadron. The dynamics for this process is represented in Fig. 2.1.

We want to obtain the expression for the cross section of inclusive DIS in terms of its fundamental components. The invariant amplitude for this process is calculated as

$$\mathcal{M} = \bar{u}(k', \lambda')\gamma^\mu u(k, \lambda)\frac{e^2}{Q^2}\langle P_X|J_\mu(0)|PS\rangle \tag{2.3}$$

where a sum over spins of the unobserved state X is implied.

The square of this amplitude, related to the cross section of a lepton-nucleon scattering, can be separated into a tensor L representing the leptonic part and a hadronic tensor W:

$$\left|\mathcal{M}^2\right| = \frac{e^4}{Q^4}L_{\mu\nu}W^{\mu\nu}\,. \tag{2.4}$$

The expressions contained in Eqs. (2.3) and (2.4) are valid in the *single-photon exchange* approximation.

The total scattering cross section then becomes

$$d\sigma = \frac{1}{F}\frac{e^4}{Q^4}L_{\mu\nu}W^{\mu\nu}d\mathcal{R} \tag{2.5}$$

with the Lorentz invariant phase space

$$d\mathcal{R} = (2\pi)^4\delta^4(l + P - l' - P_X)\frac{d^3P_X}{(2\pi)^3 2P_X^0}\frac{d^3l'}{(2\pi)^3 2l'^0} \tag{2.6}$$

and the flux factor

$$F = 4\sqrt{(P\cdot k)^2 - M^2m^2} \approx 2s \tag{2.7}$$

In this way, the information on the leptonic probe, which can be described resorting only to perturbative QED, is isolated from the information on the hadronic target. Without considering the mass contributions, the leptonic tensor can be written as

$$\begin{aligned} L_{\mu\nu}(l,l',\lambda_e) =& \sum_{\lambda'_e} \left(\bar{u}\left(l',\lambda'_e\right)\right)\gamma_\mu u\left(l,\lambda_e\right))^* \left(\bar{u}\left(l',\lambda'_e\right)\right)\gamma_\nu u\left(l,\lambda_e\right)) \\ =& -Q^2 g_{\mu\nu} + 2\left(l_\mu l'_\nu + l'_\mu l_\nu\right) + 2i\lambda_e \epsilon_{\mu\nu\rho\sigma} l^\rho l'^\sigma \,. \end{aligned} \tag{2.8}$$

The hadronic tensor $W^{\mu\nu}$ depends on the process considered and in general cannot be explicitly calculated from first principles, due to nonperturbative contributions from the nucleon. However it can be parametrized exploiting the known symmetric property of QCD. The hadronic tensor can be calculated as the product of currents that, acting on the on-shell initial state, generate the final state of the process, containing some remnant X of the hadron, eventually in addition to some other on-shell particles, depending on the process considered:

$$W^{\mu\nu} \sim \langle in|j^\mu(0)|out\rangle\langle out|j^\nu(0)|in\rangle \tag{2.9}$$

The hadronic tensor, after integration over all possible on-shell remnant X that is not observed in the final state, is given by

$$\begin{aligned} 2MW^{\mu\nu}(q,P,S) =& \frac{1}{2\pi}\sum_X \int \frac{d^3P_X}{(2\pi)^3 E_X}(2\pi)^4\delta^{(4)}\left(q+P-P_X\right) h^{\mu\nu}(P,S,P_X) \\ =& \frac{1}{2\pi}\int d^4x e^{iq\cdot x}\langle P,S|[J^\mu(x),J^\nu(0)]|P,S\rangle \end{aligned} \tag{2.10}$$

which contains the product of hadronic current matrix elements

$$H^{\mu\nu}(P,S;P_X) = \langle P,S|J^\mu(0)|P_X\rangle\langle P_X|J^\nu(0)|P,S\rangle \,. \tag{2.11}$$

It is useful to define a basis of vector suitable for the theoretical description of the hadron structure. In order to study deep inelastic processes it is convenient to define a Cartesian set of vectors derived from the particles momenta. For lepton-hadron scattering the vector basis is constructed starting from the spacelike direction of the momentum transfer q^μ. The next step is to define vectors $\tilde{a}$ that are orthogonal to q, obtained subtracting from some initial vector a its projections along q:

$$\tilde{a}^\mu = \tilde{g}^{\mu\nu}a_\nu = a^\mu - \frac{a\cdot q}{q^2}q^\mu \tag{2.12}$$

$$\tilde{g}^{\mu\nu} = g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2} \tag{2.13}$$

In particular, using the target hadron momentum P^μ the orthogonal vector becomes $\tilde{P}^\mu = P^\mu - (P\cdot q/q^2)q^\mu$. This vector is timelike and satisfies

$$\tilde{P}^2 = \kappa P\cdot q\,, \qquad\qquad \kappa = 1 + \frac{M^2Q^2}{(P\cdot q)^2} \tag{2.14}$$

where only mass corrections of the order (M^2/Q^2) are taken into account. To study the expansion of the hadronic tensor to increasing orders of $1/Q$ it is convenient to define

$$Z^\mu \equiv - q^\mu \tag{2.15}$$

$$T^\mu \equiv \frac{Q^2}{P\cdot q}\frac{\tilde{P}^\mu}{\sqrt{\kappa}} = \frac{q^\mu + 2x_B P^\mu}{\sqrt{\kappa}} \tag{2.16}$$

and we observe that their modules are directly related to the value of Q as

$$Z^2 = -Q^2 \qquad\qquad T^2 = Q^2 \ . \tag{2.17}$$

We will mainly use the normalized vectors $\hat{z}^\mu = Z^\mu/Q$ and $\hat{t}^\mu = T^\mu/Q$. Finally, in the space orthogonal to $\hat{z}$ and $\hat{t}$ one has the tensors

$$g_\perp^{\mu\nu} \equiv g^\mu\nu + \hat{q}^\mu\hat{q}^\nu - \hat{t}^\mu\hat{t}^\nu\epsilon_\perp^{\mu\nu} \equiv \epsilon^{\mu\nu\rho\sigma}\hat{t}_\rho\hat{q}_\sigma = \frac{1}{(P\cdot q)\sqrt{\kappa}}\epsilon^{\mu\nu\rho\sigma}P_\rho q_\sigma \tag{2.18}$$

Using these tensors we can introduce a normalized vector related to the lepton momentum, written as

$$\hat{l}^\mu = -\frac{g_\perp^{\mu\nu}l_\nu}{|g_\perp^{\mu\nu}l_\nu|} \tag{2.19}$$

The vector for the lepton momentum l can be expressed using the normalized vectors, as

$$l^\mu = \frac{Q}{2}\hat{q}^\mu + \frac{(2-y)}{2y}\hat{t}^\mu + \frac{Q\sqrt{1-y}}{y}\hat{l}^\mu \tag{2.20}$$

It is possible to rewrite the leptonic tensor using these normalized vectors:

$$\begin{aligned} L_{\mu\nu} = \frac{2Q^2}{y^2}\Bigg[& -\left(1-y+\frac{1}{2}y^2\right)g_{\perp\mu\nu} + 2(1-y)\hat{t}_\mu\hat{t}_\nu \\ & +2(1-y)(\hat{l}_\mu\hat{l}_\nu + \frac{1}{2}g_{\perp\mu\nu}) + (2-y)\sqrt{1-y}(\hat{t}_\mu\hat{l}_\nu + \hat{t}_\nu\hat{l}_\mu) \\ & +\frac{i}{2}\lambda_e y(2-y)\epsilon_{\perp\mu\nu} - i\lambda_e y\sqrt{1-y}(\hat{t}_\mu\epsilon_{\perp\nu} + \hat{t}_\nu\epsilon_{\perp\mu})\hat{l}^\rho \end{aligned} \tag{2.21}$$

To express also the hadronic tensors in terms of the normalized vectors, we observe that it has to satisfies the properties of Hermiticity, parity and time-reversal, respectively:

$$W^*_{\mu\nu}(q,P,S) = W_{\nu\mu}(q,P,S) \ , \tag{2.22}$$

$$L^\rho_\mu L^\sigma_\nu W_{\rho\sigma}(q,P,S) = W_{\mu\nu}(\tilde{q},\tilde{P},-\tilde{S}) \ , \tag{2.23}$$

$$L^\rho_\mu L^\sigma_\nu W_{\rho\sigma}(q,P,S) = W_{\mu\nu}(\tilde{q},\tilde{P},\tilde{S}) \tag{2.24}$$

where $\tilde{q}^\nu = L^\nu_\rho q^\rho$ and the parity reversal transformation L is defined as

$$L^\rho_\sigma = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{2.25}$$

A combination of the momentum vectors that fulfills the previous conditions, could be

$$2MW^{\mu\nu} = 2M\left[Ag^{\mu\nu} + Bq^\mu q^\nu + C\frac{P^\mu P^\nu}{M^2} + D\frac{P^\mu q^\nu + q^\mu P^\nu}{M^2}\right], \tag{2.26}$$

where each term might have a dependence on Q^2 and x_B. We have an additional constraint given by electromagnetic gauge invariance, that requires $q_\mu W^{\mu\nu} = W^{\mu\nu} q_\nu = 0$ and gives the following relations between the coefficients:

$$D = -\frac{P\cdot q}{q^2}B, \qquad\qquad C = \left(\frac{P\cdot q}{q^2}B + \frac{M^2}{q^2}A\right) \tag{2.27}$$

It follows then that only two coefficients are independent and we can define only two independent structure functions for the hadronic tensor. In terms of the normalized vectors we have

$$2MW^{\mu\nu}(q,P,S) = \frac{1}{x_B}\left[-g^{\mu\nu}_\perp F_T(x_B,Q^2) + \tilde{t}^\mu\tilde{t}^\nu F_L(x_B,Q^2)\right] \tag{2.28}$$

We take into account also the polarization for the inclusive DIS process, we can introduce four independent structure functions. A possible definition is

$$\begin{aligned} 2MW^{\mu\nu}(q,P,S) = \frac{1}{x_B}\Big[&-g^{\mu\nu}_\perp F_T(x_B,Q^2) + \hat{t}^\mu\hat{t}^\nu F_L(x_B,Q^2) \\ &+iS_L\epsilon^{\mu\nu}_\perp 2x_B\left(g_1(x_B;Q^2) - \gamma^2 g_2(x_b,Q^2)\right) \\ &+i\hat{t}^\mu\epsilon^{\nu\rho}_\perp S_\rho 2x_B\gamma\left(g_1(x_B,Q^2) + g_2(x_B,Q^2)\right)\Big]. \end{aligned} \tag{2.29}$$

In this discussion we should have distinguished the component of S parallel or orthogonal to the lepton beam instead of the virtual photon, but the difference between them at M/Q level is negligible.

The relation between these structure functions and the standard unpolarized ones is

$$F_T(x_B,Q^2) = 2x_B F_1(x_B,Q^2) \tag{2.30}$$

$$F_L(x_B,Q^2) = (1-\gamma^2)F_2(x_B,Q^2) - 2x_B F_1(x_B,Q^2) \tag{2.31}$$

Finally, contracting the leptonic and hadronic tensor, expressed in this way, for the inclusive DIS cross section we found

$$\begin{aligned}\frac{d\sigma}{dx_B dy d\phi_S} =& \frac{2\alpha^2}{x_B y Q^2}\frac{y^2}{2(1-\varepsilon)}\{F_T + \varepsilon F_L + S_\parallel \lambda_e \sqrt{1-\varepsilon^2}\, 2x_B(g_1 - \gamma^2 g_2) \\ & -|S_\perp|\lambda_e\sqrt{2\varepsilon(1-\varepsilon)}\cos\phi_S\, 2x_B\,\gamma(g_1+g_2)\}\end{aligned} \tag{2.32}$$

where we have introduced the ratio ϵ of longitudinal and transverse photon flux:

$$\varepsilon = \frac{1 - y - \frac{1}{4}\gamma^2 y^2}{1 - y + \frac{1}{2}y^2 + \frac{1}{4}\gamma^2 y^2} \qquad\qquad \gamma = \frac{2Mx_B}{Q} \tag{2.33}$$

2.1.1 DIS cross section in parton model

For the explicit calculation of DIS cross section in terms of parton distribution functions is convenient to choose a frame of reference where there are no transverse component of proton and photon momenta. Working in this framework and considering processes at high Q^2, we can describe DIS through a parton model, where we assume that the electrons scatter on an almost free quark of mass m inside the nucleons,producing a final state X. This state can be separated in a quark with momentum k and a state X with momentum P_X. In this picture the scattering process can be factorized in a hard part and a non-perturbative part, where the first one expresses the electromagnetic interaction between the lepton and the struck parton and the latter refers to the probability to find such a parton inside the nucleon.

In this model the hadronic tensor can be written as

$$\begin{aligned}2MW^{\mu\nu}(q,P,S) =& \frac{1}{2\pi}\sum_q e_q^2 \sum_X \int \frac{d^3\boldsymbol{P}_X}{(2\pi)^3 2P_X^0}\int \frac{d^3\boldsymbol{p}}{(2\pi)^3 2p^0}(2\pi)^4 \\ & \times\delta^{(4)}(P+q-p-P_X)\langle P,S|\bar\psi_i(0)|X\rangle\langle X|\psi_j(0)|P,S\rangle \\ & \times\gamma^\mu_{ik}(\not{p}+m)_{kl}\,\gamma^\nu_{lj}\,,\end{aligned} \tag{2.34}$$

considering only tree-level contributions for the electron-quark interaction. Using completeness to get rid of the unobserved states X and integrating over the phase space of the final-state quark we can rearrange the hadronic tensor in the following way:

$$\begin{aligned}2MW^{\mu\nu}(q,P,S) =& \sum_q e_q^2 \int d^4p\,\delta((k+q)^2 - m^2)\theta(k^0+q^0-m) \\ & \times \mathrm{Tr}[\Phi(k,P,S)\gamma^\mu(\not{k}+\not{q}+m)\gamma^\nu]\end{aligned} \tag{2.35}$$

by introducing the quark-quark correlation function Φ, defined as

$$\Phi_{ji}(k,P,S) = \sum_X \int \frac{d^3\boldsymbol{P}_X}{(2\pi)^3 2P_X^0}\langle P,S|\bar\psi_i(0)|X\rangle\langle X|\psi_j(0)|P,S\rangle\delta^{(4)}(P-k-P_X),$$

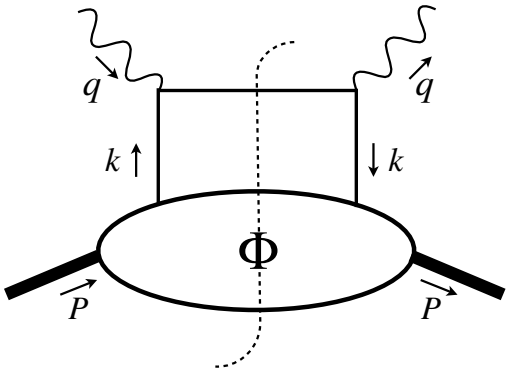


Figure 2.2: Cut diagram describing the hadronic tensor for inclusive DIS at tree level.

A graphical representation of the hadronic tensor at tree level in the parton model is given by the *handbag diagram* in Fig. 2.2. In this diagram the hard part, related to the electromagnetic interaction of the virtual photon with the struck parton (in this case a quark) is factorized from the non-perturbative quark-quark correlator. A sum over all possible final states is understood in the cut of both the on-shell quark propagator and the soft correlator Φ.

To further investigate the contributions of the different components of the quark momentum to the cross section, it is convenient to parametrize k as

$$k^\mu = \left[\frac{k^2 + |\boldsymbol{k}_\perp|^2}{2xP^+}, xP^+, \boldsymbol{k}_\perp\right] . \tag{2.36}$$

The approach that we are adopting is valid only if we assume that the virtuality of the quark k^2 and its transverse momentum squared $|\boldsymbol{k}_\perp|^2$ are small in comparison to the scale of the probe Q^2.

If we perform a $1/Q$ expansion of the complete expression we see that the dominant terms in a collinear frame are combination of plus component in the correlation function and minus component in the outgoing quark momentum. Without considering term that are $1/Q$ suppressed the final form for the hadronic tensor at leading twist is

$$2MW^{\mu\nu}(q, P, S) \simeq \sum\, e_q^2 \frac{1}{2} \,\mathrm{Tr}[\Phi^q(x_B, S)\gamma^\mu\gamma^+\gamma^\nu] \,, \tag{2.37}$$

with the integrated correlation function

$$\Phi^q_{ji}(x, S) = \int d^2\boldsymbol{k}_\perp dk^- \Phi^q_{ji}(k, P, S)\Big|_{k^+=xP^+} \tag{2.38}$$

$$= \int \frac{d\xi^-}{2\pi} e^{-ik\cdot\xi} \left\langle P, S\right| \bar\psi^q_i(\xi)\psi^q_j(0) \left|P, S\right\rangle\Big|_{\xi^+=\xi_T=0} . \tag{2.39}$$

2.1.2 Correlators and structure functions

The hadronic tensor now is expressed mainly in terms of the correlation function Φ^q. This function is a Dirac matrix and thus can be generally decomposed on a basis of Dirac terms, to analyze more precisely its structure. The terms of the decomposition are combinations of the Lorentz vectors p and P, the pseudovectors S, the vector n_- (related to the direction of the gauge link) and the Dirac structures $\mathbf{1}, \gamma_5, \gamma^\mu, \gamma^\mu\gamma_5, (i/2)[\gamma^\mu, \gamma^\nu]\gamma_5$.

Considering only the leading twist terms, the integrated correlator function in Eq. (2.38) for a parton q in the polarized case becomes [41]

$$\Phi^q(x) = f_1^q(x)\frac{\not{n}_+}{2}\,, \tag{2.40}$$

where we introduced the integrated parton distribution function

$$f_1^q(x) = \int d^2\boldsymbol{k}_\perp\, dk^2\, d(2k\cdot P)\delta(\boldsymbol{k}_\perp^2 + x^2M^2 + k^2 - 2xk\cdot P)\,[A_2^q + xA_3^q]\ . \tag{2.41}$$

The amplitudes A_i are real scalar functions $A_i = A_i(k\cdot P, k^2)$ with dimension $1/[m]^4$. If the analysis is extended to the polarized case the decomposition takes into account also combination with the S vector. The general expression of the integrated correlation function in the unpolarized case is

$$\Phi(x,S) = \frac{1}{2}\left\{ f_1\not{n}_+ + S_L g_{1L}\,\gamma_5\not{n}_+ + h_1\frac{[\not{S}_T, \not{n}_+]\gamma_5}{2}\right\}, \tag{2.42}$$

and it contains, together with $f_1(x)$ defined in eq. (2.41), the integrated parton distribution functions:

$$\begin{aligned} g_{1L}(x) = &\int d^2\boldsymbol{k}_\perp d^2k\, d(2k\cdot P)\delta\left(\boldsymbol{k}_\perp^2 + k^2 - 2xk\cdot P\right) \\ &\times\left[-A_6 - \left(\frac{k\cdot P}{M^2} - x\right)(A_7 + xA_8)\right] \end{aligned} \tag{2.43}$$

$$\begin{aligned} h_1(x) = &\int d^2\boldsymbol{k}_\perp d^2k\, d(2k\cdot P)\delta\left(\boldsymbol{k}_\perp^2 + k^2 - 2xk\cdot P\right) \\ &\times\left[-A_9 - xA_{10} + \frac{\boldsymbol{k}_\perp^2}{2M^2}A_{11}\right]. \end{aligned} \tag{2.44}$$

where the amplitudes $A_i = A_i(k\cdot P, k^2)$ are real. The function g_{1L}^q is the *helicity* distribution, while h_1 is known as the *transversity* distribution. The individual distribution functions can be isolated from $\Phi(x,S)$ through the projection

$$\Phi^{[\Gamma]} \equiv \frac{1}{2}\,\mathrm{Tr}(\Phi\Gamma), \tag{2.45}$$

where Γ denotes a specific Dirac structure.

The functions f_1 and g_{1L} in the chiral representation can be interpreted as a description of the probability to find a quark with a certain chirality inside the target, while the transversity distribution h_1 is related to the interference between two probability amplitudes of quark chirality [42, 43].

2.1.3 DIS structure functions in the parton model

The general decomposition of the correlation function Φ^q can be used to calculate the hadronic tensor of a DIS process. In the unpolarized case the correlator $\Phi(x)$ is described by Eq. (2.40); inserting this expression in Eq. (2.37) and evaluating the traces of the Dirac matrices we find an expression for the hadronic tensor, similar to Eq. (2.32). The unpolarized structure functions derived from this expression are

$$F_T = x_B \sum_q e_q^2 f_1^q(x_B) \tag{2.46}$$

$$F_L = 0\,. \tag{2.47}$$

In the polarized case to calculate the hadronic tensor we have to consider also the spin related contributions to the quark correlator, contained in Eq. (2.42). The resulting structure functions are:

$$F_1 = \frac{1}{2} \sum_q e_q^2 f_1^q(x_B)\,, \tag{2.48}$$

$$F_L = 0\,, \tag{2.49}$$

$$g_1 = \frac{1}{2} \sum_q e_q^2 g_{1L}^q(x_B)\,, \tag{2.50}$$

$$g_1 + g_2 = 0\,. \tag{2.51}$$

Once we have obtained the explicit expression of the hadronic tensor it is possible to calculate the cross section for inclusive DIS, using Eq. (2.5). The final results of the cross section can be obtained from the explicit expression of the structure functions.

2.2 Semi-inclusive DIS

Our analysis is focused on the extraction of unpolarized TMD parton distribution functions and fragmentation functions from semi-inclusive deep inelastic scattering (SIDIS), with and without target polarization, and Drell–Yan processes. As an example, we discuss SIDIS, given that its descriptions needs both unpolarized TMDs. Semi-inclusive DIS can be represented as

$$\ell(l) + N(P) \to \ell(l') + h(P_h) + X \tag{2.52}$$

where ℓ denotes the incoming lepton, N the nucleon target, h the observed hadron and X the unobserved final states, with four-momenta inside the parentheses. We introduce the kinematic variable z, defined as

$$z = \frac{P \cdot P_h}{P \cdot q}\,. \tag{2.53}$$

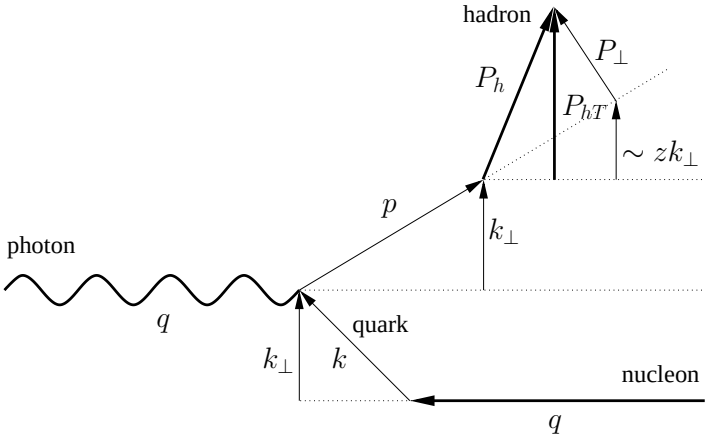


Figure 2.3: Diagram describing the relevant momenta involved in a semi-inclusive DIS event (see also Ref. [45]): a virtual photon (defining the reference axis) strikes a parton inside a proton. The parton has a transverse momentum $\boldsymbol{k}_\perp$ (not measured). The struck parton fragments into a hadron, which acquires a further transverse momentum $\boldsymbol{P}_\perp$ (not measured). The total measured transverse-momentum of the final hadron is $\boldsymbol{P}_{hT}$. When Q^2 is very large, the longitudinal components are all much larger than the transverse components. In this regime, $\boldsymbol{P}_{hT} \approx z\boldsymbol{k}_\perp + \boldsymbol{P}_\perp$.

We consider the possible polarization of beam and target, but not the eventuality of having a polarized final state; we investigate only the situation in which h has spin zero or its polarization is simply not observed. Adopting the one-photon approximation, we use the conventional variables for semi-inclusive DIS, introduced in Eq. (2.2). The masses for the nucleon target and the produced hadron h are denoted by M and M_h, respectively; we neglect contributions from lepton mass.

We define the transverse part of the hadron momentum P_h, denoted by P_{hT}, as orthogonal with respect to P and q, choosing a frame where these momenta are collinear. In a similar way, we define the transverse part $S_\perp$ of the spin vector and its longitudinal projection $S_\parallel$ along P. To study the Sivers distribution later, we have to consider also the azimuthal angles ϕ_h and ϕ_S, related to the momentum P and the spin S of the target with respect to the lepton plane. The definitions that we use in our formalism are in accordance with the Trento conventions [44] and are shown in Fig. 2.3 for the unpolarized case.

We want to express the SIDIS cross section in terms of structure functions. As a starting point, the cross section can be written as the product of a leptonic and a hadronic tensor:

$$\frac{d^6\sigma}{dx_B dy dz_h d\phi_S d^2\boldsymbol{P}_{hT}} = \frac{\alpha^2}{4zsx_B Q^2} L_{\mu\nu}\left(l, l', \lambda_e\right) 2MW^{\mu\nu}\left(q, P, S, P_h\right) \qquad (2.54)$$

The expression for the leptonic tensor is the same as in the inclusive DIS

process, while the hadronic tensor should now include also the information on the outgoing hadron h:

$$2MW^{\mu\nu}(q,P,S,P_h) = \frac{1}{(2\pi)^4}\sum_{X'}\int \frac{d^3\boldsymbol{P}_{X'}}{2P^0_{X'}} 2\pi\delta^{(4)}(q+P-P_{X'}-P_h) H^{\mu\nu}(P,S,P_{X'},P_h), \tag{2.55}$$

$$H^{\mu\nu}(P,S,P_{X'},P_h) = \langle P,S|J^\mu(0)|P_h,X'\rangle\langle P_h,X'|J^\nu(0)|P,S\rangle. \tag{2.56}$$

To decompose the hadronic tensor $M^{\mu\nu}$ we require that it satisfies Hermiticity, parity, and time reversal, as done for the inclusive process. However, we cannot impose the same condition of time-reversal invariance because, differently from the inclusive case, we should take into consideration final-state interactions on the outgoing hadron detected. This prevents the possibility of switching the initial and final state, as required by time-reversal. Instead we apply a simplified version, that consists in a simple change of sign of all vectors time components and complex conjugation:

$$L^\rho_\mu L^\sigma_\nu W^*_{\rho\sigma}(q,P,S) = W_{\mu\nu}(\tilde{q},\tilde{P},\tilde{S})\ . \tag{2.57}$$

The lepton-hadron cross section can then be parametrized in terms of 18 different structure functions [46], which depend on x_B, z, P_{hT}, Q^2. If we limit ourselves to an unpolarized target, the hadronic tensor can be parametrized in terms of 5 structure functions:

$$F_{UU,T},\ F_{UU,L},\ F_{UU}^{\cos\phi_h},\ F_{UU}^{\cos 2\phi_h},\ F_{UU}^{\sin\phi_h}\ . \tag{2.58}$$

The different structure functions F are characterized by polarization: the first two indices U, L and T refers to the polarization of the beam and the target, while a possible third index denotes the polarization of the virtual photon. Using these structure functions we can write

$$\begin{aligned} 2MW^{\mu\nu}(q,P,S) = \frac{2z}{x_B}\Big[& -g^{\mu\nu}_\perp F_{UU,T}(x_B,z,P^2_{hT},Q^2) \\ & + \hat{t}^\mu\hat{t}^\nu F_{UU,L}(x_B,z,P^2_{hT},Q^2) \\ & + (\hat{t}^\mu\hat{h}^\nu + \hat{t}^\nu\hat{h}^\mu) F_{UU}^{\cos\phi_h}(x_B,z,P^2_{hT},Q^2) \\ & + (\hat{h}^\mu\hat{h}^\mu + g^{\mu\nu}_\perp) F_{UU}^{\cos 2\phi_h}(x_B,z,P^2_{hT},Q^2) \\ & - i(\hat{t}^\mu\hat{h}^\nu - \hat{t}^\nu\hat{h}^\mu) F_{LU}^{\sin\phi_h}(x_B,z,P^2_{hT},Q^2)\Big] \end{aligned} \tag{2.59}$$

where we use the normalized vector $\hat{\boldsymbol{h}} = \boldsymbol{P}_{hT}/|\boldsymbol{P}_{hT}|$.

After the contraction of the hadronic tensor with the leptonic tensor the cross section results

$$\begin{aligned} \frac{d\sigma}{dx_B dy dz d\phi_h d\boldsymbol{P}_{hT}} = \frac{2\pi\alpha^2 y^2}{x_B y Q^2(1-\varepsilon)} \times \Big\{ & F_{UU,T} + \varepsilon\, F_{UU,L} \\ & + \sqrt{2\varepsilon(1+\varepsilon)}\cos\phi_h F_{UU}^{\cos\phi_h} + \varepsilon\cos(2\phi_h) F_{UU}^{\cos 2\phi_h} + \lambda_e\sqrt{2\varepsilon(1-\varepsilon)}\sin\phi_h F_{LU}^{\sin\phi_h}\Big\} \end{aligned} \tag{2.60}$$

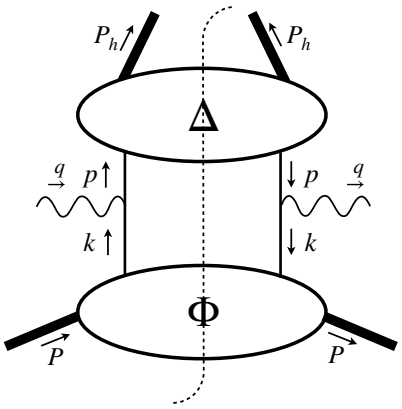


Figure 2.4: Cut diagram, describing the hadronic tensor at tree level.

The last term in Eq. (2.60) vanishes if the lepton beam is not polarized; the third and fourth term vanish if we integrate over the angle ϕ_h.

At sufficiently high Q^2 we can adopt the parton model again. Semi-inclusive DIS can be seen as a virtual photon interacting with a quark of mass m inside the nucleon; as a result the struck quark fragments producing an outgoing hadron h. The interaction dynamics of these different stages of the scattering process can be analyzed separately, it is then possible to factorize the differential cross section into two nonperturbative hadronic parts connected by a hard scattering part, as depicted in the so-called *bull diagram* in Fig. 2.4. The correlation function Φ describes the distribution of quarks, while Δ is related to quark fragmentation. The hadronic tensor at tree level can be written as

$$2MW^{\mu\nu} = 2\sum_a e_a^2 \int d^2\boldsymbol{k}_\perp d^2\frac{\boldsymbol{P}_\perp}{z}\delta^2(z\boldsymbol{k}_\perp + P_\perp - P_{hT}) \\ \times \mathrm{Tr}\left[\Phi^a(x,\boldsymbol{k}_\perp,S)\gamma^\mu\Delta^a(z,\boldsymbol{P}_\perp)\gamma^\nu\right] \tag{2.61}$$

The quark-quark distribution correlation functions at tree level is defined as [47]:

$$\Phi_{ij}(x,k_\perp,S) = \int \frac{d\xi^- d^2\xi_T}{2\pi^3} e^{ik\cdot\xi}\langle P,S|\bar{\psi}_j(0)\mathcal{U}^{n-}_{(0,+\infty)}\mathcal{U}^{n-}_{(+\infty,\xi)}\psi(\xi)|P,S\rangle\Big|_{\xi^+=0}. \tag{2.62}$$

The terms $\mathcal{U}$, called gauge links or Wilson lines, are inserted between quark fields to ensure that the definition of the correlator is gauge invariant. In principle Wilson lines should be included also in the integrated correlator in Eq. (2.38), however, they are usually neglected because in the light-cone gauge they are equal to 1. The structure of the gauge link changes for different processes and later we will discuss in more detail the differences and their observable effects. In the collinear case we do not observe any effect related to process dependence.

In the case of the fragmentation correlation we have

$$\Delta_{ij}(z, \boldsymbol{P}_\perp) = \frac{1}{2}\sum_X \int \frac{d\xi^+ d^2\xi_T}{(2\pi)^3} e^{ip\cdot\xi} \langle 0|\mathcal{U}^{n^+}_{(+\infty,\xi)}\psi_i(\xi)|h, X\rangle \times \langle h, X|\bar{\psi}_j(0)\mathcal{U}^{n^+}_{(0,+\infty)}|0\rangle\Big|_{\xi^-=0} . \tag{2.63}$$

To express the quark distribution correlator in terms of TMDs we follow a similar procedure to the one used for the inclusive case, but including this time also transverse momentum effects. The complete parametrization of the quark-quark correlation function has been given in Ref. [46]. Limiting ourselves to the leading-twist term, we get, for example:

$$\Phi(x, \boldsymbol{k}_\perp) = \left(f_1(x, \boldsymbol{k}_\perp^2) + i h_1^\perp(x, \boldsymbol{k}_\perp^2) \frac{\not{k}_\perp}{M} \right) \frac{\not{n}_+}{2} , \tag{2.64}$$

which contains the unpolarized TMD parton distribution function $f_1(x, \boldsymbol{k}_\perp^2)$, which will be the object of our extraction in the next chapter, and the *Boer-Mulders* TMD $h_1^\perp(x, \boldsymbol{k}_\perp^2)$ for transversely polarized quarks [17]. The relation between quark correlators and structure function can be found through projections of the type $\Phi^{[\Gamma]}$, where Γ is a Dirac structure function. As an example, if we consider $\Phi^{[\not{n}]}$ we can obtain the unpolarized $f_1^a(x, \boldsymbol{k}_\perp^2)$.

The same procedure can be applied, with opportune modifications, to the fragmentation correlator, to get

$$\Delta(z, \boldsymbol{P}_\perp^2) = \left(D_1(z, \boldsymbol{P}_\perp^2) + i H_1^\perp(z, \boldsymbol{P}_\perp^2) \frac{\not{P}_\perp}{M_h} \right) \frac{\not{n}_-}{2} . \tag{2.65}$$

The function $D_1(z, \boldsymbol{P}_\perp^2)$ is the unpolarized TMD fragmentation function, while $H_1^\perp(z, \boldsymbol{P}_\perp^2)$ is called *Collins function* [37].

Once we have the parametrization for the different correlators we can write explicitly the hadronic tensor in Eq. (2.61) and calculate the leptoproduction cross section. At this point we can express the structure functions contained in the expression of the cross section in terms of TMDs. The unpolarized cross section for SIDIS, integrated over the azimuthal angle, becomes

$$\frac{d\sigma}{dx dy dz d\boldsymbol{P}_{hT}^2} = \frac{4\pi^2\alpha^2}{xQ^2}\frac{y}{2(1-\varepsilon)} \left(F_{UU,T}(x, z, \boldsymbol{P}_{hT}^2, Q^2) + \varepsilon F_{UU,L}(x, z, \boldsymbol{P}_{hT}^2, Q^2) \right) . \tag{2.66}$$

The structure functions $F_{UU,T}$ and $F_{UU,L}$ can be written explicitly as

$$F_{UU,T} = \sum_a e_a^2 \mathcal{C}\left[f_1^a D_1^a \right] , \qquad F_{UU,L} = \mathcal{O}\left(\frac{M^2}{Q^2}, \frac{\boldsymbol{P}_{hT}^2}{Q^2} \right) \tag{2.67}$$

with a compact notation for the convolution

$$\mathcal{C}\left[f D \right] = x_B \int d^2\boldsymbol{k}_\perp \frac{d^2 P_\perp}{z} \delta^{(2)}(z\boldsymbol{k}_\perp - \boldsymbol{P}_{hT} + \boldsymbol{P}_\perp) f^a(x_B, \boldsymbol{k}_\perp^2) D^a(z, \boldsymbol{P}_\perp^2) . \tag{2.68}$$

In order to later discuss the formalism for the evolution of TMDs, which is developed in the space conjugate to the momentum, it is useful to calculate the Fourier transform of this convolution. The structure function in this case reduces to [48]:

$$\begin{aligned} &F_{UU,T}(x, z, \boldsymbol{P}_{hT}^2, Q^2) \\ &\quad \approx 2\pi \sum_a e_a^2 x \int_0^\infty d\xi_T \xi_T J_0\big(\xi_T |\boldsymbol{P}_{hT}|/z\big) \tilde{f}_1^a\big(x, \xi_T^2; Q^2\big) \tilde{D}_1^{a\to h}\big(z, \xi_T^2; Q^2\big)\,, \end{aligned} \tag{2.69}$$

where we introduced the Fourier transforms of the TMD PDF and FF:

$$\begin{aligned} f_1(x, \boldsymbol{k}_\perp^2; Q^2) =& \frac{1}{2\pi} \int_{-\infty}^{\infty} d\xi_T^2 e^{i\xi_T \cdot \boldsymbol{k}_\perp^2} \tilde{f}_1(x, \xi_T^2; Q^2) \\ =& \frac{1}{2\pi} \int_0^{2\pi} \int_{-\infty}^{+\infty} d\phi \frac{d^2\xi_T}{2} e^{i|\xi_T||\boldsymbol{k}_\perp|\cos\phi} \tilde{f}_1(x, \xi_T^2; Q^2)\,. \end{aligned} \tag{2.70}$$

Using the definition of the Bessel function of the first kind

$$J_0(z) = \frac{1}{2} \int_0^{2\pi} e^{iz\cos\theta} d\theta \tag{2.71}$$

we can express the Fourier transform of the distribution function

$$\tilde{f}_1^a\big(x, \xi_T^2; Q^2\big) = \int_0^\infty d|\boldsymbol{k}_\perp| |\boldsymbol{k}_\perp| J_0\big(\xi_T |\boldsymbol{k}_\perp|\big) f_1^a\big(x, \boldsymbol{k}_\perp^2; Q^2\big), \tag{2.72}$$

and, in a similar way, the TMD fragmentation function

$$\tilde{D}_1^{a\to h}\big(z, \xi_T^2; Q^2\big) = \int_0^\infty \frac{d|\boldsymbol{P}_\perp|}{z^2} |\boldsymbol{P}_\perp| J_0\big(\xi_T |\boldsymbol{P}_\perp|/z\big) D_1^{a\to h}\big(z, \boldsymbol{P}_\perp^2; Q^2\big)\,. \tag{2.73}$$

2.3 Drell–Yan

In our global extraction, in addition to the semi-inclusive DIS unpolarized data, we include also measurements of Drell–Yan processes (DY). We consider separately Drell–Yan processes with a virtual photon or with Z boson production in the intermediate state. One of the most interesting feature of TMD distributions is that their description of the partonic internal structure does not depend on the specific process considered, and should be able to give an equally good description for both SIDIS and Drell–Yan processes. In the case of TMDs the study of universality becomes more nuanced, because in some cases, such as the sign change for the Sivers function, it is not completely respected.

In the following discussion we will derive the cross section in terms of TMDs of dilepton production from high-energy scattering of two hadrons

$$A\left(P_A, S_A\right) + B\left(P_B, S_B\right) \to \ell^-(l, \lambda) + \ell^+\left(l', \lambda'\right) + X \tag{2.74}$$

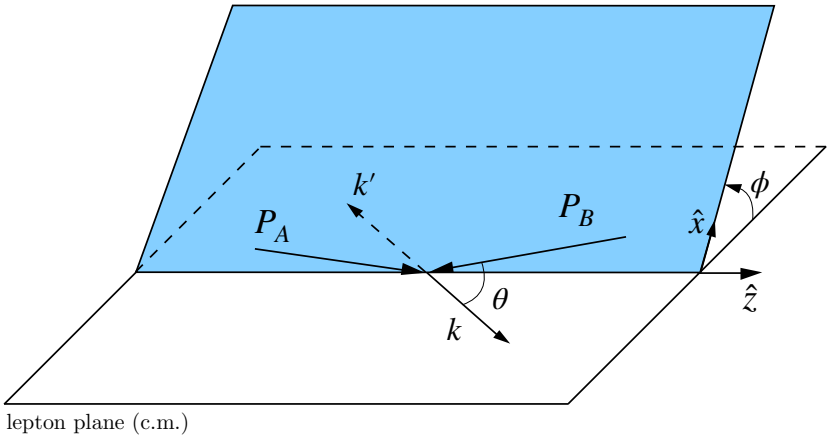


Figure 2.5: Kinematics of Drell–Yan process in the dilepton rest frame. P_A, B denote the momenta of the nucleons in the initial state, while k and $k' = q - k$ are the momenta of the produced leptons.

where $(P_{A,B}, S_{A,B})$ denotes the four-momenta and spin vectors of the initial state hadrons and λ, λ' are the helicities of the produced leptons [49]. The momenta and spin vectors of the two hadrons follow the relation $P^2_{A,B} = M^2_{A,B}$, $P_{A,B} \cdot S_{A,B} = 0$ and $S^2_A = -1$. The mass of the leptons are neglected also in this case. The quantities used to describe the kinematic of the DY process are represented in Fig. 2.5.

For a sufficiently high value of the dilepton invariant mass we can adopt the parton model framework. In this approach a quark from the hadron A and an antiquark from hadron B annihilate and produce a virtual photon in the intermediate state, which consequently decays into an e^+e^- pair; these steps can be schematized as

$$A + B \to \gamma^*(q) + X \to \ell^- + \ell^+ + X \tag{2.75}$$

where $q = l + l'$ again indicates the virtuality of the photon.

Adopting the one-photon approximation the cross section of the Drell–Yan processes can be written in a frame-independent way as

$$\frac{l^0 l'^0 d\sigma}{d^3\boldsymbol{l} d^3\boldsymbol{l'}} = \frac{\alpha^2_{em}}{F q^4} L_{\mu\nu} W^{\mu\nu} \tag{2.76}$$

with the flux F of the initial hadrons defined as

$$F = 4\sqrt{(P_A \cdot P_B)^2 - M^2_A M^2_B}\,. \tag{2.77}$$

To gather information about the internal structure of nucleons, we need to study the angular distribution of the DY cross section. The most convenient frame to conduct this analysis is a dilepton rest frame, such as the Collins–Soper frame [50]. In this kind of framework the differential cross section becomes

$$\frac{d\sigma}{d^4 q d\Omega} = \frac{\alpha^2_{em}}{2F q^4} L_{\mu\nu} W^{\mu\nu}\,. \tag{2.78}$$

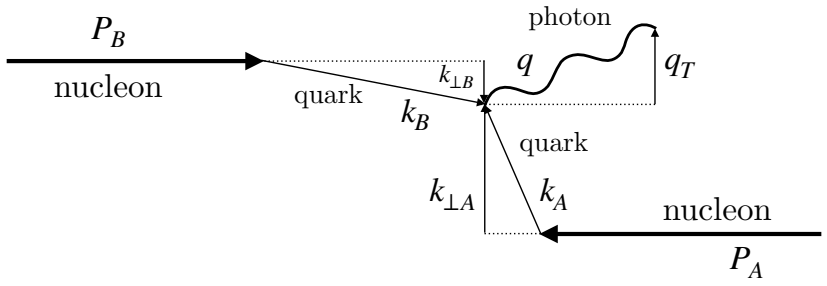


Figure 2.6: Diagram describing the relevant momenta involved in a Drell–Yan event: two partons from the nucleons A and B collide. They have transverse momenta $\boldsymbol{k}_{\perp A}$ and $\boldsymbol{k}_{\perp B}$ (not measured) and produce a virtual photon with (measured) transverse momentum $\boldsymbol{q}_T = \boldsymbol{k}_{\perp A} + \boldsymbol{k}_{\perp B}$ with respect to the hadron collision axis.

where the solid angle Ω determines the orientation of the produced leptons.

The spin-averaged leptonic tensor $L^{\mu\nu}$ is written explicitly as

$$L^{\mu\nu} = \sum_{\lambda,\lambda'} \left(\overline{u}(l,\lambda)\gamma^\mu \nu\left(l',\lambda'\right)\right)\left(\overline{u}(l,\lambda)\gamma^\nu \nu\left(l',\lambda'\right)\right)^* = 4\left(l^\mu l'^\nu + l^\nu l^\mu - \frac{q^2}{2} g^{\mu\nu}\right) \tag{2.79}$$

and the hadronic tensor $W^{\mu\nu}$ in the cross section reads

$$\begin{aligned} &W^{\mu\nu}\left(P_A, S_A; P_B, S_B; q\right) \\ &\quad = \frac{1}{(2\pi)^4} \int d^4x e^{iq\cdot x} \left\langle P_A, S_A; P_B, S_B \left| J^\nu_{em}(0) J^\nu_{em}(x) \right| P_A, S_A; P_B, S_B \right\rangle . \end{aligned} \tag{2.80}$$

As we already discussed, the hadronic tensor naturally follows the requirements of electromagnetic gauge invariance, parity and Hermiticity. Exploiting these properties is possible to write the hadronic tensor in terms of structure functions. The full decomposition of the hadronic tensor $W^{\mu\nu}$ in terms of structure functions can be found in Ref. [49].

In the unpolarized case the hadronic tensor depends on the vectors q^μ, P^μ_A, and P^μ_B, shown in Fig. 2.6. Using the possible combination of these four-vectors one can create a basis of seven independent tensors that satisfy the parity property. Through this tensor basis it is now possible to decompose the hadronic tensor as

$$W^{\mu\nu}_i = \sum_{i=1}^{7} h^{\mu\nu}_{i,j} \tilde{V}_{i,j} \tag{2.81}$$

where $\tilde{V}_{i,j}$ are structure functions that depend on the invariants q^2 and $P_{A,B}\cdot q$. If we apply the gauge invariance constraints, we observe that only four of the

structure functions $\tilde{V}_{i,j}$ are independent. It is useful, then, to redefine the tensor basis using the projection operators:

$$P^{\mu\nu} = g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2} \tag{2.82}$$

They are constructed so that $q_\mu P^{\mu\nu} = P^{\mu\nu} q_\nu = 0$. When the operator P acts on the tensor basis, according to $P^\mu_\rho h^{\rho\sigma}_{i,j} P^\nu_\sigma$, we are left with the appropriate basis of four tensors, gauge invariant by construction. The hadronic tensor decomposed using this basis becomes

$$W^{\mu\nu}_i = \sum_{j=1}^{4} t^{\mu\nu}_{i,j} V_{i,j} \,. \tag{2.83}$$

The tensor $W^{\mu\nu}$ in Eq. (2.81) is frame independent and alternative forms can be used equivalently.

Using this general form for the hadronic tensor, we can obtain the DY cross section in terms of structure functions related to the angular distribution of the particles involved in the process. This can be done in any reference frame in principle, because the hadronic tensor is frame independent; we adopt the CS frame, where the angular distribution assumes the most compact form. Contracting the leptonic tensor in Eq. (2.79) with the hadronic tensor one obtains the following expression for Eq. (2.78) :

$$\begin{aligned}\frac{d\sigma}{d^4q d\Omega} = \frac{\alpha^2_{em}}{Fq^2} \times \Big[&\left(1 + \cos^2\theta\right) F^1_{UU} + \left(1 - \cos^2\theta\right) F^2_{UU} \\ &+ \sin 2\theta \cos\phi F^{\cos\phi}_{UU} + \sin^2\theta \cos 2\phi F^{\cos 2\phi}_{UU} \Big] \,.\end{aligned} \tag{2.84}$$

where ϕ is the angle between the collision and the lepton plane, and θ is the angle between the direction $\hat{z}$ of the beam and the momentum of the outgoing lepton l (as shown in Fig. 2.5).

In the kinematic region where the order the transverse photon momentum q_T is much smaller than the hard scale Q, the DY process can be described in terms of TMD functions. At leading twist, i.e., considering the leading order of an expansion in terms of $1/Q$, the hadronic part of the cross section can be written in terms of unintegrated quark-quark correlators, as in

$$\begin{aligned}W^{\mu\nu} = \frac{1}{N_c} \sum_q e_q^2 \int d^4k_A d^4k_B \delta^{(4)}\left(q - k_A - k_B\right) \\ \times \mathrm{Tr}\left[\gamma^\mu \Phi^q\left(k_A, P_A, S_A\right) \gamma^\nu \bar{\Phi}^q\left(k_B, P_B, S_B\right)\right] + \{\Phi \leftrightarrow \bar{\Phi}\}\end{aligned} \tag{2.85}$$

where the correlator for a specific hadron A is defined in a way similar to the SIDIS case. The last term $\{\Phi \leftrightarrow \bar{\Phi}\}$ of Eq. (2.85) represents the contributions of the DY process diagrams where the hadrons A and B switch their roles in the annihilation, with A exchanging an antiquark instead of a quark and

viceversa for B; formally it is obtained interchanging the Φ correlators that appear in the first term.

Adopting the light-cone vectors, we observe that the initial partons in a parton model are assumed to move approximately along the direction of their parent hadron, consequently we can choose a frame where the hadron A has a large component k_A^+ and the hadron B instead has a large component k_B^- . Considering only the leading terms the hadronic tensor then reduces to

$$\begin{aligned} W^{\mu\nu} =& \frac{1}{N_c}\sum_q e_q^2 \int d^2\boldsymbol{k}_{\perp A} d^2\boldsymbol{k}_{\perp B}\delta^{(2)}\left(\boldsymbol{q}_T - \boldsymbol{k}_{\perp A} - \boldsymbol{k}_{\perp B}\right) \\ & \times \mathrm{Tr}\left[\gamma^\mu \Phi^q\left(x_A, \boldsymbol{k}_{\perp A}, S_A\right)\gamma^\nu \bar{\Phi}^q\left(x_B, \boldsymbol{k}_{\perp B}, S_B\right)\right] + \left\{\Phi \leftrightarrow \bar{\Phi}\right\} \end{aligned} \tag{2.86}$$

with the usual DY fractional momenta variables

$$x_A = \frac{q^2}{2P_A\cdot q} \approx \frac{k_A^+}{P_A^+}, \qquad\qquad x_B = \frac{q^2}{2P_B\cdot q} \approx \frac{k_B^-}{P_B^-}\,. \tag{2.87}$$

and the quark correlators is analogous to Eq. (2.62), with a different gauge link $\mathcal{W}_{\mathrm{DY}}$. It can be written as

$$\begin{aligned} \Phi_{ij}^q\left(x_A, \boldsymbol{k}_{\perp A}, S_A\right) =& \int \frac{d\xi^- d^2\xi_T}{(2\pi)^3} e^{ik_A\cdot\xi} \\ & \times \langle P_A, S_A|\bar{\psi}_j^q(0)\mathcal{U}_{(0,-\infty)}^{n_-}\mathcal{U}_{(-\infty,\xi)}^{n_-}\psi_i^q(\xi)|P_A, S_A\rangle|_{\xi^+=0}\,. \end{aligned} \tag{2.88}$$

The quark-quark correlator in the kinematic region that we discussed, can be parametrized through TMD parton distribution functions [15] calculated through the traces $\Phi^{[\Gamma]}$. The traces for the antiquarks can be treated in the same way, however, since we have chosen a collinear framework, the role of the plus and minus components in h_B have to be interchanged compared to h_A, consequently producing an opposite sign, which appears in front of ε_T^{ij} in the traces $\bar{\Phi}^{[\Gamma]}$.

Inserting them in the hadronic tensor (2.86) it is now possible to calculate the leading-twist observables. For the unpolarized case in the CS frame we find

$$F_{UU}^1 = \mathcal{C}\left[f_1\bar{f}_1\right], \tag{2.89}$$

$$F_{UU}^2 = \frac{1}{2}\mathcal{C}\left[\frac{2\left(\boldsymbol{h}\cdot\boldsymbol{k}_{\perp A}\right)\left(\boldsymbol{h}\cdot\boldsymbol{k}_{\perp B}\right) - \boldsymbol{k}_{\perp A}\cdot\boldsymbol{k}_{\perp B}}{M_A M_B} h_1^\perp \bar{h}_1^\perp\right] \tag{2.90}$$

where we made use of the following notation for the convolution of TMDs in momentum space:

$$\begin{aligned} \mathcal{C}_w\left[w\left(\boldsymbol{k}_{\perp A}, \boldsymbol{k}_{\perp B}\right) f_1\overline{f}_2\right] \equiv& \frac{1}{N_c}\sum_q e_q^2 \int d^2\boldsymbol{k}_{\perp A} d^2\boldsymbol{k}_{\perp B}\delta^{(2)}\left(\boldsymbol{q}_T - \boldsymbol{k}_{\perp A} - \boldsymbol{k}_{\perp B}\right) \\ \times w\left(\boldsymbol{k}_{\perp A}, \boldsymbol{k}_{\perp B}\right)& \left[f_1^q\left(x_A, \boldsymbol{k}_{\perp A}^2\right) f_2^{\bar{q}}\left(x_B, \boldsymbol{k}_{\perp B}^2\right) + f_1^{\bar{q}}\left(x_A, \boldsymbol{k}_{\perp A}^2\right) f_2^q\left(x_B, \boldsymbol{k}_{\perp B}^2\right)\right]. \end{aligned} \tag{2.91}$$

Also in the case of DY processes the structure function can be conveniently expressed as a Fourier transform:

$$F^1_{UU}(x_A, x_B, \boldsymbol{q}_T^2, Q^2) \approx 2\pi \sum_a \mathcal{H}^{1a}_{UU}\, x_a x_B \times \int_0^\infty d\xi_T \xi_T\, J_0\big(\xi_T |\boldsymbol{q}_T|\big) \tilde{f}_1^a\big(x_A, \xi_T^2; Q^2\big)\ \tilde{f}_1^{\bar{a}}\big(x_B, \xi_T^2; Q^2\big)\,, \quad (2.92)$$

where we introduced a hard coefficient $\mathcal{H}$ in order to be able to use the same formula for DY and Z boson production. In fact, the hard coefficients $\mathcal{H}$, in the case of DY processes and Z boson production, up to leading order in the couplings, are

$$\mathcal{H}^{1a}_{UU,\gamma}(Q^2) \approx \frac{e_a^2}{N_c}, \qquad \mathcal{H}^{1a}_{UU,Z}(Q^2) \approx \frac{V_a^2 + A_a^2}{N_c}\ , \quad (2.93)$$

where, using the value of weak isospin I_3 (equal to $+1/2$ for u, c, t and $-1/2$ for d, s and b) we have

$$V_a = I_{3a} - 2e_a \sin^2\theta_W\ , \qquad A_a = I_{3a}\ . \quad (2.94)$$

In Section 3.2.2 we discuss the choices that we adopted for the parametrization of the TMDs in the calculation of the DY cross section.

2.4 Sivers distribution function in leptoproduction

In this section we will study the cross sections for semi-inclusive DIS, including also the effects of beam, target and parton polarization. The complete discussion of SIDIS at small transverse momentum is presented in Ref. [46], which expands the work presented in Refs [15, 17, 51].

Working in the target rest frame, we use the same definitions of Eq. (2.2) for the kinematic variables and introduce the azimuthal angle ϕ_h of the outgoing hadron [44] through its projections:

$$\cos\phi_h = -\frac{l_\mu P_{h\nu} g_\perp^{\mu\nu}}{\sqrt{l_\perp^2 P_{hT}^2}}\,, \quad \sin\phi_h = -\frac{l_\mu P_{h\nu} \epsilon_\perp^{\mu\nu}}{\sqrt{l_\perp^2 P_{hT}^2}}\,. \quad (2.95)$$

It is possible to decompose the spin vector S of the target as

$$S^\mu = S_\parallel \frac{P^\mu - q^\mu M^2/(P\cdot q)}{M\sqrt{1+\gamma^2}} + S_\perp^\mu, \quad S_\parallel = \frac{S\cdot q}{P\cdot q}\frac{M}{\sqrt{1+\gamma^2}}, \quad S_\perp^\mu = g_\perp^{\mu\nu} S_\nu\,. \quad (2.96)$$

We introduce the spin azimuthal angle ϕ_S, in analogy to the definition of ϕ_h, simply substituting P_h with S in Eq. (2.95).

In one-photon exchange approximation, the cross section for semi-inclusive DIS can be expressed by a set of structure functions [52], dependent on x, Q^2,z and P_{hT}^2, as

$$\begin{aligned}
&\frac{d\sigma}{dxdyd\psi dzd\phi_h dP_{h\perp}^2} = \\
&\quad \frac{\alpha^2}{xyQ^2}\frac{y^2}{2(1-\varepsilon)}\left(1+\frac{\gamma^2}{2x}\right)\Big\{F_{UU,T} + \varepsilon F_{UU,L} + \sqrt{2\varepsilon(1+\varepsilon)}\cos\phi_h F_{UU}^{\cos\phi_h} \\
&+ \varepsilon\cos(2\phi_h)\, F_{UU}^{\cos 2\phi_h} + \lambda_e\sqrt{2\varepsilon(1-\varepsilon)}\sin\phi_h F_{LU}^{\sin\phi_h} \\
&+ S_\parallel\left[\sqrt{2\varepsilon(1+\varepsilon)}\sin\phi_h F_{UL}^{\sin\phi_h} + \varepsilon\sin(2\phi_h)\, F_{UL}^{\sin 2\phi_h}\right] \\
&+ S_\parallel\lambda_e\left[\sqrt{1-\varepsilon^2}F_{LL} + \sqrt{2\varepsilon(1-\varepsilon)}\cos\phi_h F_{LL}^{\cos\phi_h}\right] \\
&+ |\boldsymbol{S}_\perp|\Bigg[\sin(\phi_h-\phi_S)\left(F_{UT,T}^{\sin(\phi_h-\phi_S)} + \varepsilon F_{UT,L}^{\sin(\phi_h-\phi_S)}\right) \\
&+ \varepsilon\sin(\phi_h+\phi_S)\, F_{UT}^{\sin(\phi_h+\phi_S)} + \varepsilon\sin(3\phi_h-\phi_S)\, F_{UT}^{\sin(3\phi_h-\phi_S)} \\
&+ \sqrt{2\varepsilon(1+\varepsilon)}\sin\phi_S F_{UT}^{\sin\phi_S} + \sqrt{2\varepsilon(1+\varepsilon)}\sin(2\phi_h-\phi_S)\, F_{UT}^{\sin(2\phi_h-\phi_S)}\Bigg] \\
&+ |\boldsymbol{S}_\perp|\,\lambda_e\Bigg[\sqrt{1-\varepsilon^2}\cos(\phi_h-\phi_S)\, F_{LT}^{\cos(\phi_h-\phi_S)} + \sqrt{2\varepsilon(1-\varepsilon)}\cos\phi_S F_{LT}^{\cos\phi_S} \\
&+ \sqrt{2\varepsilon(1-\varepsilon)}\cos(2\phi_h-\phi_S)\, F_{LT}^{\cos(2\phi_h-\phi_S)}\Bigg]\Big\},
\end{aligned} \tag{2.97}$$

where we used the previous meaning for the indices of the functions $F_{XY,Z}$. The cross section is differential in ψ, which is the azimuthal angle of ℓ' on the lepton beam axis with respect to an arbitrary direction, usually the direction of S in the case of a transversely polarized target; in deep kinematics we can use $d\psi \approx d\phi_S$.

The relation to the collinear semi-inclusive cross section can be directly obtained integrating Eq. (2.97) over the transverse hadron momentum $\boldsymbol{P}_{h_\perp}$:

$$\begin{aligned}
\frac{d\sigma}{dxdyd\psi dz} =& \frac{2\alpha^2}{xyQ^2}\frac{y^2}{2(1-\varepsilon)}\left(1+\frac{\gamma^2}{2x}\right)\{F_{UU,T} + \varepsilon F_{UU,L} + S_\parallel\lambda_e\sqrt{1-\varepsilon^2}F_{LL} \\
&+ |S_\perp|\sqrt{2\varepsilon(1+\varepsilon)}\sin\phi_S F_{UT}^{\sin\phi_S} + |S_\perp|\,\lambda_e\sqrt{2\varepsilon(1-\varepsilon)}\cos\phi_S F_{LT}^{\cos\phi_S}\}
\end{aligned} \tag{2.98}$$

where the relation with the previous unintegrated structure functions is

$$F_{UU,T}\left(x,z,Q^2\right) = \int d^2\boldsymbol{P}_{hT} F_{UU,T}\left(x,z,P_{hT}^2,Q^2\right). \tag{2.99}$$

To explicit the relation between the structure functions with polarization and the parton distribution and fragmentation functions, we need to calculate the leptoproduction cross section, expressed as the product of a leptonic and a hadronic tensor:

$$\frac{d\sigma}{dxdyd\psi dzd\phi_h dP_{hT}^2} = \frac{\alpha^2 y}{8zQ^4} 2MW^{\mu\nu} L_{\mu\nu}\,, \tag{2.100}$$

where the leptonic and hadronic tensor have been already discussed in the previous sections.

The transverse direction is defined with respect to the momenta of the target and of the produced hadron, and the photon acquires a transverse momentum q_T. In some occasion, it is more convenient to use a frame where the he initial quark and the virtual photon are collinear and the transverse direction is defined in relation to them and the final hadron has a transverse momentum P_{hT}. At leading twist the two frames are related by $q_T = -P_{hT}/z$. The term Δ or Φ that appears in the hadronic tensor, expressed as in Eq. (2.61), are related to the quark distributions and quark fragmentation, respectively. Focusing on the Sivers function, we will discuss more in detail the quark-quark correlation function Φ, defined in Eq. (2.62).

If we consider also target polarization, neglected until now, the leading part of the transverse momentum dependent correlator becomes [15]

$$\begin{aligned}\Phi\left(x, \boldsymbol{k}_\perp\right) = \frac{1}{2}\Bigg\{ & f_1 \not{n}_+ + f_{1T}^\perp \frac{\epsilon_T^{\rho\sigma} S_{T\rho} k_{\perp\sigma}}{M} \not{n}_+ + g_{1S}\gamma_5 \not{n}_+ + h_{1T} \frac{\left[\not{S}_T, \not{n}_+\right]\gamma_5}{2} \\ & + h_{1s}^\perp \frac{\left[\not{k}_\perp, \not{n}_+\right]\gamma_5}{2M} + \mathrm{i} h_1^\perp \frac{\left[\not{k}_\perp, \not{n}_+\right]}{2M} \Bigg\}\end{aligned} \tag{2.101}$$

where we used $\epsilon_T^{\alpha\beta} = \epsilon^{\alpha\beta\rho\sigma} n_{+\rho} n_{-\sigma}$ and we introduced the following shorthand notation for the spin-dependent function (i.e., the one with an index S):

$$g_{1S}\left(x, \boldsymbol{k}_\perp\right) = S_L g_{1L}\left(x, \boldsymbol{k}_\perp^2\right) - \frac{\boldsymbol{S}_T \cdot \boldsymbol{k}_\perp}{M} g_{1T}\left(x, \boldsymbol{k}_\perp^2\right) \tag{2.102}$$

It is useful define combination of the original functions which can later related to experimental observables, as in the case of the transversity function, which is defined as

$$h_1(x, \boldsymbol{k}_\perp^2) \equiv h_{1T}(x, \boldsymbol{k}_\perp^2) + h_{1T}^{\perp(1)}(x, \boldsymbol{k}_\perp^2)\,. \tag{2.103}$$

We can define eight different TMD distribution functions, related to the polarization of the nucleon target and of the struck quark.

The relation between quark correlators and structure function can again be found through projections of the type $\Phi^{[\Gamma]}$, where Γ is a Dirac matrix $\gamma^\mu, \gamma^\mu\gamma_5, i\sigma^{\mu\nu}\gamma_5, 1, i\gamma_5$.

The explicit expression for all the projections that include twist–2 TMDs are

$$\Phi^{a\left[\gamma^{+}\right]} = f_1^a\left(x, \boldsymbol{k}_\perp^2\right) - \frac{\varepsilon_T^{ij} k_\perp^i S_\perp^j}{M} f_{1T}^{\perp a}\left(x, \boldsymbol{k}_\perp^2\right) , \tag{2.104}$$

$$\Phi^{a\left[\gamma^{+}\gamma_5\right]} = S_L g_{1L}^q\left(x_A, \boldsymbol{k}_\perp^2\right) + \frac{\boldsymbol{k}_\perp \cdot \boldsymbol{S}_\perp}{M} g_{1T}^a\left(x, \boldsymbol{k}_\perp^2\right) , \tag{2.105}$$

$$\begin{aligned}\Phi^{a\left[i\sigma^{i+}\gamma_5\right]} =& S_\perp^i h_1^a\left(x, \boldsymbol{k}_\perp^2\right) + \frac{k_\perp^i \left(\boldsymbol{k}_\perp \cdot \boldsymbol{S}_\perp\right) - \frac{1}{2}\boldsymbol{k}_{\perp A}^2 S_\perp^i}{M^2} h_{1T}^{\perp a}\left(x, \boldsymbol{k}_\perp^2\right) \\ &+ S_L \frac{k_\perp^i}{M} h_{1L}^{\perp a}\left(x, \boldsymbol{k}_\perp^2\right) + \frac{\varepsilon_T^{ij} k_\perp^j}{M} h_1^{\perp a}\left(x, \boldsymbol{k}_\perp^2\right) .\end{aligned} \tag{2.106}$$

Transverse momentum dependent distributions of leading twist can be interpreted as number densities, for example if we consider the distribution of unpolarized quarks in a polarized proton:

$$\begin{aligned}f_{a/p^\uparrow}\left(x, \boldsymbol{k}_\perp\right) =& \Phi^{\left[\not{n}_-\right]} = f_1^a\left(x, \boldsymbol{k}_\perp^2\right) - f_{1T}^{\perp a}\left(x, p_T^2\right) \frac{\epsilon^{\mu\nu\rho\sigma} P_\mu p_\nu S_\rho \left(n_-\right)_\sigma}{M\left(P \cdot n_-\right)} \\ =& f_1^a\left(x, \boldsymbol{k}_\perp^2\right) - f_{1T}^{\perp a}\left(x, \boldsymbol{k}_\perp^2\right) \frac{\left(\hat{\boldsymbol{P}} \times \boldsymbol{k}_\perp\right) \cdot \boldsymbol{S}}{M}\end{aligned} . \tag{2.107}$$

The second expression is valid in any frame where $\boldsymbol{n}$ and the direction $\hat{\boldsymbol{P}}$ of the proton momentum point into opposite directions. Therefore $f_{1T}^{\perp a} > 0$ corresponds to a preference of the quark to move to the left if the proton is moving towards the observer and the proton spin is pointing upwards.

The usual TMD structure function and their probabilistic interpretation, related to nucleon target and quark polarization was summarized in Fig. 1.1.

If we explicit the expressions of the different correlators in the hadronic tensor we can write the structure function which appears in Eq. (2.97) as convolutions of PDF and FF. In particular, the structure function $F_{UT,T}^{\sin(\phi_h - \phi_S)}$ contains the Sivers function:

$$F_{UT,T}^{\sin(\phi_h - \phi_S)} = \mathcal{C}\left[-\frac{\hat{\boldsymbol{h}} \cdot \boldsymbol{k}_\perp}{M} f_{1T}^\perp D_1\right] . \tag{2.108}$$

We introduce also the corresponding expression for the structure function in terms of the Fourier transform of the first moment of the Sivers function $f_{1T}^{\perp(1)a}$ and of the FF:

$$\begin{aligned}F_{UT,T}^{\sin(\phi_h - \phi_S)}\left(x, z, \boldsymbol{P}_{hT}^2, Q^2\right) =& -2\pi M \sum_a e_a^2 x \int_0^\infty d\xi_T^2 \xi_T^2 J_1\left(\xi_T \left|\boldsymbol{P}_{hT}\right| / z\right) \\ &\times \tilde{f}_{1T}^{\perp(1)a}\left(x, \xi_T^2; Q^2\right) \tilde{D}_1^{a\to h}\left(z, \xi_T^2; Q^2\right) .\end{aligned} \tag{2.109}$$

(a) (b)

Figure 2.7: Gauge link for semi-inclusive DIS (a) and Drell–Yan (b)

2.4.1 Time reversal of Sivers function

An interesting feature of the Sivers function $f_{1T}^{\perp}(x, k_\perp^2)$ and the Boer-Mulders function $h_1^{\perp}(x, k_\perp)$ is to be *time-reversal odd*, or *T-odd* for short. This means that when a naive time-reversal transformation is applied to the Sivers distribution the resulting function has an opposite sign. At first, this seems to suggest that the function is its own opposite, thus excluding the possibility of having a non-zero value. However, we are considering a gauge theory where Wilson lines are present and the PT transformation changes them from the past-pointing to the future-pointing direction, introducing a different term for each functions and thus avoiding their cancellation.

In the case of Drell–Yan processes we have to include past-pointing Wilson lines, opposed to the gauge links introduced in the definition of the SIDIS polarized cross section [40].

As we observed, to preserve gauge invariance in the expression of the hadronic tensor it is necessary to introduce gauge link $\mathcal{U}$, defined as

$$\mathcal{U}^{n_-}_{(0,+\infty)} = \mathcal{U}^{n_-}\left(0^-, \infty^-; \mathbf{0}_T\right) \mathcal{U}^{T}\left(\mathbf{0}_T, \boldsymbol{\infty}_T; \infty^-\right) \tag{2.110}$$

$$\mathcal{U}^{n_-}_{(+\infty,\xi)} = \mathcal{U}^{T}\left(\boldsymbol{\infty}_T, \boldsymbol{\xi}_T; \infty^-\right) \mathcal{U}^{n_-}\left(\infty^-, \xi^-, \xi_T\right) , \tag{2.111}$$

where $\mathcal{U}^{n_-}\left(a^-, b^-; \mathbf{c}_T\right)$ denotes a gauge link going from $[a^-, 0^+, \boldsymbol{c}_T]$ to $[c^-, 0^+, \boldsymbol{b}_T]$ along the minus direction. Instead, $\mathcal{U}^{T}(\boldsymbol{a}_T, \boldsymbol{b}_T; c^-)$ indicates a Wilson line running in the transverse direction. Their explicit expression is

$$\mathcal{U}^{n_-}\left(a^-, b^-; \boldsymbol{c}_T\right) = \mathcal{P} \exp\left[-ig \int_{a^-}^{b^-} \mathrm{d}\eta^- \mathrm{A}^+\left(\eta^-, 0, \boldsymbol{c}_T\right)\right] \tag{2.112}$$

$$\mathcal{U}^{T}\left(\boldsymbol{a}_T, \boldsymbol{b}_T; c^-\right) = \mathcal{P} \exp\left[-ig \int_{a_T}^{b_T} \mathrm{d}\eta_T \cdot A_T\left(c^-, 0, \boldsymbol{\eta}_T\right)\right] \tag{2.113}$$

The graphical representation of the gauge link for the SIDIS and DY case are shown in Fig. 2.7a and Fig. 2.7b, respectively. The gauge link structure can change for different processes [53, 54]. The gauge link can be calculated, choosing a specific model, considering the contributions at leading twist of diagrams similar to the ones shown in Fig. 2.8 and their conjugates.

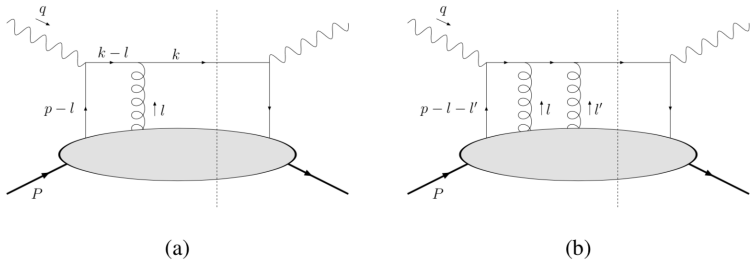


Figure 2.8: Types of diagrams contributing to the gauge link.

Understanding how the gauge links and the quark-quark correlator change in relation to the kind of process is fundamental to keep using the same universal distributions to fit different processes in a global analysis.

As an example, we outline the differences between the calculation of the amplitude for Drell–Yan scattering and SIDIS, taking into account also gauge links. We adopt the eikonal approximation, thus taking into account only the leading parts of the momenta of the quark after the photon scattering. The leading part in this calculation are the minus components.

The important difference for DY compared to the SIDIS case is that, instead of having an outgoing quark, we have an incoming antiquark. As a consequence the quark propagator in the upper part of the diagram is slightly modified. For the diagram of semi-inclusive DIS, shown in Fig. 2.9a, we get

$$\frac{i(\not{k}+\not{q}-\not{l}+m)}{(k+q-l)^2-m^2+i\epsilon} \approx i\frac{(k+q)^-\gamma^+}{-2l^+(k+q)^-+i\epsilon} = \frac{i\gamma^+}{2(-l^++i\epsilon)} \tag{2.114}$$

In the last step it is essential that we have $(k+q)^- > 0$, which is guaranteed by the fact that we have an outgoing quark with momentum $(k+q)$ in the final state.

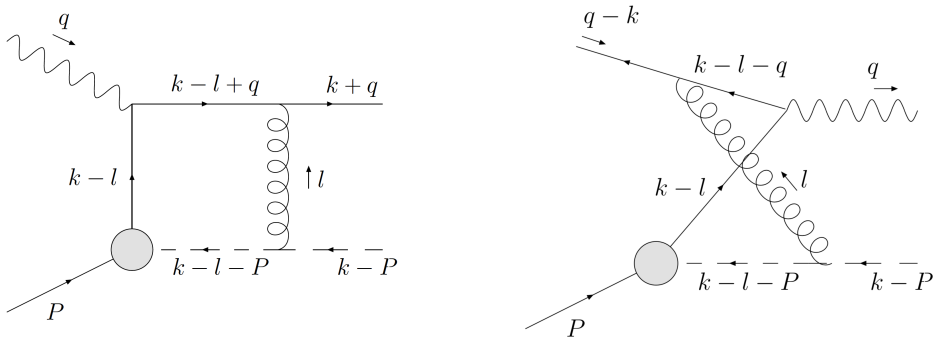
Instead, the explicit expression for the propagator in the Drell–Yan in Fig. 2.9b case is written:

$$\frac{\mathrm{i}(\not{k}-\not{q}-\not{l}+m)}{(k-q-l)^2-m^2+\mathrm{i}\epsilon} \approx \mathrm{i}\frac{-(q-k)^-\gamma^+}{2l^+(q-k)^-+\mathrm{i}\epsilon} = \frac{\mathrm{i}\gamma^+}{2\,(-l^+-\mathrm{i}\epsilon)} \tag{2.115}$$

Differently from the SIDIS case, in this case we have $(q-k)^- \geq 0$, because we consider an incoming quark with momentum $q-k$ in the initial state, which gives an opposite sign for $i\epsilon$. This detail leads to a noticeable difference for the calculation of the Sivers function, since when we apply the residue theorem to integrate over l^+, we use

$$\frac{1}{[l^+-i\epsilon]} \to 2\pi\mathrm{i}\delta\left[l^+\right] \tag{2.116}$$

while for semi-inclusive DIS we have to use



(a) (b)

Figure 2.9: Feynman diagrams for semi-inclusive DIS (a) and Drell–Yan (b)

$$\frac{1}{[l^+ + \mathrm{i}\epsilon]} \to -2\pi \mathrm{i}\delta\left[l^+\right] \tag{2.117}$$

and the final result will carry an opposite sign.

More in general, the gauge link introduced for the gauge invariance of DY processes, goes towards $-\infty$ instead of $+\infty$, or in other terms, we have a past-pointing Wilson line instead of a future-pointing one.

Therefore the DY quark correlator will assume the form

$$\Phi_{ij}\left(x, k_\perp\right) = \int \frac{\mathrm{d}\xi^- \mathrm{d}^2\xi_T}{(2\pi)^3} e^{\mathrm{i}p\cdot\xi} \langle P|\bar{\psi}_j(0)\mathcal{U}^{n_-}_{(0,-\infty)}\mathcal{U}^{n_-}_{(-\infty,\xi)}\psi_i(\xi)|P\rangle\Big|_{\xi^+=0} \tag{2.118}$$

The differences in the gauge links have the effect of modifying the T-odd distribution functions. In particular, the contribution of the Sivers and the Boer–Mulders functions will have an opposite sign for DIS and Drell–Yan.

2.5 Evolution of TMDs

2.5.1 Evolution formalism for unpolarized TMDs

To carry out a complete determination of TMD parton distribution and fragmentation functions one needs to study how they are affected by the change of parton flavor and values of the kinematic variables x, z, $P_\perp$ or q_T. To investigate their behavior, we have included measurements that cover a large kinematic region and have been collected from three different processes by different experimental collaboration.

However to have a full description of the TMDs we have to reproduce also how their shape and normalization change in relation with the energy scale Q^2, to establish in this way a connection between experimental measurements at any value of the hard scale and, at the same time, to test the validity of the

QCD formalism that describes them and predicts the scale evolution of TMDs. To fully appreciate TMD evolution, we included also Z boson production data, in order to cover a wider Q^2 interval, going from $Q \simeq 1$ GeV to $Q \simeq M_Z$.

In this section we discuss the formalism needed to describe the QCD evolution of TMDs. The resulting evolution equations derive from factorization theorems, which give us rules for the separation of soft and hard scales. Factorization is a fundamental concept in QCD phenomenology which provides much of the predictive power of the theory behind hard scattering processes, providing a solid basis for the description of perturbative and nonperturbative effects measured in various high energy experiments.

Different factorization procedures should be applied in different kinematic regions, characterized by the ratio of q_T with respect to Q. In the low-q_T region ($q_T \ll Q$) transverse momentum is generated by nonperturbative contributions and by soft and collinear gluon radiation. In this regime, TMD factorization applies and the nonperturbative objects involved are TMDs. In the high-q_T region ($q_T \gg Q$) transverse momentum is generated by hard perturbative radiation, therefore collinear factorization applies and the nonperturbative objects involved are collinear PDFs [47].

In the intermediate region, where $M \ll q_T \ll Q$, both descriptions could be valid; a prescription to match the two factorization schemes is then required [55]. A common solution is the introduction of a matching term

$$Y(q_T, M) = \left.\frac{d\sigma}{dq_T^2}\right|_{\substack{\text{fixed}\\ \text{order}}} - \left.\frac{d\sigma}{dq_T^2}\right|_{ASY} . \tag{2.119}$$

The first term is the differential cross section calculated at a fixed order n for high-q_T, while the second term subtracts the asymptotic limit $q_T \to 0$ of the fixed order, which is already included in the TMD description. By construction, the matching term should be negligible for low-q_T, increase in the intermediate region and finally match the fixed order calculation at high-q_T [56]. The transverse momentum q_T is related to P_{hT} by the relation $P_{hT} = -zq_T$.

The most interesting kinematic region to study the three-dimensional nucleon structure is the low-q_T regime, where the effects related to intrinsic transverse momenta of partons in the nucleon are more evident.

If we consider semi-inclusive DIS in a region where the transverse momentum q_T is much smaller than the hard scale Q, in general the cross section can be factorized in terms of nonperturbative elements which depends on the parton and produced hadron transverse momenta, in what is called TMD factorization:

$$d^{(n)}\sigma \approx \sum_a \mathcal{H}(Q,\mu) f_1^a(x, \boldsymbol{k}_\perp^2; \mu, \zeta) D_1^{a\to h}(z, \boldsymbol{P}_\perp^2, \mu, \zeta_h) + \mathcal{O}\left(\frac{P_{hT}}{Q}\right) . \tag{2.120}$$

Similarly for the Drell–Yan process we have

$$d^{(n)}\sigma \approx \sum_{a,b} \mathcal{H}(Q,\mu) f_1^a(x_a, p_T^2; \mu, \zeta_a) f_1^b(x, p_T^2, \mu, \zeta_b) + \mathcal{O}\left(\frac{q_T}{Q}\right). \tag{2.121}$$

where $\mathcal{H}$ is a factor that represents the hard interaction and μ is the factorization scale where separation between hard and soft part occur. The meaning of the scale ζ will be explained in the next sections.

One of the most commonly used method to deal with divergences that arise from the choices for the scales is the transverse momentum resummation, intially developed by Collins and Soper. In their original work [14] they derived the factorization theorem for for e^+e^- annihilation and it was later extended to other high energy processes.

The observation at the basis of the resummation procedure is that at low transverse momentum $q_T \ll Q$ fixed-order calculations are not reliable anymore [57]. The reason is that in the limit $q_T \to 0$, infra-red singularities do not cancel completely between real and virtual diagrams, leaving in the cross section $d\sigma/dq_T^2$ logarithmic terms of the form

$$\alpha_S^k \frac{\ln^m(Q^2/q_T^2)}{q_T^2} \tag{2.122}$$

where k is the perturbative order and $m = 1, \ldots, 2k-1$. For smaller q_T the coupling constant α_S is not a valid expansion parameter for the perturbative series, since it will not be able to compensate the divergence of the logarithms divided by the small value of q_T. Consequently, to have a consistent estimate of the cross section, it becomes necessary to *resum* the large logarithms to all order in α_S.

The most widespread method for resummation is the Collins-Soper-Sterman (CSS) formalism [55]. It is derived in configuration space, which is related to the momentum space by a Fourier transform. In this space the convolution of TMDs with the evolution has a simpler expression and the conservation of the soft-gluon momenta is guaranteed. The formalism was initially developed for the Drell–Yan process and then applied also to other high energy observables, including unpolarized and polarized SIDIS.

An important observation is that the CSS evolution equations used for the evolution of spin-dependent and unpolarized TMDs is the same [58].

2.5.2 Evolution operator

As we have seen the relevant structure function for the calculation of the unpolarized SIDIS cross section is $F_{UU,T}$; following the CS factorization it can be written as

$$\begin{aligned} F_{UU,T} =& \mathcal{H}\left(Q^2, \mu^2\right) \sum_a x e_a^2 \int d^2\boldsymbol{k}_\perp \frac{d^2\boldsymbol{P}_\perp}{z} \\ & \times \delta^{(2)}\left(z\boldsymbol{k}_\perp + \boldsymbol{P}_\perp + z\boldsymbol{q}_T\right) f_1^a\left(x, \boldsymbol{k}_\perp^2; \zeta\right) D_1^a\left(z, \boldsymbol{P}_\perp^2; \zeta\right). \end{aligned} \tag{2.123}$$

The corresponding formulation in ξ_T-space is

$$F_{UU,T}(x, z, \boldsymbol{P}_{hT}^2, Q^2) \approx \sum_a e_a^2 x \int d\xi_T^2 e^{i\xi_T P_{hT}/z} \tilde{f}_1{}^a(x, \xi_T^2; \mu, \zeta) \tilde{D}_1^{a\to h}(z, \xi_T^2; \mu, \zeta) \tag{2.124}$$

Having defined explicitly the convolution in ξ_T-space of TMD PDF and FF contained in the unpolarized structure function, we can now apply the CSS evolution formalism to them, and observe how they depend upon the renormalization scale μ and the rapidity scale ζ, related to two different kind of divergences of the perturbative part of $F_{UU,T}$. The ultraviolet divergences in the cross section calculation are regulated by introducing a renormalization scale μ. The dependence of the nonperturbative components on this scale is dictated by renormalization group equations (RGEs).

In addition to the standard UV divergences, the TMD distributions contain also the so-called *light-cone* or *rapidity* divergences [59]. They arise when the gauge links required for the definitions of the TMD correlators are oriented exactly in light-like directions. This type of singularities is not present in the collinear correlation functions because they cancel out when the integration over q_T is performed, which naturally is not carried out in the TMD case. Light-cone divergences correspond to gluons moving with infinite rapidity in the direction opposite to the hadron which contains them, and are not regulated by the use of an infrared cut-off.

The common approach to deal with the light-cone divergences to act on the direction of the Wilson lines. Initially they are deviated slightly from the light-cone directions, removing in this way the rapidity divergences. This deviation however introduces a dependence on the arbitrary rapidity scale ζ. This scale can be used to separate three kinematic regions and introduce three nonperturbative objects: a *soft factor* and two TMDs. To have a consistent definition of TMDs it is necessary to consider also the role of soft gluons in the formula for TMD factorization, which brings a separate soft factor, defined as the expectation value of the Wilson lines on bare fields. The soft factor, which would introduce an universal normalization in the calculation of the cross section of different processes, is usually adsorbed into the TMDs, now called *subtracted* TMDs, to distinguish them from the previous *unsubtracted* ones.

As we have seen, in general the QCD evolution with respect to μ follows standard renormalization group equations (RGE), while the evolution in ζ is determined by a different set of equations[60].

The evolution equation of the TMDs with respect to the rapidity scale can be written in terms of a rapidity-independent kernel

$$\frac{\partial \ln \tilde{f}(x, \xi_T; \mu, \zeta)}{\partial \ln \zeta} = K(\xi_T; \mu) \tilde{f}(x, \xi_T; \mu, \zeta)\,. \tag{2.125}$$

This is called the Collins-Soper kernel and it regulates the rapidity evolution for any quark TMDs, as its definitions is connected to the soft factor, which

is entirely related only to soft-gluon emission and hence spin and flavor independent. Considering that we want to apply this evolution kernel to factorized factors that will be evaluated at different μ, it is important to consider also its RG evolution, expressed by

$$\frac{dK(\xi_T;\mu)}{d\ln\mu} = \Gamma_{cusp}[\alpha_S(\mu)]\,. \tag{2.126}$$

The evolution of TMD distribution function from μ_b to another value of μ is governed by the standard renormalization group equations [61]:

$$\frac{d\tilde{f}(x,\xi_T;\mu,\zeta)}{d\ln\mu} = \gamma_F\left(\alpha_s(\mu);\zeta/\mu^2\right)\tilde{f}(x,\xi_T;\mu,\zeta) \tag{2.127}$$

where γ_F is called *anomalous dimension* and is defined as

$$\gamma_F = -\left(\Gamma_{cusp}\ln\frac{\zeta}{\mu^2} + \gamma^V\right). \tag{2.128}$$

The terms Γ_{cusp} and γ^V can be written as power series in α_S in the $\overline{MS}$ scheme [60].

Solving the Collins-Soper equation for rapidity and the RG equations for renormalization, we find the expression for the TMD distribution evolved from the initial scale $\{\mu_i,\zeta_i\}$ to $\{\mu_f,\zeta_f\}$:

$$\begin{aligned}\tilde{f}(x,\xi_T;\mu_f,\zeta_f) = \sum_{a=q,\bar{q},g}(C_i\otimes f_1^a)(x,\xi_T,\mu_i,\zeta_i)\left(\frac{\zeta_f}{\zeta_i}\right)^{-K(\xi_T,\mu_i)} \\ \times\exp\left\{\int_{\mu_i}^{\mu_f}\frac{d\mu}{\mu}\gamma_F\right\} f^a_{1\mathrm{NP}}(x,\xi_T)\end{aligned} \tag{2.129}$$

The terms in exponential is called Sudakov factor and written in a more compact form as $e^{S(\mu_i^2,Q)}$.

The explicit expression for the Sudakov form factor $S(\mu_b^2,Q)$ reads

$$S(\mu_b^2,Q^2) = -\frac{1}{2}\int_{\mu_b^2}^{Q_b^2}\frac{d\mu^2}{\mu^2}\left[A\Big(\alpha_S(\mu^2)\Big)\ln\left(\frac{Q^2}{\mu^2}\right) + B\Big(\alpha_S(\mu^2)\Big)\right], \tag{2.130}$$

where the functions A and B can be written as perturbative expansions in α_s

$$A(\alpha_S) = \sum_{k=1}^{\infty}A_k\left(\frac{\alpha_S}{\pi}\right)^k \qquad B(\alpha_S) = \sum_{k=1}^{\infty}B_k\left(\frac{\alpha_S}{\pi}\right)^k \tag{2.131}$$

If we consider the resummation only at order Next-to-Leading Logarithms (NLL), only the coefficient $A_{1,2}$ and B_1 are needed. Their explicit expressions are [62, 55]

$$A_1 = C_F, \qquad A_2 = \frac{1}{2}\left[C_A\left(\frac{67}{18}-\frac{\pi^2}{6}\right)-\frac{5}{9}N_f\right], \qquad B_1 = -\frac{3}{2}C_F\,, \tag{2.132}$$

where C_A and C_F are related to the number of colors N_c by

$$C_A = N_c, \qquad C_F = (N_c^2 - 1)/2N_c \tag{2.133}$$

and N_f is related to the number of active quarks

The TMD evolution framework leaves some arbitrary choices that can lead to different possible scenarios, according for example to the choice of the initial starting value for the factorization scale, and of the model for the nonperturbative part of the evolution kernel. We will discuss in the next chapter also the specific choices that we made for our parametrization of the TMDs.

2.5.3 Analytical calculation of Sudakov factor

The explicit expression of the Sudakov factor can be obtained at LL solving the integral

$$\int_{\mu_b^2}^{Q_b^2} \frac{1}{2\mu^2} C_F \frac{1}{\pi} \frac{4\pi}{\beta_0 \log\left(\mu^2/\Lambda^2\right)} \log\left(\frac{Q^2}{\mu^2}\right) \tag{2.134}$$

where C_F has been defined already and β_0 is a function dependent on the number of flavors available for the particles involved in the process at leading order:

$$\beta_0 = \beta_{LO}(Q) = 11 - \frac{2}{3} n_f(Q)\,. \tag{2.135}$$

The variable $n_f(Q)$ gives the number of available flavors for different energy thresholds:

$$n_f(Q) = \begin{cases} 3 & \text{if } Q < 1.4 \text{ GeV} \\ 4 & \text{if } 1.4 < Q < 4.5 \text{ GeV} \\ 5 & \text{if } Q > 4.5 \text{ GeV} \end{cases} \tag{2.136}$$

To calculate the integral Eq. (2.134) we assumed Q^2, μ_b^2, Λ^2 to be real and positive. Moreover these quantities are assumed to have the following relations:

$$\mu_b^2 < \Lambda^2; \quad Q_b^2 < \mu_b^2 \tag{2.137}$$

.

After integrating Eq. (2.134), the expression for the Sudakov factor becomes

$$\frac{2C_F}{\beta_0}\left[\log(Q_b^2)\log\left(\frac{\log(Q_b^2/\Lambda^2)}{\log(\mu_b^2/\Lambda^2)}\right) + \log(\Lambda^2)\log\left(\frac{\log(\mu_b^2/\Lambda^2)}{\log(Q_b^2/\Lambda^2)}\right) + \log\left(\frac{\mu_b^2}{Q_b^2}\right)\right]. \tag{2.138}$$

This can be rewritten in a more condensed way as

$$\exp\left\{-\left[2C_F\left(\log\left(\frac{\mu_b^2}{Q_b^2}\right) + \log\left(\frac{Q^2}{\Lambda^2}\right)\log\left(\frac{\log(Q_b^2/\Lambda^2)}{\log(\mu_b^2)/\Lambda^2)}\right)\cdot\frac{1}{\beta_0}\right]\right\}. \tag{2.139}$$

2.5.4 Wilson coefficients

To study the evolution of TMDs in configuration space it is useful to separate their perturbative part, related mostly to the interaction of the hadron with the probe, from the nonperturbative part.

To factorize these terms we apply the Operator Product Expansion, proposed by Wilson [63]. It states that a product of local operators have an expansion of the form

$$A(x)B(y) = \sum_n C_n(x-y)O_n(x) \tag{2.140}$$

when $|x-y|$ is sufficiently small. $O_n(x)$ is a basis of local operators. In the same way, a generic TMD function can be expanded over a basis of collinear PDFs using the appropriately named Wilson coefficients C. The TMD parton distribution functions, and similarly also the TMD FF, can be represented as [64]

$$f_1(x, \xi_T; \mu, \zeta) = [C \otimes f_1](x, \xi_T; \mu, \zeta) + \mathcal{O}(\xi_T \Lambda_{QCD}) \tag{2.141}$$

where the first factor of the sum can be calculated perturbatively at small ξ_T and the second term contains the contributions relevant for $\xi_T \gg \Lambda_{QCD}$.

The convolution is defined as

$$(C \otimes f_1)(x, \xi_T; \mu, \zeta) = \int_z^1 \frac{du}{u} C\left(\frac{x}{u}, \xi_T; \mu, \zeta\right) f_1(u; \mu)\,. \tag{2.142}$$

This formula is valid only for the small ξ_T region, as the explicit expression for the Wilson coefficients consists in a power series in $\alpha_S \ln(\mu^2/\mu_b^2)$, and the OPE is valid only when the logarithms are not divergent. The definition of μ_b is

$$\mu_b = \frac{2e^{-\gamma_E}}{\xi_T} \tag{2.143}$$

and the term γ_E in the exponent is the Euler gamma constant.

We want to have a consistent definition for the perturbative expression, thus minimizing all the divergent logarithm. A natural choice for the scale μ that takes into account the logarithms in Eq. (2.122) and in the Wilson coefficient is $\mu = \mu_b$.

To include in Eq. (2.142) also the contributions at large ξ_T we introduce an additional nonperturbative term $f_{1\mathrm{NP}}$, obtaining the following expression for the TMD distribution at the initial scales μ_i and ζ_i:

$$\tilde{f}_1^a(x, \xi_T; \mu_i, \zeta_i) = \sum_{i=q,\bar{q},g} C_{a/i}(x, \xi_T; \mu_i, \zeta_i) \otimes f_1^i(x; \mu_i) \tilde{f}_{1\mathrm{NP}}^a(x, \xi_T)\,. \tag{2.144}$$

The model $f_{1\mathrm{NP}}$ might be a function of x, the flavor parton a. Then the functional form of TMDs has different characteristic in different regimes, depending on the value of ξ_T. For small values it can be calculated in perturbative QCD, for the nonperturbative part at large ξ_T the calculations cannot be performed and it must be constrained by fitting experimental data.

2.5.5 Parametrization of renormalization scale

As shown in the previous section the cross section in transverse momentum space is obtained by an inverse Fourier transform of $f(x, \xi_T; Q^2)$. This involves an integration over $\xi_T = |\xi_T|$ from 0 to infinity. Because of the Landau pole of the perturbative strong coupling in the Sudakov exponent at $\xi_T \sim \Lambda_{QCD}$, this integration is not well defined for some of the possible value of ξ_T. The ξ_T-integration extends over both perturbative ($\xi_T \ll 1/\Lambda_{QCD}$) and nonperturbative ($\xi_T \geq \Lambda_{QCD}$) regions. In order to define a perturbative resummed cross section, a prescription for the ξ_T integration is required to prevent ξ_T from becoming too large and to avoid the Landau pole.

The most common prescription introduces a new parameter, called ξ_{max}, that represent the limit for ξ_T, beyond which we do not trust the perturbative calculation [59]. Then we introduce a new variable ξ^* for the OPE that saturates at the value of ξ_{max} even when ξ_T becomes large:

$$\lim_{\xi_T \to \infty} \xi_T^*(\xi_T) = \xi_{max}. \tag{2.145}$$

For $\xi_T \leq \xi_{max}$ the evolution is controlled by the function $K(\xi_T^*; \mu_b)$ that now depends on the new variable ξ_T^*) and on

$$\mu_b = \frac{2e^{-\gamma_E}}{\xi_T^*}.$$

The nonperturbative part introduces a new term in the Sudakov factor

$$e^{S(\xi_T, Q)} \to e^{S(\xi_T, Q) - g_{np}\xi_T^2} \tag{2.146}$$

where g_{np} is a coefficient that must be determined by comparison to data.
We can write this nonperturbative part also using the K function

$$g_{np}(\xi_T) = -K(\xi_T^*; \mu_b) + K(\xi_T; \mu_b)\,. \tag{2.147}$$

This definition implies a dependence on ξ_{max}, the parameter that separates the perturbative and nonperturbative regimes. This is more evident if we expand perturbatively the function $K(\xi_T, \mu_b)$ at lowest order, obtaining

$$g_{np}(\xi_T) \simeq \frac{\alpha_S(\mu_b) C_F}{\pi} \ln\left(1 + \frac{\xi_T^2}{\xi_{max}^2}\right). \tag{2.148}$$

In the regime $\xi_T \ll \xi_{max}$ we parametrize this term as $g_{np}(\xi_T) \simeq \frac{1}{2} g_2 \xi_T^2$.
We observe that the parameter g_2 is not totally arbitrary, because it is anti-correlated to the value of ξ_{max}. In general this is due to the fact that both $K(\xi_T; \mu_b)$ and the TMDs should not depend on the arbitrary choice of the value of ξ_{max}. This means that ξ_{max} has to be considered an arbitrary scale that separates different regimes, and not as a free parameter to fit. A variation in ξ_{max} requests that all terms in the convolution are rearranged, so that

the TMD PDF does not change. However in practical situations changing the value of ξ_{max} could give different results.

With the parametrization just discussed and using $\zeta_i = \mu_i = \mu_{\hat{b}}^2$ and $\zeta_f = \mu_f^2 = Q^2$ we obtain the equation Eq. (2.129). We can write this expression including all the evolution terms in a single evolution operator, called R, that acts on the initial TMD PDF, evaluated at the scale μ_b:

$$f_1(x, \xi_T; Q^2) = R(\xi_T; Q^2, \mu_b^2) f(x, \xi_T; \mu_b^2). \tag{2.149}$$

The effect of evolution is represented through the action of an evolution operator on an input TMD distribution evaluated at the scale μ_b, which depends on ξ_T. This dependence assures that there is a smooth matching between the perturbative region at small ξ_T and the nonperturbative one at large ξ_T. In addition, from Eq. (2.129) it can be seen that choosing different models for the nonperturbative part affects the whole ξ_T spectrum, not only the large ξ_T region.

The procedure used for the evolution of TMD PDFs can be reproduced in a similar way for the TMD fragmentation function.

Finally, having defined all the necessary elements, we can write the convolution in configuration space of the evolved D_1 and f_1 that appears in $F_{UU,T}$:

$$\begin{aligned} F_{UU,T}(x, z, \boldsymbol{P}_{hT}^2, Q^2) =& \frac{1}{2\pi z^2} \int_0^\infty d\xi_T |\xi_T| J_0\left(|\xi_T| |\frac{\boldsymbol{P}_{hT}}{z}| \right) f_1(x, \mu_b^2) R(\xi_T; Q^2, \mu_b^2) \\ & \times D_1(z, \mu_b^2) z^2 R(\xi_T; Q^2, \mu_b^2) \exp\left(- \frac{\boldsymbol{P}_{hT}^2 \xi_T^2}{4z^2} \right). \end{aligned} \tag{2.150}$$

2.5.6 TMD evolution and Sivers distribution

The Sivers distribution function is a fundamental tool to study how the internal structure of nucleons influences scattering processes with polarized targets. To have a complete picture of their behavior is fundamental to comprehend how it is affected by changes in the energy scale of the process.

The formalism used to describe evolution of unpolarized TMDs can be applied also in the spin-dependent case. The evolution of the Sivers function is usually defined in terms of its first derivative in configuration space $\tilde{f}_{1T}^{\perp(1)}(x, \xi_T^2)$. The first moment is defined as

$$f_{1T}^{\perp(1)}(x) = \int d^2\boldsymbol{k}_\perp \frac{k_\perp^2}{2M_p^2} f_{1T}^\perp(x, \boldsymbol{k}_\perp^2) = \frac{\langle k_\perp^2 \rangle}{2M_p^2} f_{1T}^\perp(x)\,. \tag{2.151}$$

As we observed, the evolution of TMDs depends on two different energy scales. The Collins-Soper (CS) equation describes the evolution regulated by the rapidity ζ and the renormalization group (RG) equations give evolution with respect to μ.

The CS evolution equation applied to the first moment of the Sivers function reads

$$\frac{\partial \ln \tilde{f}_{1T}^{\perp(1)a}(x,\xi_T^2;\mu,\zeta)}{\partial \ln \zeta} = K\left(\xi_T;\mu\right) \tag{2.152}$$

where we use the same evolution kernel $K(\xi_T,\mu)$, defined in Eq. (2.126), and anomalous dimension $\gamma_F(\alpha_S(\mu);\ln(\zeta/\mu^2))$. Instead, the RG equation for the first moment assumes the form:

$$\frac{d\tilde{f}_{1T}^{\perp(1)a}(x,\xi_T^2;\mu,\zeta)}{d\ln\mu} = \gamma_F\left(\alpha_s(\mu);\zeta/\mu^2\right)\tilde{f}_{1T}^{\perp(1)a}(x,\xi_T^2;\mu,\zeta)\,. \tag{2.153}$$

The TMD evolution for the Sivers function $f_{1T}^{\perp(1)}$ follows the same equations of the unpolarized TMD f_1, however, there is a difference in the OPE expansion: while for the evolution of the unpolarized distributions we consider the convolution with the Wilson coefficient C_i, in the case of Sivers function we have a different coefficient C^{Siv}. Then, we will apply the evolution equations to the object

$$\tilde{f}_{1T}^{\perp a}(x,\xi_T;\mu_b,\zeta) = \sum_{i=q,\bar{q},g} C_{a/i}^{Siv}(x,\xi_T,\mu_b,\zeta)\otimes f_{1T}^{\perp(1)i}(x,\mu_b)\tilde{f}_{1T\,\mathrm{NP}}^{\perp a}(x,\xi_T)\,. \tag{2.154}$$

We substitute the first moment of the Sivers function with the twist-3 Qiu-Sterman function $T_{aF}(x,x,\mu)$ which can be considered the collinear counterpart of the Sivers function [65]. The first moment $f_{1T}^{\perp(1)}(x)$ is related to the Qiu-Sterman function by the following relation [66]:

$$f_{1T}^{\perp(1)}(x) = -\frac{1}{2M}T_F(x,x)\,. \tag{2.155}$$

The μ_b evolution for this term follows the Efremov–Teryaev–Qiu–Sterman (ETQS) evolution equations [67]. In our case the Wilson coefficients are calculated at leading-order and then they assume the form [68]:

$$C^{Siv(0)} = \delta(1-x)\delta^{ai}\,. \tag{2.156}$$

Moreover, we assume that $f_{1T}^{\perp(1)}(x)$ evolves in the same way as $f_1(x)$, which constitutes, however, a simplification [69].

Chapter 3

Phenomenology of unpolarized TMDs

3.1 Introduction

The determination of TMD parton distribution and fragmentation functions, due to their nonperturbative nature, cannot be done only through analytical methods but has to be extracted form experimental observables measured in hard scattering processes. Through this extraction from experimental data, we want to study the evolution of TMDs over a large enough range of energy, and to test their universality among different processes. To achieve this we included measurements taken from SIDIS, Drell–Yan and Z boson production from different experimental collaborations at different energy scales.

As we explained in the previous chapter, if we take into account target and probe polarizations, we can define several different TMDs [15, 17, 18, 19, 20]. In the past there have been many different attempts to extract some of them, with different levels of accuracy and sometimes without considering the contribution of TMD evolution [23, 24, 25, 26, 27, 28, 29, 30, 31]. In this chapter we present our work [70] on the extraction of the simplest ones of all the possible transverse distributions, the unpolarized TMD PDF $f_1^q(x, k_\perp^2)$ and the unpolarized TMD FF $D_1^{q\to h}(z, P_\perp^2)$.

Even if they appear simple, the phenomenology of these unpolarized TMDs presents several challenges [71], such as the choice of a functional form for the nonperturbative components of TMDs, the inclusion of a possible dependence on partonic flavor [72], the implementation of TMD evolution [60, 73], the matching to fixed-order calculations in collinear factorization [74].

In this chapter we present the first attempt at a global fit of unpolarized TMDs, including their evolution. We have considered three kinds of processes: SIDIS, Drell–Yan processes (DY) and the production of Z bosons. To date, they represent all possible processes where experimental information is available for unpolarized TMD extractions. The only important process currently

missing is electron-positron annihilation, which is particularly important for the determination of TMD FFs [60]. The expression for the observables measured in SIDIS and DY experiments contains at least a TMD parton distribution, while the cross section of e^+e^- annihilation is proportional to the convolution of two TMD fragmentation functions. Therefore, when annihilation data will be available, including them in an eventual global fit will allow to extract the FF without an inherent bias from the coupled PDF. The measurements for the included processes have been taken by different experimental collaboration, such as HERMES at DESY, COMPASS at CERN and E288, E605, CDF, D0 at Fermilab.

3.2 Formalism

3.2.1 Semi-inclusive DIS

The first step taken in our study was the independent analysis of semi-inclusive DIS measured at HERMES the same group of data considered in a previous work of the Pavia group [72], using a similar Gaussian functional form for the TMDs but with an important addition, that is the contributions of TMD evolution. To enrich the statistics of SIDIS data, for the next stage we decided to include the large number of COMPASS measurements, even if they were not as sensible to parton flavor as the HERMES data. In order to make the analysis more sensible to the change of energy, subsequently we considered also data sets for DY and Z boson production cross section, collected at Fermilab, which covers a larger interval of Q^2. Following the initial steps of our analysis, which started from the HERMES data, in this section we start with outlining the formalism necessary to the analysis of the semi-inclusive DIS multiplicities.

The available SIDIS data from HERMES and COMPASS, refer to SIDIS hadron multiplicities, i.e., to the differential number of hadrons produced per corresponding inclusive DIS event. In terms of cross sections, the multiplicities can be defined as

$$m_N^h(x, z, |\boldsymbol{P}_{hT}|, Q^2) = \frac{d\sigma_N^h/(dxdzd|\boldsymbol{P}_{hT}|dQ^2)}{d\sigma_{\rm DIS}/(dxdQ^2)}\,, \tag{3.1}$$

where $d\sigma_N^h$ is the differential cross section for the SIDIS process and $d\sigma_{\rm DIS}$ is the corresponding inclusive one, and where $\boldsymbol{P}_{hT}$ is the component of $\boldsymbol{P}_h$ transverse to $\boldsymbol{q}$ (we follow here the notation suggested in Ref. [75]). We already discussed how the unpolarized inclusive and semi-inclusive DIS cross sections can be expressed in terms of structure functions. In the single-photon-exchange approximation, the multiplicities can be written as ratios of structure functions

(see Ref. [46] for details):

$$m_N^h(x,z,|\boldsymbol{P}_{hT}|,Q^2) = \frac{2\pi\,|\boldsymbol{P}_{hT}|F_{UU,T}(x,z,\boldsymbol{P}_{hT}^2,Q^2) + 2\pi\varepsilon|\boldsymbol{P}_{hT}|F_{UU,L}(x,z,\boldsymbol{P}_{hT}^2,Q^2)}{F_T(x,Q^2)+\varepsilon F_L(x,Q^2)}\,. \quad (3.2)$$

The semi-inclusive cross section can be expressed in a factorized form in terms of TMDs only in the kinematic limits $M^2 \ll Q^2$ and $\boldsymbol{P}_{hT}^2 \ll Q^2$. In these limits, the structure function $F_{UU,L}$ of Eq. (3.2) can be neglected [47]. The structure function F_L in the denominator contains contributions involving powers of the strong coupling constant α_S at an order that goes beyond the level reached in this analysis; hence, it will be consistently neglected (for measurements and estimates of the F_L structure function see, e.g., Refs. [76, 77, 78, 79] and references therein).

To express the structure functions in terms of TMD PDFs and FFs, we rely on the factorized formula for SIDIS [14, 55, 80, 81, 22, 59, 82, 83, 84] (see Fig. 2.3 for a graphical representation of the involved transverse momenta):

$$\begin{aligned} F_{UU,T}(x,z,\boldsymbol{P}_{hT}^2,Q^2) &= \sum_a \mathcal{H}^a_{UU,T}(Q^2) \\ &\times x\int d^2\boldsymbol{k}_\perp\, d^2\boldsymbol{P}_\perp\, f_1^a\big(x,\boldsymbol{k}_\perp^2;Q^2\big)\, D_1^{a\to h}\big(z,\boldsymbol{P}_\perp^2;Q^2\big)\,\delta^{(2)}\big(z\boldsymbol{k}_\perp - \boldsymbol{P}_{hT}+\boldsymbol{P}_\perp\big) \\ &+ Y_{UU,T}\big(Q^2,\boldsymbol{P}_{hT}^2\big) + \mathcal{O}\big(M^2/Q^2\big)\,. \end{aligned} \quad (3.3)$$

Here, $\mathcal{H}_{UU,T}$ is the hard scattering part. The Fourier transform in momentum space of this structure function is

$$\begin{aligned} F_{UU,T}(x,z,\boldsymbol{P}_{hT}^2,Q^2) &= 2\pi\sum_a \mathcal{H}^a_{UU,T}(Q^2) \\ &\times x\int_0^\infty d\xi_T\xi_T J_0\big(\xi_T|\boldsymbol{P}_{hT}|/z\big)\,\tilde{f}_1^a\big(x,\xi_T^2;Q^2\big)\,\tilde{D}_1^{a\to h}\big(z,\xi_T^2;Q^2\big) \\ &+ Y_{UU,T}\big(Q^2,\boldsymbol{P}_{hT}^2\big) + \mathcal{O}\big(M^2/Q^2\big)\,. \end{aligned} \quad (3.4)$$

For the moment we have included the possibility of higher order contributions in $\mathcal{H}_{UU,T}$ and the presence of the matching term $Y_{UU,T}$ of Eq. (2.119).

TMDs generally depend on two energy scales [22], which enter via the renormalization of ultraviolet and rapidity divergences. In this chapter we choose them to be equal and set them to Q^2. The term $Y_{UU,T}$ is introduced to ensure a matching to the perturbative fixed-order calculations at higher transverse momenta.

In our analysis, we neglect any correction of the order of M^2/Q^2 or higher to Eq. (3.3). At large Q^2 this is well justified. However, fixed-target DIS experiments typically collect a large amount of data at relatively low Q^2 values, where these assumptions should be all tested in future studies. The reliability of the theoretical description of SIDIS at low Q^2 has been recently discussed in Refs. [56, 85].

Eq. (3.3) can be expanded in powers of α_S. In the present analysis, we will consider only the terms at order α_S^0. In this case $\mathcal{H}^a_{UU,T}(Q^2) \approx e_a^2$ and $Y_{UU,T} \approx 0$. However, perturbative corrections include large logarithms $L \equiv \log\left(z^2Q^2/P_{hT}^2\right)$, so that $\alpha_S L \approx 1$. In the present analysis, we will take into account all leading and NLL. We remark that formulas at NNLL are available in the literature [86].

In these approximations (α_S^0 and NLL), only the first term in Eq. (3.3) is relevant (often in the literature this has been called W term). We expect this term to provide a good description of the structure function only in the region where $P_{hT}^2 \ll Q^2$. It can happen that $Y_{UU,T}$, defined in the standard way (see, e.g., Ref. [55]), gives large contributions also in this region, but it is admissible to redefine it in order to avoid this problem [74].

3.2.2 Drell–Yan and Z production

To test the universality of the TMD parton distribution functions we included in our analysis also measurements of Drell–Yan cross section. As we have seen, the description of a Drell–Yan cross section involves the convolution of two PDFs, one for each of the hadrons present in the initial state. Due to the universality property, we use the same expression for the distributions involved in the modelization of the Drell–Yan and SIDIS process. In a Drell–Yan process, two hadrons A and B with momenta P_A and P_B collide at a center-of-mass energy squared $s = (P_A + P_B)^2$ and produce a virtual photon or a Z boson plus hadrons. The boson decays into a lepton-antilepton pair. The invariant mass of the virtual photon is $Q^2 = q^2$ with $q = l + l'$. We introduce the rapidity of the virtual photon/Z boson

$$\eta = \frac{1}{2}\log\left(\frac{q^0 + q_z}{q^0 - q_z}\right). \tag{3.5}$$

where the z direction is defined along the momentum of hadron A (see Fig. 2.6).

The cross section can be written in terms of structure functions [87, 49]. For our purposes, we need the unpolarized cross section integrated over $d\Omega$ and over the azimuthal angle of the virtual photon,

$$\frac{d\sigma}{dQ^2\, dq_T^2\, d\eta} = \sigma_0^{\gamma,Z}\left(F_{UU}^1 + \frac{1}{2}F_{UU}^2\right). \tag{3.6}$$

The elementary cross sections are

$$\sigma_0^\gamma = \frac{4\pi^2\alpha_{\rm em}^2}{3Q^2 s}, \quad \sigma_0^Z = \frac{\pi^2\alpha_{\rm em}}{s\sin^2\theta_W\cos^2\theta_W} B_R(Z\to\ell^+\ell^-)\delta(Q^2 - M_Z^2), \tag{3.7}$$

where θ_W is Weinberg's angle, M_Z is the mass of the Z boson, and $B_R(Z \to \ell^+\ell^-)$ is the branching ratio for the Z boson decay in two leptons.We used the values $\sin^2\theta_W = 0.2313$, $M_Z = 91.18$ GeV, and $B_R(Z \to \ell^+\ell^-) = 3.366$ [88]. We adopted the narrow-width approximation, i.e., we neglect contributions for

$Q^2 \neq M_Z^2$. Similarly to the SIDIS case, in the kinematic limit $q_T^2 \ll Q^2$ the structure function F_{UU}^2 can be neglected (for measurement and estimates of this structure function see, e.g., Ref. [89] and references therein).

The longitudinal momentum fractions of the annihilating quarks can be written in terms of rapidity in the following way

$$x_A = \frac{Q}{\sqrt{s}} e^{\eta}, \qquad\qquad x_B = \frac{Q}{\sqrt{s}} e^{-\eta}. \tag{3.8}$$

Some experiments use the variable x_F, which is connected to the other variables by the following relations

$$\eta = \sinh^{-1}\left(\frac{\sqrt{s}}{Q}\frac{x_F}{2}\right), \quad x_A = \sqrt{\frac{Q^2}{s} + \frac{x_F^2}{4}} + \frac{x_F}{2}, \quad x_B = x_A - x_F. \tag{3.9}$$

The structure function F_{UU}^1 can be written as (see Fig. 2.6 for a graphical representation of the involved transverse momenta)

$$\begin{aligned} F_{UU}^1(x_A, x_B, \boldsymbol{q}_T^2, Q^2) &= \sum_a \mathcal{H}_{UU}^{1a}(Q^2) \\ &\times \int d^2\boldsymbol{k}_{\perp A}\, d^2\boldsymbol{k}_{\perp B}\, f_1^a\big(x_A, \boldsymbol{k}_{\perp A}^2; Q^2\big)\, f_1^{\bar{a}}\big(x_B, \boldsymbol{k}_{\perp B}^2; Q^2\big) \\ &\times \delta^{(2)}\big(\boldsymbol{k}_{\perp A} - \boldsymbol{q}_T + \boldsymbol{k}_{\perp B}\big) + Y_{UU}^1\big(Q^2, \boldsymbol{q}_T^2\big) + \mathcal{O}\big(M^2/Q^2\big)\,. \end{aligned} \tag{3.10}$$

As in the SIDIS case, in our analysis we neglect the Y_{UU} term. The hard coefficients $\mathcal{H}_{UU}$ have been defined in Eq. (2.93).

The Drell–Yan data considered cover a wide range of Q^2 and they are fundamental for the factor g_K which appears in Eq. (2.129). Also in this case, it is useful to express the convolution of PDF in ξ_T-space.

The Fourier transform of the right-hand side of Eq. (3.10) has been given in Eq. (2.92).

3.2.3 Choices for unpolarized TMDs and their evolution

In this section we will discuss the explicit expression for the terms in Eq. (2.92), and the choices adopted for the arbitrary parameters of the TMD evolution formalism. Evolution equations quantitatively describe the connection between different values for the energy scales. In the following we will set their initial values to μ_b^2 and their final values as Q^2, so that only these quantities will be relevant in a TMD distribution. Following the formalism of Refs. [22, 59], the fragmentation functions in configuration space for a parton with flavor a at a certain scale Q^2 can be written as

$$\begin{aligned} \widetilde{D}_1^{a\to h}(z, \xi_T^2; Q^2) &= \sum_{i=q,\bar{q},g} \big(\hat{C}_{a/i} \otimes D_1^{i\to h}\big)(z, \bar{\xi}_*, \mu_b^2) \\ &\times e^{S(\mu_b^2, Q^2)} \left(\frac{Q^2}{\mu_b^2}\right)^{-K(\bar{\xi}_*;\mu_b)} \left(\frac{Q^2}{Q_0^2}\right)^{g_K(\xi_T)} \widetilde{D}_{1\mathrm{NP}}^{a\to h}(z, \xi_T^2)\,. \end{aligned} \tag{3.11}$$

and the corresponding expression for the TMD PDF has been given in Eq. (2.129). We make the canonical choice for the scale μ_b to be

$$\mu_b = \frac{2e^{-\gamma_E}}{\bar{\xi}_*} \,, \tag{3.12}$$

where γ_E is the Euler constant and

$$\bar{\xi}_* \equiv \bar{\xi}_*(\xi_T; \xi_{\min}, \xi_{\max}) = \xi_{\max} \left(\frac{1 - \exp(-\xi_T^4/\xi_{\max}^4)}{1 - \exp(-\xi_T^4/\xi_{\min}^4)} \right)^{1/4} . \tag{3.13}$$

The variable $\bar{\xi}_*$ replaces the simple dependence upon ξ_T in the perturbative parts of the TMD definitions of Eqs. (2.129) and (3.11). This substitution is necessary because at large ξ_T these parts are no longer reliable. Therefore, $\bar{\xi}_*$ is constructed so that it saturates on the maximum value $\xi_{\max}$, as suggested by the CSS formalism [22, 59]. On the other hand, at small ξ_T the TMD formalism is not valid and should be matched to the fixed-order collinear calculations. The way the matching is implemented is not unique. In any case, the TMD contribution can be arbitrarily modified at small ξ_T, considering that resummed calculations are not reliable in that region. In our approach, we mirror the general idea that leads to the introduction of the parameter $\xi_{\max}$ and we choose to saturate $\bar{\xi}_*$ at the minimum value $\xi_{\min} \propto 1/Q$. With the appropriate choices, for $\xi_T = 0$ the Sudakov exponent vanishes, as it should [90, 91]. Our choice partially corresponds to modifying the resummed logarithms as in Ref. [92] and to other similar modifications proposed in the literature [93, 74]. These kind of prescriptions are in part justified by a unitarity constraint, so that by integrating over the impact parameter ξ_T, the collinear expression for the cross section in terms of collinear PDFs is recovered, at least at leading order [74]. In the original CSS paper the functional form adopted was based on a square-root function that saturated at a certain value. With our definition of, $\bar{\xi}_*$ rapidly saturates to its limits $\xi_{\max}$ and $\xi_{\min}$, leaving a reasonably wide region where it reduces to ξ_T. It becomes particularly critical for small Q, where $\xi_{\max}$ and $\xi_{\min}$ assume really close values.

There are alternative schemes available to deal with the high-ξ_T region, which differ from our approach, such as the the so-called "complex-ξ prescription" [94] or another one based on an extrapolation of the perturbative small-ξ_T calculation to the large ξ_T region based on dynamical power corrections [95].

The values of $\xi_{\max}$ and $\xi_{\min}$ could be regarded as arbitrary scales separating perturbative from nonperturbative regimes. Their definition is not completely unconstrained, but at the same time there is not a method to accurately determine them. We choose to fix them to the values

$$\xi_{\max} = 2e^{-\gamma_E}\ \text{GeV}^{-1} \approx 1.123\ \text{GeV}^{-1}, \qquad \xi_{\min} = 2e^{-\gamma_E}/Q \,. \tag{3.14}$$

The motivations behind these choices are the following:

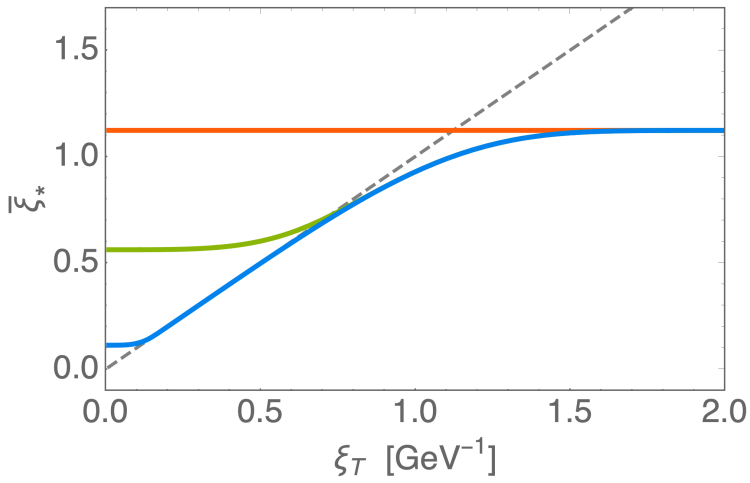


Figure 3.1: Values of $\bar{\xi}^*$ as a function of ξ_T for different energy scales Q.

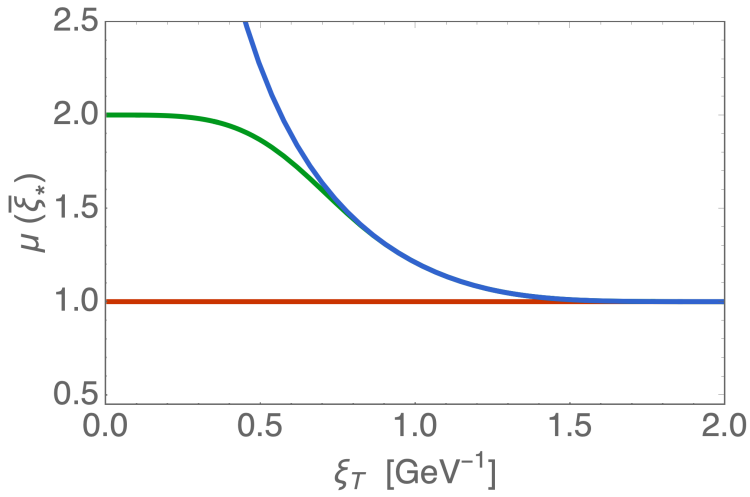


Figure 3.2: Values of $\mu(\bar{\xi}^*)$ as a function of ξ_T for different energy scales Q.

- the scale μ_b is constrained between 1 GeV and Q, so that the collinear PDFs are never computed at a scale lower than 1 GeV and the lower limit of the integrals contained in the definition of the perturbative Sudakov factor (see Eq. (2.130)) can never become larger than the upper limit;

- at the initial energy scale, by definition, we do not expect any contribution of TMD evolution. In fact, if we set the energy scale at the initial value $Q = Q_0 = 1$ GeV, $\xi_{\max} = \xi_{\min}$ and there are no evolution effects: the TMD is simply given by the corresponding collinear function multiplied by a nonperturbative contribution depending on $k_\perp$ (plus possible corrections of order α_S from the Wilson coefficients).

In Fig. 3.1 are shown the values for $\bar{\xi}^*(\xi_T, \xi_{\max}, \xi_{\min})$ as a function of ξ_T

for $Q = 1, 2, 10$ GeV. We can observe how this variable reach the saturation value more slowly with higher values of Q and in particular, for the initial scale $Q = 1$ GeV we reach immediately the saturation level. The consequence of this saturation can be observed in Fig. 3.2, where the corresponding values of $\mu(\bar{\xi}^*)$ are presented. In this case for $Q = 1$ GeV, the scale μ_b remain at the initial value and there are no evolution effects, while at higher Q we observe a larger region where the scale is not saturated, as expected from our definition. Alternative choices for the parametrization of the scale μ_b and their consequences will be discussed in the next sections.

Having discussed the parametrization of the energy scale, we need now to discuss the explicit expression of the other terms related to the evolution, which appear in Eqs. (2.129) and (3.11). At NLL accuracy, for our choice of scales the evolution kernel $K(\bar{\xi}_*, \mu_b)$ for the rapidity scale ζ is equal to zero. Similarly, C and $\hat{C}$ are perturbatively calculable Wilson coefficients for the TMD distribution and fragmentation functions, respectively. We remind that in the present analysis, we consider only the leading-order term in the α_S expansion for C and $\hat{C}$, i.e.,

$$C_{a/i}\Big(\frac{x}{u}, \bar{\xi}_*, \alpha_S\big(\mu_b^2\big)\Big) \approx \delta_{ai}\delta(1-x/u), \quad \hat{C}_{a/i}\Big(\frac{z}{u}, \bar{\xi}_*, \alpha_S\big(\mu_b^2\big)\Big) \approx \delta_{ai}\delta(1-z/u). \tag{3.15}$$

As a consequence of the choices we made, the expression for the evolved TMD functions reduces to

$$\widetilde{f}_1^a(x, \xi_T^2; Q^2) = f_1^a(x; \mu_b^2)\; e^{S(\mu_b^2, Q^2)}\; e^{g_K(\xi_T)\ln(Q^2/Q_0^2)}\; \widetilde{f}_{1\mathrm{NP}}^a(x, \xi_T^2)\;, \tag{3.16}$$

$$\widetilde{D}_1^{a\to h}(z, \xi_T^2; Q^2) = D_1^{a\to h}(z; \mu_b^2)\; e^{S(\mu_b^2, Q^2)}\; e^{g_K(\xi_T)\ln(Q^2/Q_0^2)}\; \widetilde{D}_{1\mathrm{NP}}^{a\to h}(z, \xi_T^2)\;. \tag{3.17}$$

The explicit expression for the Sudakov exponent S was given in Eq. (2.130)[1].

We use the approximate analytic expression for α_S at NLO with $\Lambda_{\mathrm{QCD}} =$ 340 MeV, 296 MeV, 214 MeV for three, four, five flavors, respectively, corresponding to a value of $\alpha_S(M_Z) = 0.117$. We fix the flavor thresholds at $m_c = 1.5$ GeV and $m_b = 4.7$ GeV. The integration of the Sudakov exponent in Eq. (2.130) can be done analytically (for the complete expressions see, e.g., Refs. [96, 97, 83]) as discussed in Section 2.5.

Following Refs. [98, 99, 100], for the nonperturbative Sudakov factor we use the most common choice found in the literature, which is:

$$g_K(\xi_T) = -g_2\xi_T^2/4 \tag{3.18}$$

with g_2 a free parameter. Recently, several alternative forms have been proposed [101, 102]. Also, recent theoretical studies tried to calculate this term using nonperturbative methods [103]. In Ref. [104], a good agreement with

[1]We warn the reader that the definition of the Sudakov factor in (2.3) of [70] is missing a term 1/2

data was achieved even without this term, but this is not possible when including data at low Q^2.

We still need to give an explicit form for f_1^a and $f_{1\mathrm{NP}}^a$ to have a complete expression of $F_{UU,T}$; the collinear PDF f_1^a has been studied in detail by many theoretical groups and its expression, even if nonperturbative, is well known at the moment. The collection of these collinear extractions, obtained by different groups, is available through the LHAPDF library [105]. In this analysis, for the collinear PDFs f_1^a we adopt the GJR08FFnloE set [106], and for the collinear fragmentation functions the DSS14 NLO set for pions [107] and the DSS07 NLO set for kaons [108]. After the completion of our analysis, a new set of kaon fragmentation functions was presented in Ref. [109]. We will comment on the use of other PDF sets in Sec. 3.5.3. The Compass data distinguish only negatively and positively charged hadrons, and it was chosen to identify them with pions. It could be possible to get a plausible estimate of the percentage of kaons in the sample and include more accurately their contribution. However, we stress that our analysis is not flavor dependent and, when considering the necessity to distinguish between pions and kaons contributions, we evaluated that the smaller percentage of kaons should not introduce a significant error in the final prediction for the multiplicities. However, in future studies we will take into account this aspect of the collinear FFs when including Compass data.

Having determined the terms of the expression that are well known or perturbatively calculable, we have now to deal with the contributions at large ξ_T and choose a reasonable model for $f_{1\mathrm{NP}}$ and $D_{1\mathrm{NP}}$. We parametrize the intrinsic nonperturbative parts of the TMDs in the following ways

$$\widetilde{f}_{1\mathrm{NP}}^a(x,\xi_T^2) = \frac{1}{2\pi} e^{-g_{1a}\frac{\xi_T^2}{4}} \left(1 - \frac{\lambda g_{1a}^2}{1+\lambda g_{1a}} \frac{\xi_T^2}{4}\right) , \tag{3.19}$$

$$\widetilde{D}_{1\mathrm{NP}}^{a\to h}(z,\xi_T^2) = \frac{g_{3a\to h}\; e^{-g_{3a\to h}\frac{\xi_T^2}{4z^2}} + \big(\lambda_F/z^2\big) g_{4a\to h}^2 \left(1 - g_{4a\to h}\frac{\xi_T^2}{4z^2}\right)\; e^{-g_{4a\to h}^2\frac{\xi_T^2}{4z^2}}}{2\pi z^2 \Big(g_{3a\to h} + \big(\lambda_F/z^2\big) g_{4a\to h}^2\Big)} . \tag{3.20}$$

After performing the anti-Fourier transform, the $f_{1\mathrm{NP}}$ and $D_{1\mathrm{NP}}$ in momentum space correspond to

$$f_{1\mathrm{NP}}^a(x,\boldsymbol{k}_\perp^2) = \frac{1}{\pi}\; \frac{\big(1+\lambda \boldsymbol{k}_\perp^2\big)}{g_{1a} + \lambda\; g_{1a}^2}\; e^{-\frac{\boldsymbol{k}_\perp^2}{g_{1a}}} , \tag{3.21}$$

$$D_{1\mathrm{NP}}^{a\to h}(z,\boldsymbol{P}_\perp^2) = \frac{1}{\pi}\; \frac{1}{g_{3a\to h} + \big(\lambda_F/z^2\big) g_{4a\to h}^2} \left(e^{-\frac{\boldsymbol{P}_\perp^2}{g_{3a\to h}}} + \lambda_F \frac{\boldsymbol{P}_\perp^2}{z^2}\; e^{-\frac{\boldsymbol{P}_\perp^2}{g_{4a\to h}}}\right) . \tag{3.22}$$

The TMD PDF at the starting scale is therefore a normalized sum of a Gaussian with variance g_1 and the same Gaussian weighted by a factor $\lambda \boldsymbol{k}_\perp^2$. The

TMD FF at the starting scale is a normalized sum of a Gaussian with variance g_3 and a second Gaussian with variance g_4 weighted by a factor $\lambda_F \boldsymbol{P}_\perp^2/z^2$. The choice of this particular functional forms is motivated by model calculations: the weighted Gaussian in the TMD PDF could arise from the presence of components of the quark wave function with angular momentum $L = 1$ [110, 111, 112, 113, 114]. Similar features occur in models of fragmentation functions [115, 110, 45].

The Gaussian width of the TMD distributions may depend on the parton flavor a [72, 45, 116]. In the present analysis, however, we assume they are flavor independent. The motivation for this choice is that most of the data we are considering are not sufficiently sensitive to flavor differences, leading to unclear results. We will devote attention to this issue in further studies.

Finally, we assume that the Gaussian width of the TMD depends on the fractional longitudinal momentum x according to

$$g_1(x) = N_1 \, \frac{(1-x)^\alpha \; x^\sigma}{(1-\hat{x})^\alpha \; \hat{x}^\sigma} \,, \tag{3.23}$$

where α, σ, and $N_1 \equiv g_1(\hat{x})$ with $\hat{x} = 0.1$, are free parameters. Similarly, for fragmentation functions we have

$$g_{3,4}(z) = N_{3,4} \, \frac{(z^\beta + \delta) \; (1-z)^\gamma}{(\hat{z}^\beta + \delta) \; (1-\hat{z})^\gamma} \,, \tag{3.24}$$

where β, γ, δ, and $N_{3,4} \equiv g_{3,4}(\hat{z})$ with $\hat{z} = 0.5$ are free parameters.

The average transverse momentum squared for the distributions in Eq. (3.21) and (3.22) can be computed analytically:

$$\left\langle \boldsymbol{k}_\perp^2 \right\rangle(x) = \frac{g_1(x) + 2\lambda g_1^2(x)}{1 + \lambda g_1(x)}, \qquad \left\langle \boldsymbol{P}_\perp^2 \right\rangle(z) = \frac{g_3^2(z) + 2\lambda_F g_4^3(z)}{g_3(z) + \lambda_F g_4^2(z)}. \tag{3.25}$$

The form used for $g_1(x)$ and $g_{3,4}(z)$ was initially introduced in Ref. [71], where a model with a single Gaussian was chosen for TMDs. The Gaussians used the average momenta $\left\langle \boldsymbol{k}_\perp^2 \right\rangle(x)$ and $\left\langle \boldsymbol{P}_\perp^2 \right\rangle(z)$ as widths. At the fixed values $\hat{x}$ and $\hat{z}$ their expressions reduce to the normalization only. In this way, the normalizations are fixed at the values assumed at $\hat{x}$ and $\hat{z}$, connecting them to a real value and allowing us to check if they reproduce the momenta measured experimentally. The same idea inspired the definition of $g_1(x)$ and $g_{3,4}(z)$, however in their case the relation between normalization and average momenta is not so straight-forward, but follows Eq. (3.25). This choice for the parametrization is arbitrary and while it could influence the results for the free parameters, it does not change the final result for the Gaussian widths.

In total our parametrization for the TMDs uses 11 different free parameters:

- 4 parameters to describe the PDFs (N_1, α, σ, λ)
- 6 parameters to determine the two Gaussians that constitute the FFs (N_3,N_4, β, δ, γ, λ_F)

- 1 additional parameter characterizes the nonperturbative part of the TMD evolution.

In the present analysis these parameters do not have a dependence on the flavors of the constituents.

3.3 Data analysis

In this section we review the features of the data sets considered for each process. We will also discussed the kinematic cuts applied to the data and justify why they are necessary.

Table 3.1 refers to the measurements for SIDIS off a proton target (HERMES experiment) and presents their kinematic ranges. Tables 3.2 to 3.4 present the characteristics of data for SIDIS off deuteron (HERMES and COMPASS experiments), Drell–Yan events at low energy, and Z boson production respectively. If not specified otherwise, the theoretical formulas are computed at the average values of the kinematic variables in each bin.

3.3.1 Semi-inclusive DIS data

The data for semi-inclusive DIS are taken from HERMES [117] and COMPASS [118] collaborations. Both data sets have already been analyzed in previous works, e.g., Refs. [72, 119], however they have never been fitted together, including also the contributions deriving from TMD evolution.

The application of the TMD formalism to SIDIS depends on the capability of identifying the current fragmentation region. This task has been recently discussed in Ref. [56], where the authors point out a possible overlap among different fragmentation regions when the hard scale Q is sufficiently low. For the moment, we do not tackle this problem and we leave it to future studies. As described in Tabs. 3.1 and 3.2, we identify the current fragmentation region operating a cut on z only, namely $0.2 < z < 0.74$.

Another requirement for the applicability of TMD factorization is the presence of two separate scales in the process, which allows the identification of a hard and a soft part of the interaction. In SIDIS, those are the energy Q^2 and transverse momentum P_{hT}^2, which should satisfy the condition $P_{hT}^2 \ll Q^2$, or more precisely $P_{hT}^2/z^2 \ll Q^2$. We implement this condition by selecting the data that satisfy $P_{hT} < \min[0.2\ Q, 0.7\ Qz] + 0.5$ GeV. With this choice, P_{hT}^2 is always smaller than $Q^2/3$, but in a few bins (at low Q^2 and z) P_{hT}^2/z^2 may become larger than Q^2. The applicability of TMD factorization in this case could be questioned.

However, as we will explain further in Sec. 3.5.3, the best fit parameters that we obtained using this selection rules give a good description of the phenomena over a wide region of P_{hT}, while at the same time performing very well in a restricted region, where TMD factorization is certainly valid. All these choices are summarized in Tabs. 3.1 and 3.2.

HERMES data

HERMES hadron multiplicities are measured in a fixed–target experiment, using a 27.6 GeV lepton beam which collides on a hydrogen (p) or deuterium (D) gas target. The data set consist in a total of 2688 points.

They are grouped in bins of (x, z, Q^2, P_{hT}) with the values of the binnning of (x, Q^2) ranging from about $(0.023, 1.25\ \text{GeV}^2)$ to $(0.6, 9.2\ \text{GeV}^2)$. The collinear energy fraction z assumes values in the interval in $0.1 \leq z \leq 0.8$, while the transverse momentum of the detected hadron covers the interval $0.1\ \text{GeV} \leq |P_{hT}| \leq 1.2\ \text{GeV}$.

The peculiarity of HERMES SIDIS experiment lies in the ability of its detector to distinguish between pions and kaons in the final state, in addition to determining their momenta and charges. This setup allows to consider eight different combinations of target (p, D) and detected charged hadron ($\pi^{\pm}$, $K^{\pm}$).

The HERMES collaboration published two distinct group of data, characterized by the inclusion or subtraction of the vector meson contribution. In our work we considered only the data set where this contribution has been subtracted. We made this choice in order to avoid higher-twist contributions, however, at the moment we are not completely sure how much they are affected by vector meson contributions. We observe that, in any case, there is not much difference in the data at low and medium z.

Compass data

The COMPASS collaboration extracted multiplicities produced in SIDIS off a deuteron (^{6}LiD) target [118]. Differently from the HERMES experiment, the hadrons in the final state are not identified, but only separated into negatively and positively charged. The number of data points is an order of magnitude higher compared to the HERMES experiment. The data are organized in multidimensional bins of (x, z, Q^2, P_{hT}), which cover a range in (x, Q^2) from about $(0.005, 1.11\ \text{GeV}^2)$ to $(0.09, 7.57\ \text{GeV}^2)$ and the interval $0.2 \leq z \leq 0.8$. The multiplicities published by COMPASS are affected by normalization errors (see the *erratum* to Ref. [118]). In order to avoid this issue, we divide the data in each bin in (x, z, Q^2) by the data point with the lowest P_{hT}^2 in the bin. As a result, we define the *normalized* multiplicity as

$$m_{\text{norm}}(x, z, \boldsymbol{P}_{hT}^2, Q^2) = \frac{m_N^h(x, z, \boldsymbol{P}_{hT}^2, Q^2)}{m_N^h(x, z, \min[\boldsymbol{P}_{hT}^2], Q^2)}\ , \tag{3.26}$$

where the multiplicity m_N^h is defined in Eq. (3.1). When fitting normalized multiplicities, the first data point of each bin is considered as a fixed constraint and excluded from the degrees of freedom. Alternatively, we could consider the normalizations as extra free parameters, while maintaining the total number of points. The final number of d.o.f. would turn out to be the same.

Recently COMPASS collaboration has published a new semi-inclusive measurement of charged hadron multiplicities from muon-deuteron DIS [120]. Comparing this dataset to the one used in our previously discussed analysis [118], we observe that they cover a wider kinematic range, reaching in x up to 0.4 instead of 0.12, in Q^2 up to 81 GeV^2 instead of 10 GeV^2 and in P_{hT}^2 up to 3 GeV^2 instead of about 1 GeV^2. Even if the data points are reduced in quantity they have a much higher statistics, leading to a significantly decreased systematic uncertainties on the normalization of the P_{hT}^2-integrated multiplicities. This feature could be an important improvement for our analysis, considering that at the moment it seems to be necessary to normalize each x, z, Q^2 bin of COMPASS data. We conducted a preliminary study of this new data set to test its agreement with the result obtained in our global fit. The initial outcomes suggest that they are compatible, however further consideration on the normalization and the description of the z dependence are necessary for a future improved global fit.

3.3.2 Low-energy Drell–Yan data

Our global analysis includes Drell–Yan events collected by fixed-target experiments at low-energy. These data sets have been considered also in previous works, e.g., in Ref. [121, 99, 100, 122]. We used data sets from the E288 experiment [123], which measured the invariant dimuon cross section $Ed^3\sigma/d^3q$ for the production of $\mu^+\mu^-$ pairs from the collision of a proton beam with a fixed target, either composed of Cu or Pt. The measurements were performed using proton beam energies of 200, 300 and 400 GeV, producing three different data sets. Their respective center of mass energies are $\sqrt{s} = 19.4, 23.8, 27.4$ GeV. We also included the set of measurements $Ed^3\sigma/d^3q$ from E605 [124], extracted from the collision of a proton beam with an energy of 800 GeV ($\sqrt{s} = 38.8$ GeV) with a copper target .

The explored Q values are higher compared to the SIDIS case, as can be seen in Table 3.3. E288 provides data at fixed rapidity, whereas E605 provides data at fixed $x_F = 0.1$. We can apply TMD factorization if $q_T^2 \ll Q^2$, where q_T is the transverse momentum of the intermediate electroweak boson, reconstructed from the kinematics of the final–state leptons. We choose $q_T < 0.2\ Q + 0.5$ GeV.

As suggested in Ref. [123], we consider the target nuclei as an incoherent ensemble composed 40% by protons and 60% by neutrons. We considered eventual nuclear corrections to the PDFs to be negligible at our level of accuracy.

As we already observed, results from E288 and E605 experiments are reported as $\frac{Ed^3\sigma}{d^3q}$; this variable is related to the differential cross section of Eq. (3.6) in the following way:

$$\frac{Ed^3\sigma}{d^3q} = \frac{d^3\sigma}{d\phi d\eta q_T dq_T} \Rightarrow \frac{d^2\sigma}{\pi d\eta d(q_T^2)}, \tag{3.27}$$

where ϕ is the polar angle of q_T and the third term is the average over ϕ. Therefore, the invariant dimuon cross section can be obtained from Eq. (3.6) integrating over Q^2 and adding a factor $1/\pi$ to the result

$$\frac{Ed^3\sigma}{d^3q} = \frac{1}{\pi}\int dQ^2 \frac{d\sigma}{dQ^2 dq_T^2 d\eta}. \tag{3.28}$$

Numerically we checked that integrating in Q^2 only the prefactor σ_q^γ in Eq. (3.7) introduces only a negligible error in the theoretical estimates. We also assume that $\alpha_{\rm em}$ does not change within the experimental bin. Therefore, for Drell–Yan we obtain

$$\frac{1}{\pi}\int dQ^2 \frac{d\sigma}{dQ^2 dq_T^2 d\eta} \approx \frac{4\pi\alpha_{\rm em}^2}{3s} \ln\left(\frac{Q_f^2}{Q_i^2}\right) F_{UU}^1. \tag{3.29}$$

where $Q_{i,f}$ are the lower and upper values in the experimental bin.

3.3.3 Z-boson production data

In order to reach higher Q and q_T values, we also consider Z boson production measured in collider experiments at Tevatron. We analyze data from the CDF and D0 collaborations, collected during Tevatron Run I [125, 126] at $\sqrt{s} =$ 1.8 TeV and Run II [127, 128] at $\sqrt{s} = 1.96$ TeV. CDF and D0 groups studied the differential cross section for the production of an e^+e^- pair through an intermediate Z vector boson from the collision of a proton and an antiproton, namely $p\bar{p} \to Z \to e^+e^- + X$.

Naturally, the invariant mass distribution peaks at the Z-pole, $Q \approx M_Z$, while the transverse momentum of the exchanged Z ranges in $0 < q_T <$ 20 GeV.

We use the same kinematic cut applied to Drell–Yan events: $q_T < 0.2\ Q + 0.5$ GeV. This choice produces the same cut for every data point at $q_T = 18.7$ GeV, since Q is fixed to M_Z.

The observable measured in CDF and D0 is

$$\begin{aligned}\frac{d\sigma}{dq_T} &= \int dQ^2 d\eta 2q_T \frac{d\sigma}{dQ^2 dq_T^2 d\eta} \\ &\approx \frac{\pi^2\alpha_{\rm em}}{s\sin^2\theta_W\cos^2\theta_W} B_R(Z\to\ell^+\ell^-) 2q_T \int d\eta F_{UU}^1,\end{aligned} \tag{3.30}$$

except in the case of D0 Run II, for which the published data refer to $1/\sigma \times d\sigma/dq_T$. In order to work with the same observable, we multiply the D0 Run II data by the total cross section of the process $\sigma_{exp} = 255.8 \pm 16$ pb [129]. To take into account the error introduced by this correction, we add in quadrature the uncertainties of the total cross section and of the published data.

We normalize our functional form with the factors listed in Table 3.4. These are the same normalization factors used in Ref. [122], computed by comparing

the experimental total cross section with the theoretical results based on the code of Ref. [130]. These factors are not entirely consistent with our formulas. In fact, as we will discuss in Sec. 3.5.3 a 5% increase in these factors would improve the agreement with data, without affecting the TMD parameters.

	HERMES $p \to \pi^+$	HERMES $p \to \pi^-$	HERMES $p \to K^+$	HERMES $p \to K^-$
Reference	[117]			
Cuts	$Q^2 > 1.4$ GeV^2 $0.20 < z < 0.74$ $P_{hT} < \text{Min}[0.2\ Q, 0.7\ Qz] + 0.5$ GeV			
Points	190	190	189	187
Max. Q^2	9.2 GeV^2			
x range	$0.04 < x < 0.4$			

Table 3.1: SIDIS proton-target data (HERMES experiment).

	HERMES $D \to \pi^+$	HERMES $D \to \pi^-$	HERMES $D \to K^+$	HERMES $D \to K^-$	COMPASS $D \to h^+$	COMPASS $D \to h^-$
Reference	[117]				[118]	
Cuts	$Q^2 > 1.4$ GeV^2 $0.20 < z < 0.74$ $P_{hT} < \text{Min}[0.2\ Q, 0.7\ Qz] + 0.5$ GeV					
Points	190	190	189	189	3125	3127
Max. Q^2	9.2 GeV^2				10 GeV^2	
x range	$0.04 < x < 0.4$				$0.005 < x < 0.12$	

Table 3.2: SIDIS deuteron-target data (HERMES and COMPASS). For COMPASS data the observable considered is m_{norm}, defined in Eq. (3.26)

.

3.4 The replica method

In our minimization we apply the *replica method* to deal with the error analysis of our theoretical predictions.

	E288 200	E288 300	E288 400	E605
Reference	[123]	[123]	[123]	[124]
Cuts	$q_T < 0.2\ Q + 0.5$ GeV			
Points	45	45	78	35
$\sqrt{s}$	19.4 GeV	23.8 GeV	27.4 GeV	38.8 GeV
Q range	4-9 GeV	4-9 GeV	5-9, 11-14 GeV	7-9, 10.5-11.5 GeV
Kin. var.	η=0.40	η=0.21	η=0.03	$x_F = 0.1$

Table 3.3: Low energy Drell–Yan data collected by the E288 and E605 experiments at Tevatron, with different center-of-mass energies.

	CDF Run I	D0 Run I	CDF Run II	D0 Run II
Reference	[125]	[126]	[127]	[128]
Cuts	$q_T < 0.2\ Q + 0.5$ GeV $= 18.7$ GeV			
Points	31	14	37	8
$\sqrt{s}$	1.8 TeV	1.8 TeV	1.96 TeV	1.96 TeV
Normalization	1.114	0.992	1.049	1.048

Table 3.4: Z boson production data collected by the CDF and D0 experiments at Tevatron, with different center-of-mass energies.

The replica method is a Monte Carlo approach similar to Refs. [131, 72, 132] and inspired by the work on collinear parton distribution extraction of the Neural-Network PDF (NNPDF) collaboration, e.g. in Refs. [133, 134, 135]. The steps necessary for the replica methodology are summarized in Fig. 3.3, where it is applied on some arbitrary bins of HERMES . The general idea consists in creating $\mathcal{M}$ replicas of the original data points and then conduct the minimization on each different replica, to create a confidence band for our theoretical predictions. We start from the original data points with their error band, grouped in different kinematic bins (Fig. 3.3(a)). To create each single replica (denoted by the index r), every data point i is shifted by a Gaussian noise with the same variance as the measurement (b). This procedure is repeated $\mathcal{M}$ times choosing different random values for the Gaussian noise, obtaining different groups of replicated data. Each one of these replica, therefore, represents a possible outcome of an independent experimental measurement, which we denote by $m^h_{N,r}(x, z, \boldsymbol{P}^2_{hT}, Q^2)$.

The number of replicas is chosen so that the mean and standard deviation of the set of replicas accurately reproduces the original data points. In this case we checked that 200 replicas are sufficient for the purpose.

The error for each replicated data point is taken to be equal to the error on

the corresponding original data points. This is consistent with the fact that the variance of the $\mathcal{M}$ replicas should reproduce the variance of the original data points.

In the next step, a minimization procedure is applied to each replica separately (c), by searching for the best fit of the following error function:

$$E_r^2(\{p\}) = \sum_i \frac{\left(m^h_{N,r}(x_i, z_i, \boldsymbol{P}^2_{hTi}, Q^2_i) - m^h_{N,\text{theo}}(x_i, z_i, \boldsymbol{P}^2_{hTi}; \{p\})\right)^2}{\left(\Delta m^{h\,2}_{N,\text{stat}} + \Delta m^{h\,2}_{N,\text{sys}}\right)(x_i, z_i, \boldsymbol{P}^2_{hTi}, Q^2_i) + \left(\Delta m^h_{N,\text{theo}}(x_i, z_i, \boldsymbol{P}^2_{hTi})\right)^2}\,. \tag{3.31}$$

The sum runs over the i experimental points, including all species of targets N and final-state hadrons h. We choose different starting points for the fit parameters, in order to better sample the parameter space and avoid stagnation in a local minimum for our predictions.

Another aspect of our parametrization that is changed for every replica, are the values of the collinear fragmentation functions $D_1^{a\to h}$, which are independently modified with a Gaussian noise, different for each z bin, with standard deviation equal to the theoretical error $\Delta D_1^{a\to h}$. In this work we rely on different parametrizations for $D_1^{a\to h}$: for pions we use the DSEHS analysis [107] at NLO in α_S; for kaons we use the DSS parametrization [108] at LO in α_S. The uncertainties $\Delta D_1^{a\to h}$ are estimated from the plots in Ref. [136]; they represent the only source of uncertainty in $\Delta m^h_{N,\text{theo}}$.

Statistical and systematic experimental uncertainties $\Delta m^h_{N,\text{stat}}$ and $\Delta m^h_{N,\text{sys}}$, used in the error function (3.31), are taken directly from the publications of the experimental collaborations. We do not take into account the covariance among different kinematic bins.

The minimization is carried out using a Fortran code that employs the algorithms provided by MINUIT [137].

The final outcome (d) is a set of $\mathcal{M}$ different vectors of best-fit parameters, $\{p_{0r}\}$, $r = 1, \ldots \mathcal{M}$, with which we can calculate any observable, its mean, and its standard deviation. The distribution of these values needs not to be necessarily Gaussian. In fact, in this case the $1-$sigma confidence interval is different from the 68% interval. The latter can simply be computed for each experimental point by rejecting the largest and the lowest 16% of the $\mathcal{M}$ values (e), obtaining for every bin the lower and upper bound of the 68% confidence band for our theoretical predictions.

Although the minimization is performed on the function defined in Eq. (3.31), the agreement of the $\mathcal{M}$ replicas with the original data is expressed in terms of a χ^2 function defined as in Eq. (3.31) but with the replacement $m^h_{N,r} \to m^h_N$, i.e., with respect to the original data set. If the model is able to give a good description of the data, the distribution of the $\mathcal{M}$ values of χ^2/d.o.f. should be peaked around one.

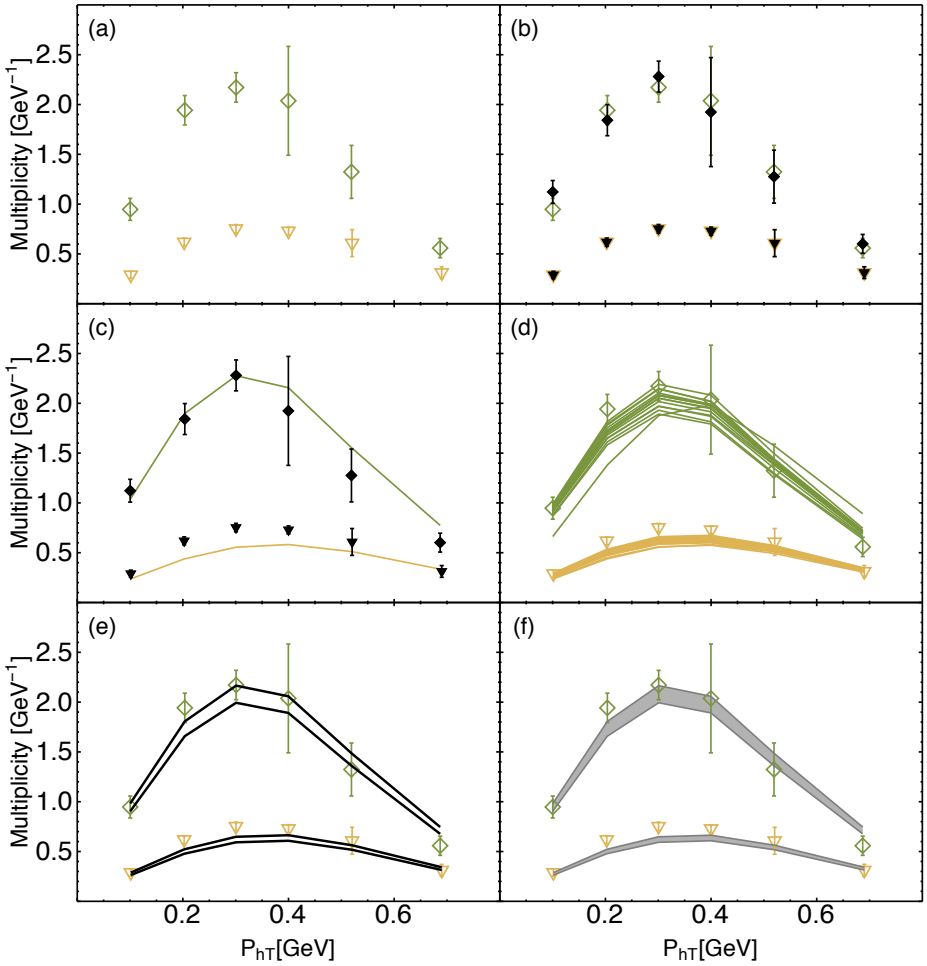


Figure 3.3: HERMES multiplicities for production of pions off a proton and a deuteron for different $\langle x \rangle$, $\langle z \rangle$, and $\langle Q^2 \rangle$ bins as a function of the transverse momentum of the detected hadron P_{hT}. For clarity, each $\langle z \rangle$ bin has been shifted by an offset indicated in the legend.

3.5 Results

Our work aims at simultaneously extracting unpolarized TMDs data sets related to different experiments for the first time. In the past, only fits related either to SIDIS or hadronic collisions have been presented.

In Ref. [72], the authors fitted HERMES multiplicities only (taking into account a total of 1538 points) without taking into account QCD evolution. In that work, a flavor decomposition in transverse momentum of the unpolarized TMDs and an analysis of the kinematic dependence of the intrinsic average

square transverse momenta were presented. In Ref. [119] the Torino group fitted HERMES and COMPASS multiplicities separately (576 and 6284 points respectively), without TMD evolution and introducing an ad-hoc normalization for COMPASS data. A fit of SIDIS data including TMD evolution was performed on measurements by the H1 collaboration of the so-called transverse energy flow [98, 138].

Looking at data from hadronic collisions, Konychev and Nadolsky [100] fitted data of low-energy Drell–Yan events and Z-boson production at Tevatron, taking into account TMD evolution at NLL accuracy (this is the most recent of a series of important papers on the subject [139, 121, 99]). They fitted in total 98 points. Contrary to our approach, Konychev and Nadolsky studied the quality of the fit as a function of $\xi_{\max}$. They found that the best value for $\xi_{\max}$ is 1.5 GeV^{-1} (to be compared to our choice $\xi_{\max} \approx 1.123$ GeV^{-1}, see Sec. 3.2.3). Comparisons of best-fit values in the nonperturbative Sudakov form factors are delicate, since the functional form is different from ours.

In 2014 D'Alesio, Echevarria, Melis, Scimemi performed a fit [122] of Drell–Yan data and Z-boson production data at Tevatron, focusing in particular on the role of the nonperturbative contribution to the kernel of TMD evolution. This is the fit with the highest accuracy in TMD evolution performed up to date (NNLL in the Sudakov exponent and $\mathcal{O}(\alpha_S)$ in the Wilson coefficients). In the same year Echevarria, Idilbi, Kang and Vitev [25] presented a parametrization of the unpolarized TMD that described qualitatively well a limited number of bins of HERMES and COMPASS data, together with Drell–Yan and Z-production data. A similar result was presented by Sun, Isaacson, Yuan and Yuan [140].

In 2017 Scimemi and Vladimirov [141] presented an extraction of the unpolarized TMD parton distribution functions from Drell–Yan and Z-boson production measurements, including the data of E288, Tevatron and LHC experiments. They choose a scheme for dealing with the scale fixation, called ζ-prescription, which tends to minimize the contribution of perturbative logarithms in the large range of scales.

In the following, we detail the results of a fit to the data sets described in Sec. 3.3 with a flavor-independent configuration for the transverse momentum dependence of unpolarized TMDs. In Table 3.5 we present the total χ^2; the number of degrees of freedom (d.o.f.) is given by the number of data points analyzed reduced by the number of free parameters in the error function (including the COMPASS data points which were used as normalization in m_{norm}). The overall quality of the fit is good, with a global $\chi^2/\text{d.o.f.} = 1.55 \pm 0.05$. Uncertainties are computed as the 68% confidence level (C.L.) from the replica methodology.

3.5.1 Agreement between data and theory

The partition of the global χ^2 among SIDIS off a proton, SIDIS off a deuteron, Drell–Yan and Z production events is given in Table 3.6, 3.7, 3.8, 3.9 respec-

Points	Parameters	χ^2	χ^2/d.o.f.
8059	11	12629 ± 363	1.55 ± 0.05

Table 3.5: Total number of points analyzed, number of free parameters and χ^2 values.

tively.

Semi-inclusive DIS

For SIDIS at HERMES off a proton, most of the contribution to the χ^2 comes from events with a π^+ in the final state. In Ref. [72] the high χ^2 was attributed to the poor agreement between experiment and theory at the level of the collinear multiplicities. In this work we use a newer parametrization of the collinear FFs (DSEHS [107]), based on a fit that includes HERMES collinear pion multiplicities. In spite of this improvement, the contribution to χ^2 from HERMES data is higher then in Ref. [72], because the present fit includes data from other experiments (HERMES represents less than 20% of the whole data set). The bins with the worst agreement are at low Q^2. As we will discuss in Sec. 3.5.3, we think that the main reason for the large χ^2 at HERMES is a normalization difference. This may also be due to the fact that we are computing our theoretical estimates at the average values of the kinematic variables, instead of integrating the multiplicities in each bin. Kaon multiplicities have in general a lower χ^2, due to the bigger statistical errors and the large uncertainties for the kaon FFs.

	HERMES $p \to \pi^+$	HERMES $p \to \pi^-$	HERMES $p \to K^+$	HERMES $p \to K^-$
Points	190	190	189	187
χ^2/points	4.83 ± 0.42	2.47 ± 0.28	0.91 ± 0.14	0.82 ± 0.17

Table 3.6: Number of points analyzed and χ^2 values for SIDIS off a proton target.

For pion production off a deuteron at HERMES the χ^2 is lower with respect to the production off a proton, but still compatible within uncertainties. For kaon production off a deuteron the χ^2 is higher with respect to the scattering off a proton. The difference is especially large for K^-.

SIDIS at COMPASS involves scattering off deuteron only, $D \to h^\pm$, and we identify $h \equiv \pi$. The quality of the agreement between theory and COMPASS data is better than in the case of pion production at HERMES. This depends on at least two factors: first, our fit is essentially driven by the COMPASS data,

which represent about 75% of the whole data set; second, the observable that we fit in this case is the normalized multiplicity, defined in Eq. (3.26). This automatically eliminates most of the discrepancy between theory and data due to normalization.

	HERMES $D \to \pi^+$	HERMES $D \to \pi^-$	HERMES $D \to K^+$	HERMES $D \to K^-$
Points	190	190	189	189
χ^2/points	3.46 ± 0.32	2.00 ± 0.17	1.31 ± 0.26	2.54 ± 0.57
	COMPASS $D \to h^+$		COMPASS $D \to h^-$	
Points	3125		3127	
χ^2/points	1.11 ± 0.03		1.61 ± 0.04	

Table 3.7: Number of points analyzed and χ^2 values for SIDIS off a deuteron target.

In section 3.7 the plots showing the agreement between experimental data and our predictions are collected. Figure 3.4 presents the agreement between the theoretical formula in (3.1) and the HERMES multiplicities for production of pions off a proton and a deuteron. Different $\langle x \rangle$, $\langle z \rangle$ and $\langle Q^2 \rangle$ bins are displayed as a function of the transverse momentum of the detected hadron P_{hT}. The grey bands are an envelope of the 200 replicas of best-fit curves. For every point in P_{hT} we apply a 68% C.L. selection criterion. Points marked with different symbols and colors correspond to different $\langle z \rangle$ values. There is a strong correlation between $\langle x \rangle$ and $\langle Q^2 \rangle$ that does not allow us to explore the x and Q^2 dependence of the TMDs separately. Studying the contributions to the χ^2/points as a function of the kinematics, we notice that the $\chi^2(Q^2)$ tends to improve as we move to higher Q^2 values, where the kinematic approximations of factorization are more reliable. Moreover, usually the $\chi^2(z)$ increases at lower z values.

Figure 3.5 has same content and notation as in Fig. 3.4 but for kaons in the final state. In this case, the trend of the agreement as a function of Q^2 is not as clear as for the case of pions: good agreement is found also at low Q^2.

In Fig. 3.6 we present COMPASS normalized multiplicities (see Eq. (3.26)) for production of negative hadrons off a deuteron for different $\langle x \rangle$, $\langle z \rangle$, and $\langle Q^2 \rangle$ bins as a function of the transverse momentum of the detected hadron P_{hT}. The open marker around the first P_{hT} point in each panel indicates that the first value is fixed and not fitted. The correlation between x and Q^2 is less strong than at HERMES and this allows us to study different $\langle x \rangle$ bins at fixed $\langle Q^2 \rangle$. For the highest Q^2 bins, the agreement is good for all $\langle x \rangle$, $\langle z \rangle$ and P_{hT}^2.

In bins at lower Q^2, the descriptions gets worse, especially at low and high z. We observe even though normalized, the low-x high-z data are not always well described, this is probably due to the first data point of the bin used as normalization, which in these specific region does not follow the general trend of the rest of the data points of the bin. This problem could be avoided using a different strategy for normalization, such as using the data in the middle of the bin to normalize instead of the initial one, or such as a normalization factor which is derived through a fit over every data point in the bin

For fixed $\langle Q^2 \rangle$ and high $\langle z \rangle$, a good agreement is recovered moving to higher $\langle x \rangle$ bins (see, e.g., the third line from the top in Fig. 3.7).

Fig. 3.7 has same content and notation as in Fig. 3.7, but for $h^- \equiv \pi^-$. The same comments on the agreement between theory and the data apply.

Drell–Yan and Z production

The low energy Drell–Yan data collected by the E288 and E605 experiments at Fermilab have large error bands (see Fig. 3.8). This is why the χ^2 values in Table 3.8 are rather low compared to the other data sets.

The agreement is also good for Z boson production, see Table 3.9. The statistics from Run-II is higher, which generates smaller experimental uncertainties and higher χ^2, especially for the CDF experiment.

	E288 [200]	E288 [300]	E288 [400]	E605
Points	45	45	78	35
χ^2/points	0.99 ± 0.09	0.84 ± 0.10	0.32 ± 0.01	1.12 ± 0.08

Table 3.8: Number of points analyzed and χ^2 values for fixed-target Drell–Yan experiments at low energy. The labels in square brackets were introduced in Sec. 3.3.2.

	CDF Run I	D0 Run I	CDF Run II	D0 Run II
Points	31	14	37	8
χ^2/points	1.36 ± 0.00	1.11 ± 0.02	2.00 ± 0.02	1.73 ± 0.01

Table 3.9: Number of points analyzed and χ^2 values for Z boson production at Tevatron.

Fig. 3.8 displays the cross section for DY events differential with respect to the transverse momentum q_T of the virtual photon, its invariant mass Q^2 and rapidity y. As for the case of SIDIS, the grey bands are the 68% C.L. envelope of the 200 replicas of the fit function. The four panels represents

different values for the rapidity y or x_F (see Eq. (3.9)). In each panel, we have plots for different Q^2 values. The lower is Q, the less points in q_T we fit (see also Sec. 3.3.2). The hard scale lies in the region $4.5 < \langle Q \rangle < 13.5$ GeV. This region is of particular importance, since these "moderate" Q values should be high enough to safely apply factorization and, at the same time, low enough in order for the nonperturbative effects to not be shaded by transverse momentum resummation.

In Fig. 3.9 we compare the cross section differential with respect to the transverse momentum q_T of the virtual Z (namely Eq. (3.6) integrated over η) with data from CDF and D0 at Tevatron Run I and II. Due to the higher $Q = M_Z$, the range explored in q_T is much larger compared to all the other observables considered. The tails of the distributions deviate from a Gaussian behavior, as it is also evident in the bins at higher Q^2 in Fig. 3.8. The band from the replica methodology in this case is much narrower, due to the reduced sensitivity to the intrinsic transverse momenta at $Q = M_Z$ and to the limited range of best-fit values for the parameter g_2, which controls soft-gluon emission. As an effect of TMD evolution, the peak shifts from ~ 1 GeV for Drell–Yan events in Fig. 3.8 to ~ 5 GeV in Fig. 3.9. The position of the peak is affected both by the perturbative and the nonperturbative part of the Sudakov exponent (see Sec. 3.2.3 and [71]). Most of the contributions to the χ^2 comes from normalization effects and not from the shape in q_T (see Sec. 3.5.3).

3.5.2 Transverse momentum dependence at 1 GeV

The variables $\xi_{\min}$ and $\xi_{\max}$ delimit the range in ξ_T where transverse momentum resummation is computed perturbatively. The g_2 parameter enters the nonperturbative Sudakov exponent and quantifies the amount of transverse momentum due to soft gluon radiation that is not included in the perturbative part of the Sudakov form factor. As already explained in Sec. 3.2.3, in this work we fix the value for $\xi_{\min}$ and $\xi_{\max}$ in such a way that at $Q = 1$ GeV the unpolarized TMDs coincide with their nonperturbative input. We leave g_2 as a fit parameter.

Table 3.10 summarizes the chosen values of $\xi_{\min}$, $\xi_{\max}$ and the best-fit value for g_2. The latter is given as an average with 68% C.L. uncertainty computed over the set of 200 replicas. We also quote the results obtained from replica 105, since its parameters are very close to the mean values of all replicas. We obtain a value $g_2 = 0.13 \pm 0.01$, smaller than the value ($g_2 = 0.184 \pm 0.018$) obtained in Ref. [100], where however no SIDIS data was taken into consideration, and smaller than the value ($g_2 = 0.16$) chosen in Ref. [25]. We stress however that our prescriptions involving both $\xi_{\min}$ and $\xi_{\max}$ are different from previous works.

Table 3.11 collects the best-fit values of parameters in the nonperturbative part of the TMDs at $Q = 1$ GeV (see Eqs. (3.19) and (3.20)); as for g_2, we give the average value over the full set of replicas and the standard deviation based on a 68% C.L. (see Sec. 3.4), and we also quote the value of replica 105.

In Fig. 3.12 we compare different extractions of partonic transverse momenta. The horizontal axis shows the value of the average transverse momentum squared for the incoming parton, $\langle \boldsymbol{k}_\perp^2 \rangle(x = 0.1)$ (see Eq. (3.25)). The vertical axis shows the value of $\langle \boldsymbol{P}_\perp^2 \rangle(z = 0.5)$, the average transverse momentum squared acquired during the fragmentation process (see Eq. (3.25)). The white square (label 1) indicates the average values of the two quantities obtained in the present analysis at $Q^2 = 1$ GeV^2. Each black dot around the white square is an outcome of one replica. The red region around the white square contains the 68% of the replicas that are closest to the average value. The same applies to the white circle and the orange region around it (label 2), related to the flavor-independent version of the analysis in Ref. [72], obtained by fitting only HERMES SIDIS data at an average $\langle Q^2 \rangle = 2.4$ GeV^2 and neglecting QCD evolution. A strong anticorrelation between the transverse momenta is evident in this older analysis. In our new analysis, the inclusion of Drell–Yan and Z production data adds physical information about TMD PDFs, free from the influence of TMD FFs. This reduces significantly the correlation between $\langle \boldsymbol{k}_\perp^2 \rangle(x = 0.1)$ and $\langle \boldsymbol{P}_\perp^2 \rangle(z = 0.5)$. The 68% confidence region is smaller than in the older analysis. The average values of $\langle \boldsymbol{k}_\perp^2 \rangle(x = 0.1)$ are similar and compatible within error bands. The values of $\langle \boldsymbol{P}_\perp^2 \rangle(z = 0.5)$ in the present analysis turn out to be larger than in the older analysis, an effect that is due mainly to COMPASS data. It must be kept in mind that the two analyses lead also to differences in the x and z dependence of the transverse momentum squared. This dependence is shown in Fig. 3.10 (a) for $\langle \boldsymbol{k}_\perp^2 \rangle(x)$ and Fig. 3.10 (b) for $\langle \boldsymbol{P}_\perp^2 \rangle(z)$. The bands are computed as the 68% C.L. envelope of the full sets of curves from the 200 replicas. Comparison with other extractions are presented and the legend is detailed in the caption of Fig. 3.12.

In Fig. 3.11 we present the same results with the complete set of curves extracted from the 200 replicas minimization. This plots can be useful also to understand the relation between the complete space covered by the replicas and the 68% C.L bands.

	$\xi_{\max}$ [GeV^{-1}] (fixed)	$\xi_{\min}$ [GeV^{-1}] (fixed)	g_2 [GeV^2]
All replicas	$2e^{-\gamma_E}$	$2e^{-\gamma_E}/Q$	0.13 ± 0.01
Replica 105	$2e^{-\gamma_E}$	$2e^{-\gamma_E}/Q$	0.128

Table 3.10: Values of parameters common to TMD PDFs and TMD FFs.

3.5.3 Stability of our results

The value that we obtained for the reduced χ^2 does not necessarily means a good agreement between measurements and predictions, if we consider the large amount of data points included. However, we are mostly interested in

TMD PDFs	N_1 [GeV^2]	α	σ		λ [GeV^{-2}]	
All replicas	0.28 ± 0.06	2.95 ± 0.05	0.17 ± 0.02		0.86 ± 0.78	
Replica 105	0.285	2.98	0.173		0.39	
TMD FFs	N_3 [GeV^2]	β	γ	δ	λ_F [GeV^{-2}]	N_4 [GeV^2]
All replicas	0.21 ± 0.01	1.65 ± 0.49	2.28 ± 0.46	0.14 ± 0.07	5.50 ± 1.23	0.03 ± 0.01
Replica 105	0.212	2.10	2.52	0.094	5.29	0.033

Table 3.11: 68% confidence intervals of best-fit values for parametrizations of TMDs at $Q = 1$ GeV.

determining the shape of the TMDs in the widest possible kinematic region. In this section we will observe how we can achieve a value for the reduced χ^2 closer to 1, simply introducing appropriate normalizations or reducing the kinematic area, without modifying the parameters found in the original fit, which define the TMDs shape.

Moreover, we will test the stability of our parametrization against some of the choices we made in our default fit. Instead of repeating the fitting procedure with different choices, we limit ourselves to checking how the χ^2 of a single replica is affected by the modifications.

As starting point for this exploration and subsequent ones, we choose replica 105, which, as discussed above, is one of the most representative among the whole replica set. The global χ^2/d.o.f. of replica 105 is 1.51. We keep all parameters fixed, without performing any new minimization, and we compute the χ^2/d.o.f. after the modifications described in the following.

First of all, we analyze HERMES data with the same strategy as COMPASS, i.e., we normalize HERMES data to the value of the first bin in P_{hT}. In this case, the global χ^2/d.o.f. reduces sharply to 1.27. The partial χ^2 for the different SIDIS processes measured at HERMES are shown in Table 3.12. This confirms that normalization effects are the main contribution to the χ^2 of SIDIS data and have minor effects on TMD-related parameters. In fact, even if we perform a new fit with this modification, the χ^2 does not improve significantly and parameters are compatible with the ones of the original fit. However, we decided not to include a normalization for HERMES data in the final results because, differently from COMPASS data, they are not affected by an error in the normalization.

We consider the effect of changing the normalization of the Z-boson data: if we increase the normalization factors quoted in the last row of Table 3.4 by 5%, the χ^2 quoted in the last row of Table 3.9 drops to 0.66, 0.52, 0.65, 0.68. This effect is also already visible by eye in Fig. 3.9: the theoretical curves are systematically below the experimental data points, but the shape is reproduced very well.

	$p \to \pi^+$	$p \to \pi^-$	$p \to K^+$	$p \to K^-$
Original	5.18	2.67	0.75	0.78
Normalized	1.94	1.13	0.57	0.29
	$D \to \pi^+$	$D \to \pi^-$	$D \to K^+$	$D \to K^-$
Original	3.63	2.31	1.12	2.27
Normalized	1.59	0.80	0.47	0.97

Table 3.12: χ^2/d.o.f. for HERMES data of replica 105 with and without normalization to the value of the first bin in P_{hT}.

We test the sensitivity of our results to the parametrizations adopted for the collinear quark PDFs. The χ^2/d.o.f. varies from its original value 1.51, obtained with the NLO GJR 2008 parametrization [106], to 1.84 using NLO MSTW 2008 [142], and 1.85 using NLO CJ12 [143]. In both cases, the agreement with HERMES and Z boson data is not affected significantly, the agreement with COMPASS data becomes slightly worse, and the agreement with DY data becomes clearly worse.

The choices we made for the parameters of the renormalization scale μ_b can be challenged and we test how alternative values for $\xi_{\max}$ affect the results of replica 105, compared to our default choice $\xi_{\max} = 2e^{-\gamma_E}$ GeV^{-1}. As a first step we set $\xi_{\max} = 0.5$ GeV^{-1} as was done in Ref. [99], and we find similarly a larger value of g_2. Repeating the analysis on a restricted number of replicas, adopting $\xi_{\max} = 0.5$ GeV^{-1}, we find in general a worse agreement compared to our original results. In fact, it has been found that this choice restricts the perturbative part of the calculation of the resummation to a range in ξ_T that is smaller than the range where perturbative methods should still be valid [59]. A more reasonable choice appears to be $\xi_{\max} = 1.5$ GeV^{-1}, chosen for example in Ref. [100]. If we repeat our analysis using this choice, g_2 assumes values compatible with our original ones and we get in general an higher value for χ^2. Finally we tried to remove $\xi_{\min}$ and repeat the minimization, in this case the outcome has a noticeably worse agreement with experimental data. Therefore, $\xi_{\min}$ is an important element of the μ_b parametrization, an it could be a signal that we are exiting the proper region for TMD factorization and we are approaching the region of collinear factorization, especially in SIDIS data at low Q.

An extremely important point is the choice of kinematic cuts. Our default choices are listed in Tabs. 3.1–3.4. We consider also more stringent kinematic cuts on SIDIS data: $Q^2 > 1.5$ GeV^2 and $0.25 < z < 0.6$ instead of $Q^2 > 1.4$ GeV^2 and $0.2 < z < 0.7$, leaving the other ones unchanged. The cut in the high-z region, even if it could be useful for TMD factorization, was made to guarantee that we are in an area which avoids The number of bins with these

cuts reduces from 8059 to 5679 and the χ^2/d.o.f. decreases to the value 1.23. In addition, if we replace the constraint $P_{hT} < \text{Min}[0.2\,Q, 0.7\,Qz] + 0.5$ GeV with $P_{hT} < \text{Min}[0.2\,Q, 0.5\,Qz] + 0.3$ GeV, the number of bins reduces to 3380 and the χ^2/d.o.f. decreases further to 0.96. By adopting the even stricter cut $P_{hT} < 0.2\,Qz$, the number of bins drops to only 477, with a χ^2/d.o.f. =1.02. We can conclude that our fit, obtained by fitting data in an extended kinematic region, where TMD factorization may be questioned, works extremely well also in a narrower region, where TMD factorization is expected to be under control.

When we planned the step of our minimization procedure, after some initial test with different numbers of replicas, we assumed that including 200 replicas in our analysis is enough to reproduce the mean and standard deviation of the original data points. In fact, with this assumption we are able to produce theoretical predictions for the cross sections that are consistent with the shape of the experimental data and contained in the range of their uncertainties. A greater number of replicas would be required if we wanted to reproduce also correlations between data. As a test for this choice, we created again the theoretical bands for our predictions, this time including only the results of 100 replicas. As expected, we found that even with a reduced number we achieve a good description ($\chi^2 = 1.53 \pm 0.05$) where the 68% bands are still contained in the uncertainty range, confirming that our assumption is sufficient to effectively reproduce the original data.

3.6 Visualization of TMDs in momentum space

We have presented our results extraction of TMDs showing the curves for theoretical multiplicities which are directly related to their values and we have shown the shape of the average momenta $k_\perp$ and $P_\perp$ as a function of x and z, respectively. In this section we present some approaches to the visualization of TMD parton distribution functions in momentum space, focusing also on how their shape changes for different values of x and Q^2. In particular visualizing the shift of the distribution peak with different values for the hard scale, could be useful to observe the effects of TMD evolution.

Figure 3.13 shows a density plot for the value of $f_1(x, k; Q^2)$ for the up quark, calculated for $Q^2 = 1$ GeV^2, using the parameters extracted from the usual replica 105. There are three different panels for increasing values of x and through the color gradient it is possible to observe how the shape of the distribution changes with respect to the projections of the momentum k_x and k_y. In Fig. 3.14 is shown the corresponding plot for the TMD fragmentation function with three different slices characterized by increasing values of z.

The shape of the TMD PDF is plotted in Fig. 3.15 for different values of x and Q^2, again using the replica 105 parameters. The peak of each one of this curves is highlighted by a red point, this way we can observe its shift with respect to the variation of the hard scale. We point out that the width of the distribution broadens with the increase of Q^2 and the position of the peak

shifts to higher $k_\perp$.

In order to give an idea of the different estimates of the TMD PDF given by the different replicas, we show in Fig. 3.16 and Fig. 3.17 some density plots for $f_1(x, k; Q^2)$ in momentum space for the up quark, calculated at $Q^2 = 1$ GeV^2. We chose to display in particular replicas that shows particular features compared to the others; For example replica 158 has a larger value of $g_1^a(x)$, while replica 149 has a large value for λ. Moreover, we included replica 149 because it shows an interesting shape for the x-dependence, not found in other replicas. Finally, we included also replica 191 which has the $\chi^2/d.o.f.$ closer to 1. ì We observe that focusing on replicas with different features, we obtain different shapes for the unpolarized TMDs, e.g., replica 158 is wider, while replica 149 is much narrower. The shape of replica 185 instead displays a dip in the middle. Even with their different shapes, all these PDFs give a good agreement with the experimental data.

The corresponding plots for the TMD fragmentation function are shown in Fig. 3.18 and Fig. 3.19, where we chose to use the same replicas shown in the PDF plots. We notice that all the FF shapes shown here present a dip in the middle, due to our choice of having different widths in the two summed Gaussian parametrization.

3.7 Appendix: Plots and Figures

3.7.1 HERMES multiplicities for unpolarized SIDIS

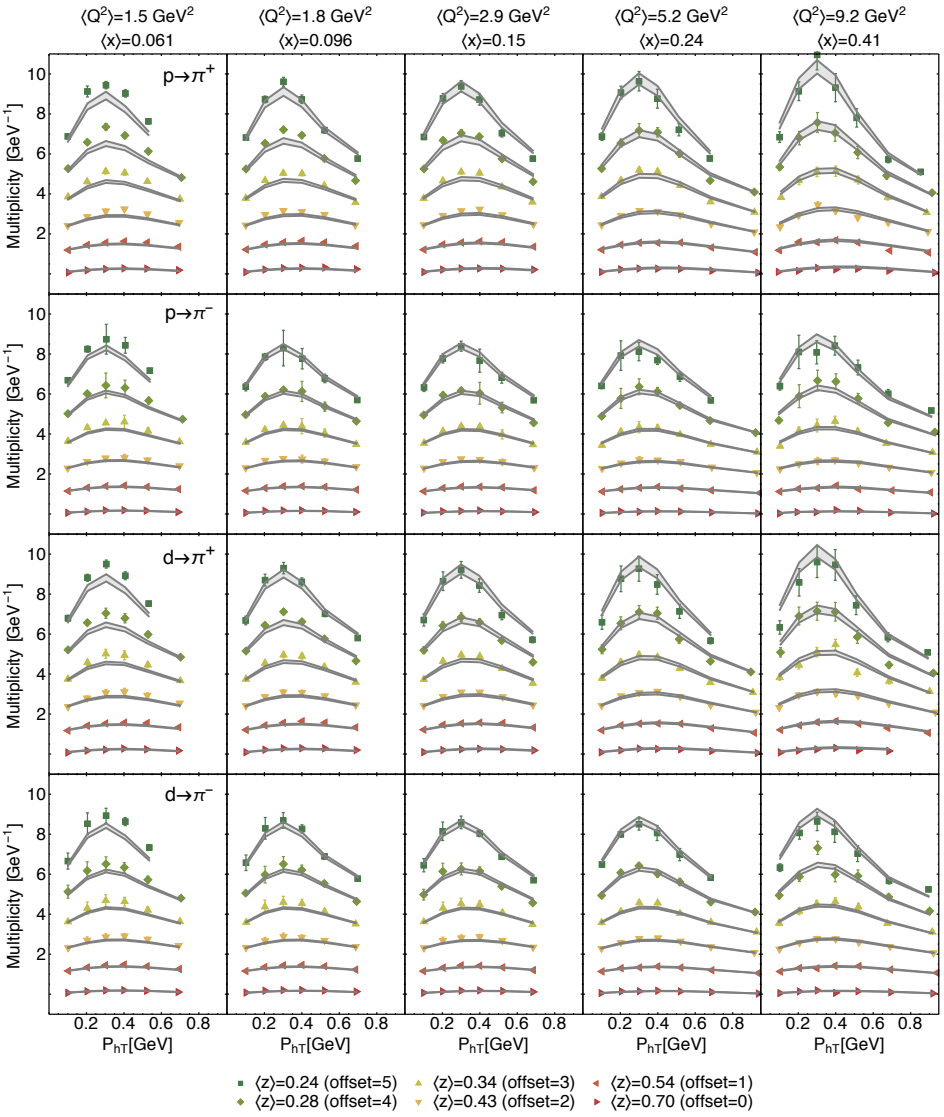


Figure 3.4: HERMES multiplicities for production of pions off a proton and a deuteron for different $\langle x \rangle$, $\langle z \rangle$, and $\langle Q^2 \rangle$ bins as a function of the transverse momentum of the detected hadron P_{hT}. For clarity, each $\langle z \rangle$ bin has been shifted by an offset indicated in the legend.

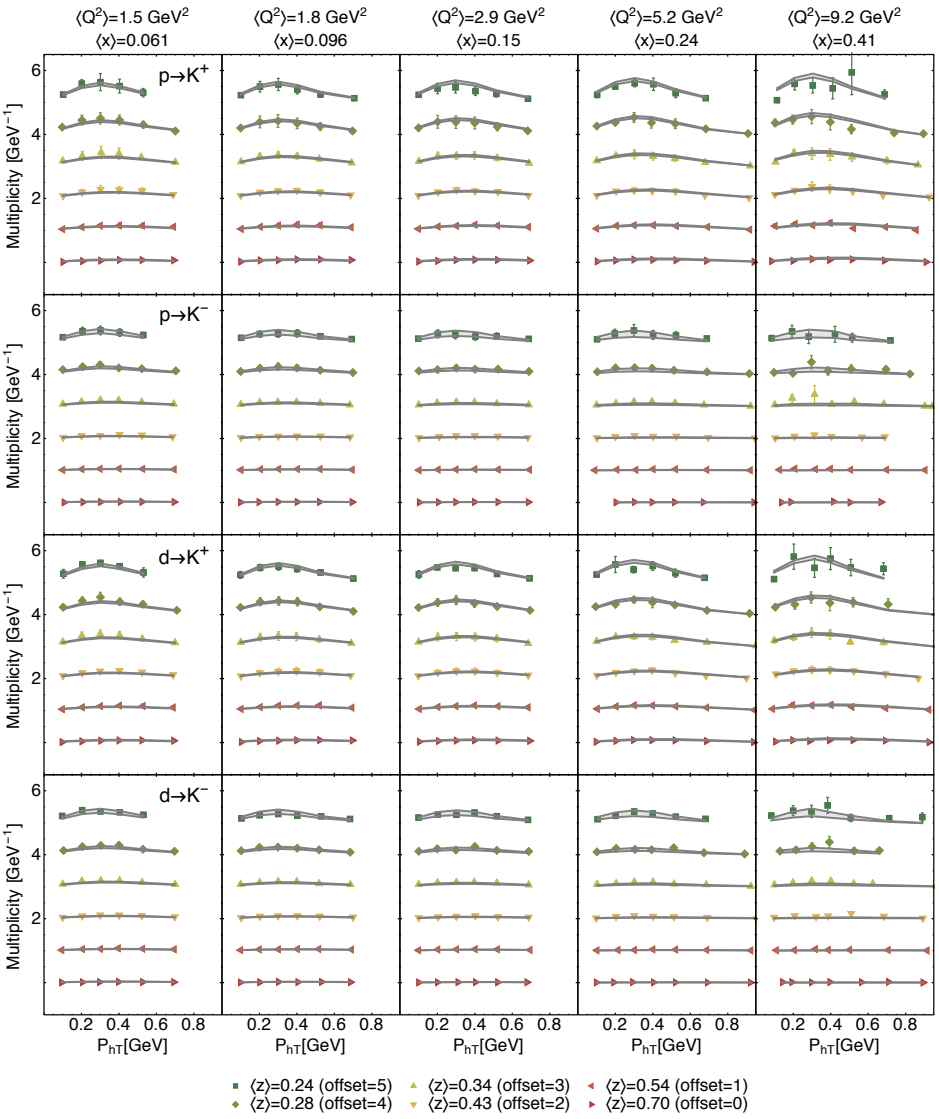


Figure 3.5: HERMES multiplicities for production of kaons off a proton and a deuteron for different $\langle x \rangle$, $\langle z \rangle$, and $\langle Q^2 \rangle$ bins as a function of the transverse momentum of the detected hadron P_{hT}. For clarity, each $\langle z \rangle$ bin has been shifted by an offset indicated in the legend.

3.7.2 COMPASS multiplicities for unpolarized SIDIS

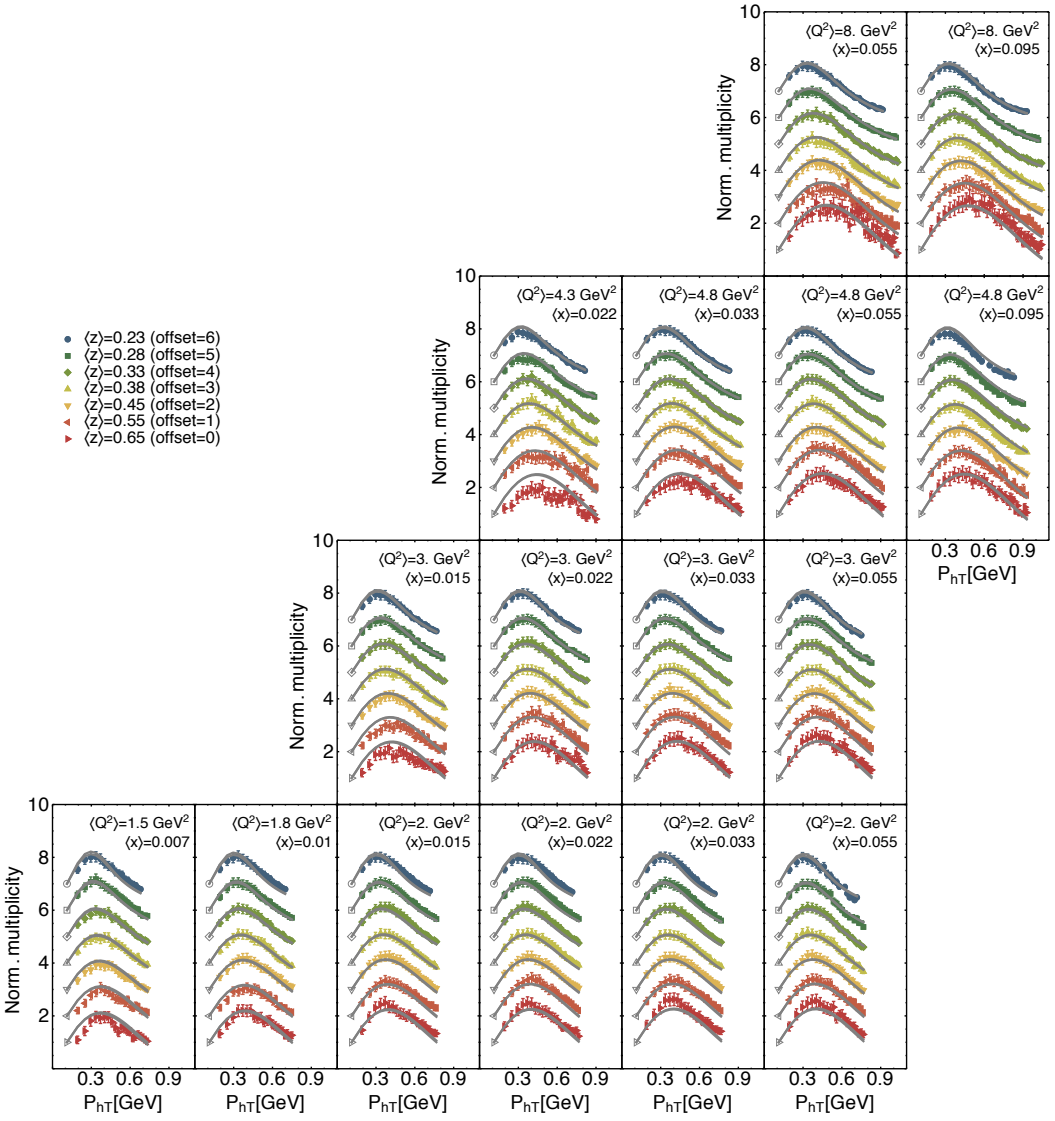


Figure 3.6: COMPASS multiplicities for production of positive hadrons (π^+) off a deuteron for different $\langle x \rangle$, $\langle z \rangle$, and $\langle Q^2 \rangle$ bins as a function of the transverse momentum of the detected hadron P_{hT}. Multiplicities are normalized to the first bin in P_{hT} for each $\langle z \rangle$ value (see Eq. (3.26)). For clarity, each $\langle z \rangle$ bin has been shifted by an offset indicated in the legend.

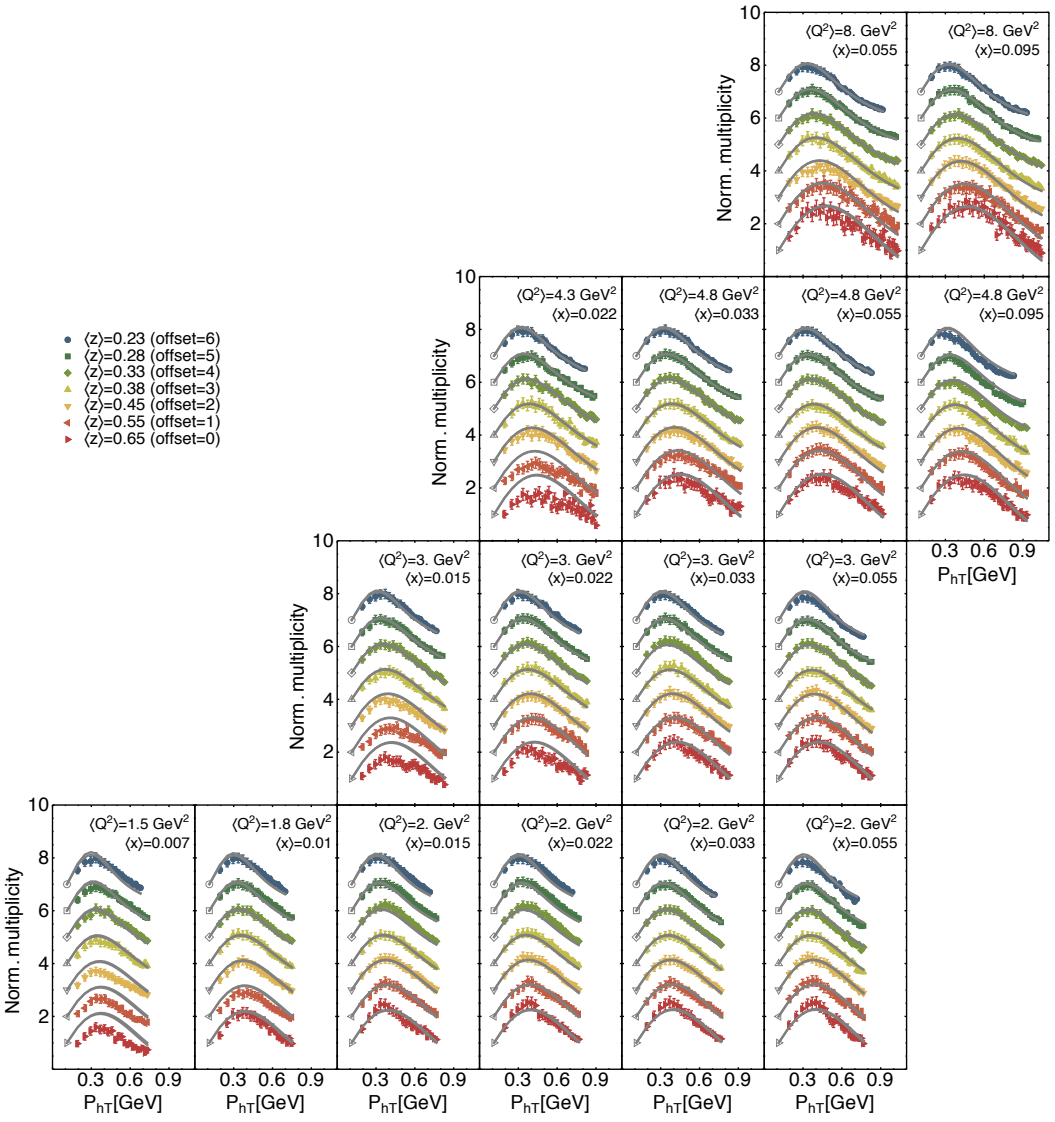


Figure 3.7: COMPASS multiplicities for production of negative hadrons (π^-) off a deuteron for different $\langle x \rangle$, $\langle z \rangle$, and $\langle Q^2 \rangle$ bins as a function of the transverse momentum of the detected hadron P_{hT}. Multiplicities are normalized to the first bin in P_{hT} for each $\langle z \rangle$ value (see Eq. (3.26)). For clarity, each $\langle z \rangle$ bin has been shifted by an offset indicated in the legend.

3.7.3 Cross section for Drell–Yan and Z boson production

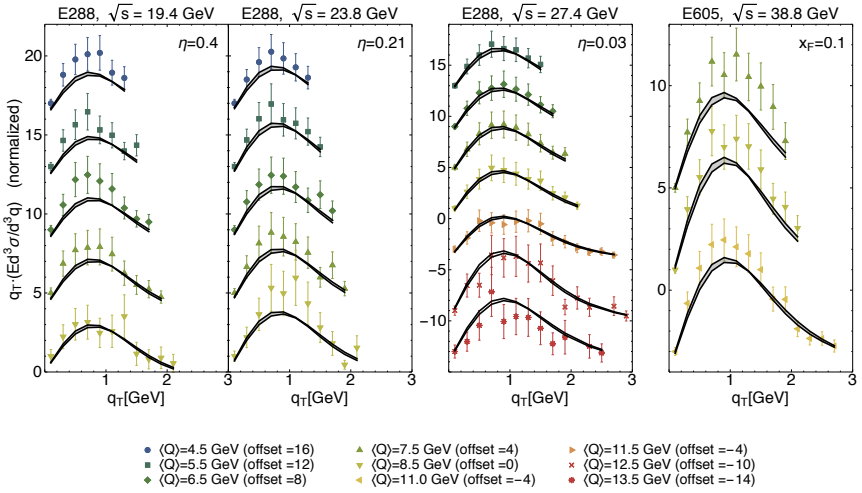


Figure 3.8: Drell–Yan differential cross section for different experiments and values of $\sqrt{s}$ and for different $\langle Q \rangle$ bins. For clarity, each $\langle Q \rangle$ bin has been normalized and then shifted by an offset.

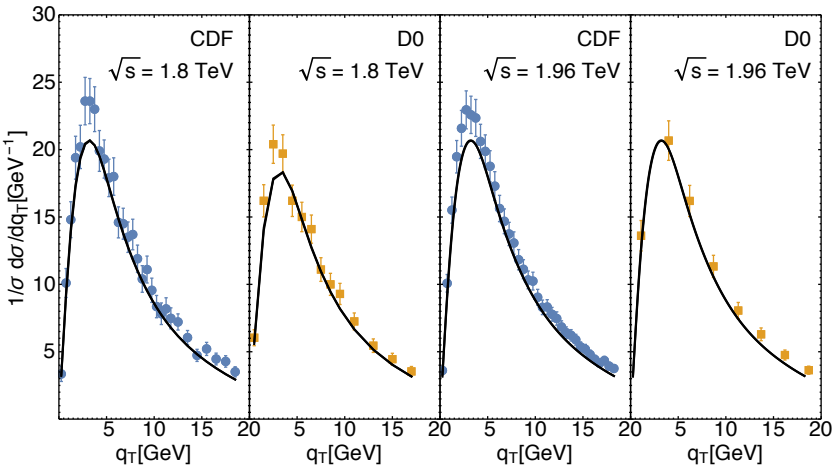
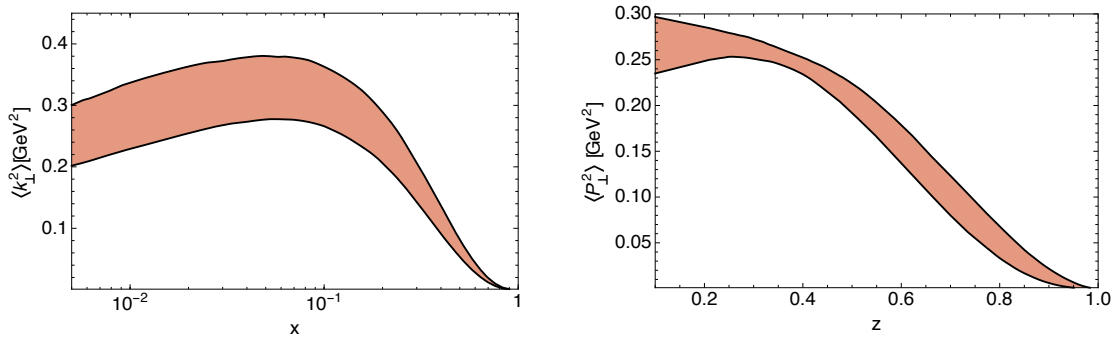


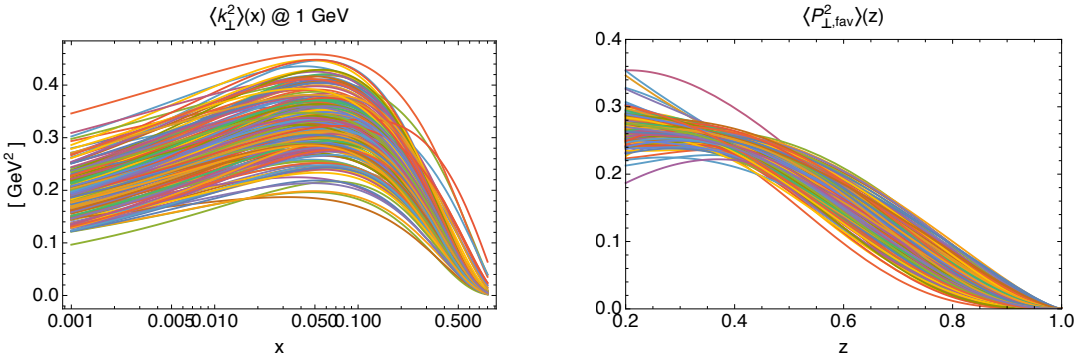
Figure 3.9: Cross section differential with respect to the transverse momentum q_T of a Z boson produced from $p\bar{p}$ collisions at Tevatron. The four panels refer to different experiments (CDF and D0) with two different values for the center-of-mass energy ($\sqrt{s} = 1.8$ TeV and $\sqrt{s} = 1.96$ TeV).

3.7.4 Behavior of transverse momenta



(a) (b)

Figure 3.10: Kinematic dependence of $\langle \boldsymbol{k}_\perp^2 \rangle(x)$ (a) and of $\langle \boldsymbol{P}_\perp^2 \rangle(z)$ (b). The bands are the 68% C.L. envelope of the full sets of best-fit curves.



(a) (b)

Figure 3.11: Kinematic dependence of $\langle \boldsymbol{k}_\perp^2 \rangle(x)$ (a) and of $\langle \boldsymbol{P}_\perp^2 \rangle(z)$ (b). Each curves represent a different set of the 200 replica parameters.

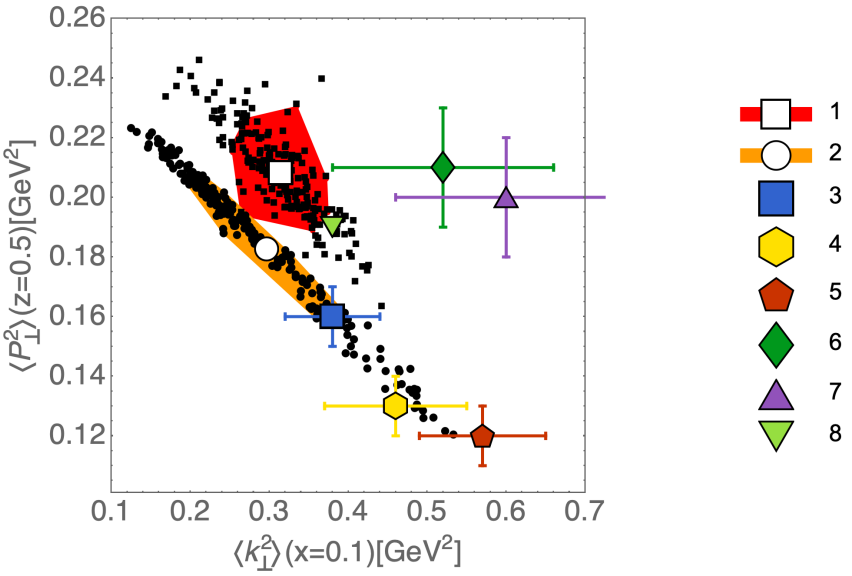


Figure 3.12: Correlation between transverse momenta in TMD FFs, $\langle P_\perp^2\rangle(z = 0.5)$, and in TMD PDFs, $\langle k_\perp^2\rangle(x = 0.1)$, in different phenomenological extractions. (1): average values (white square) obtained in the present analysis, values obtained from each replica (black dots) and 68% C.L. area (red); (2) results from Ref. [72], (3) results from Ref. [144], (4) results from Ref. [119] for HERMES data, (5) results from Ref. [119] for HERMES data at high z, (6) results from Ref. [119] for normalized COMPASS data, (7) results from Ref. [119] for normalized COMPASS data at high z, (8) results from Ref. [25].

3.7.5 Visualization of TMDs

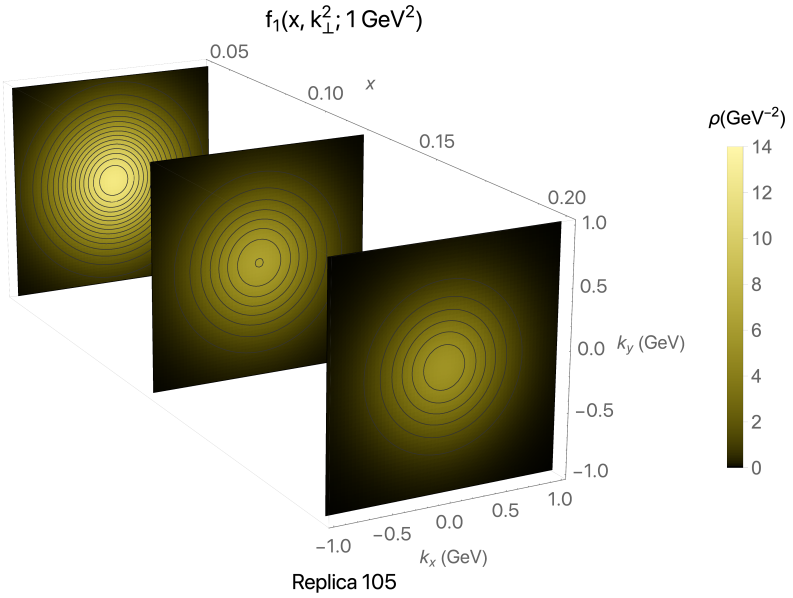


Figure 3.13: Density plot for three-dimensional visualization of $f_1(x, k_\perp^2; Q^2)$ for the up quark extracted from replica 105. The slices are characterized by different values of x.

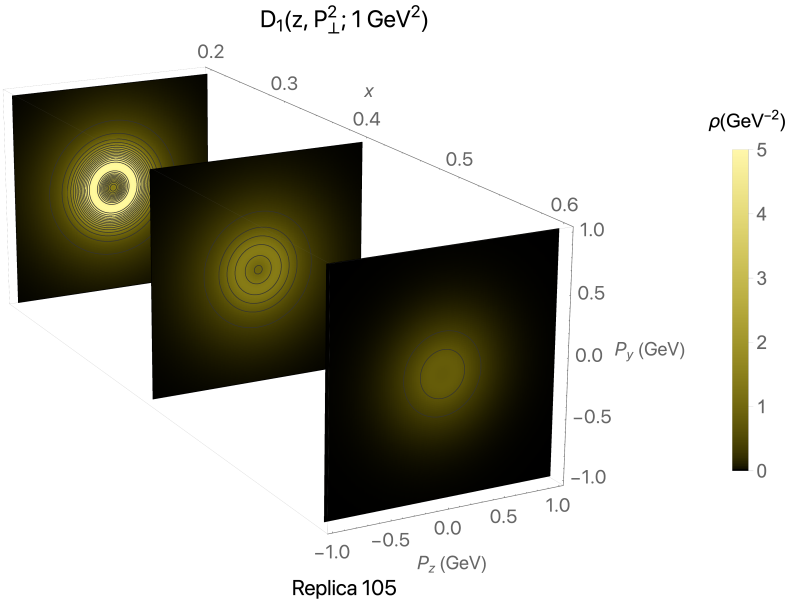


Figure 3.14: Density plot for three-dimensional visualization of $D_1(z, P_\perp^2; Q^2)$ for the up quark extracted from replica 105. The slices are characterized by different values of z.

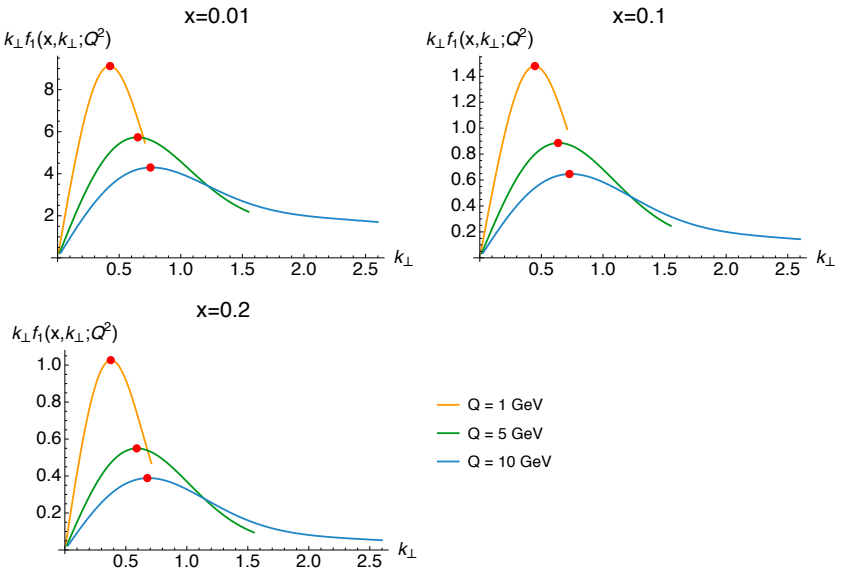


Figure 3.15: TMD parton distribution calculated with the parameters extracted from replica 105 for different values of x and Q. The red point indicates the peak for each curve.

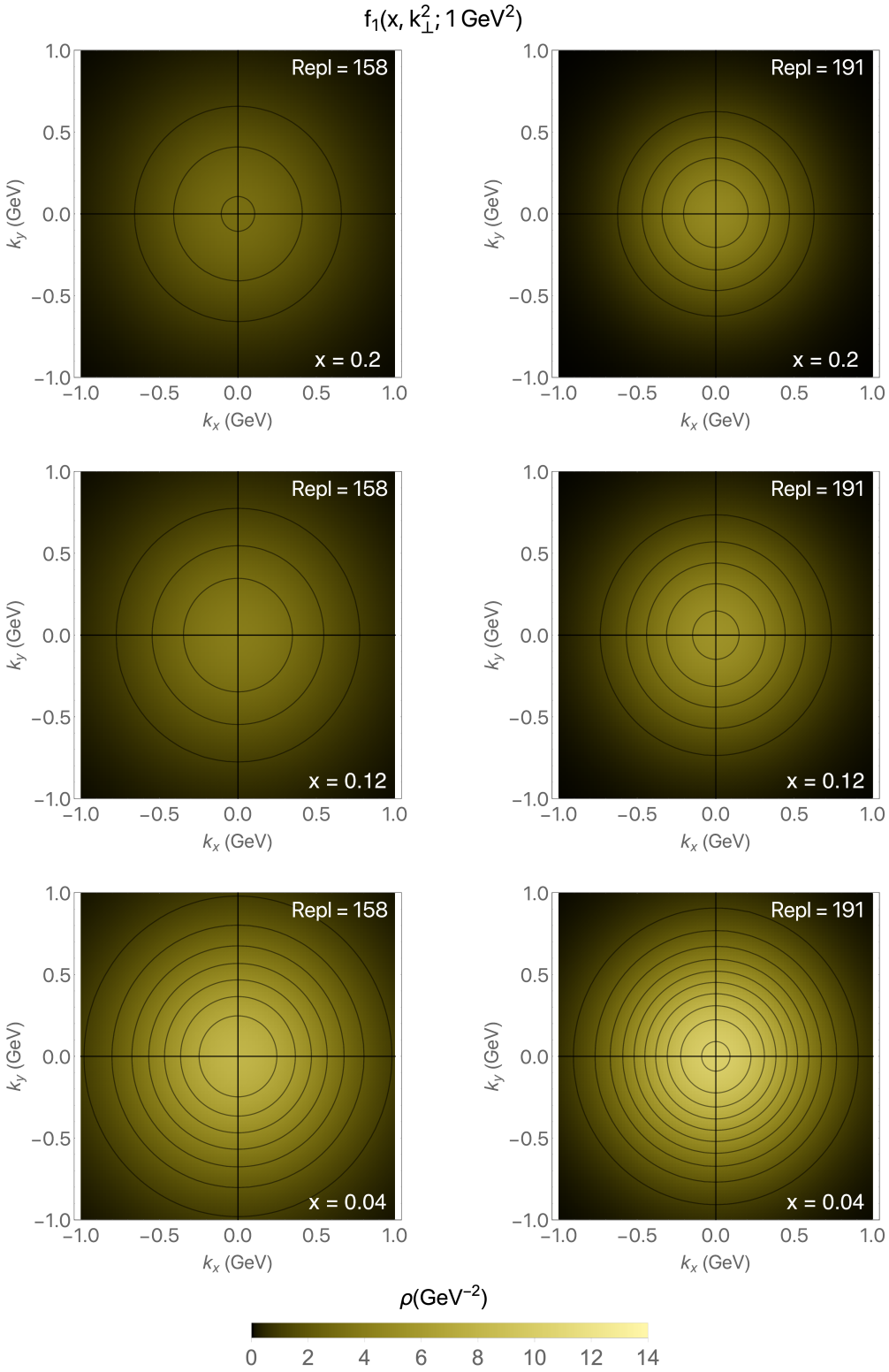


Figure 3.16: Density plots for the TMD parton distribution function in momentum space, calculated at $Q^2 = 1$ GeV^2 for the up quark. Replica parameters used for the calculation are indicated in each panel.

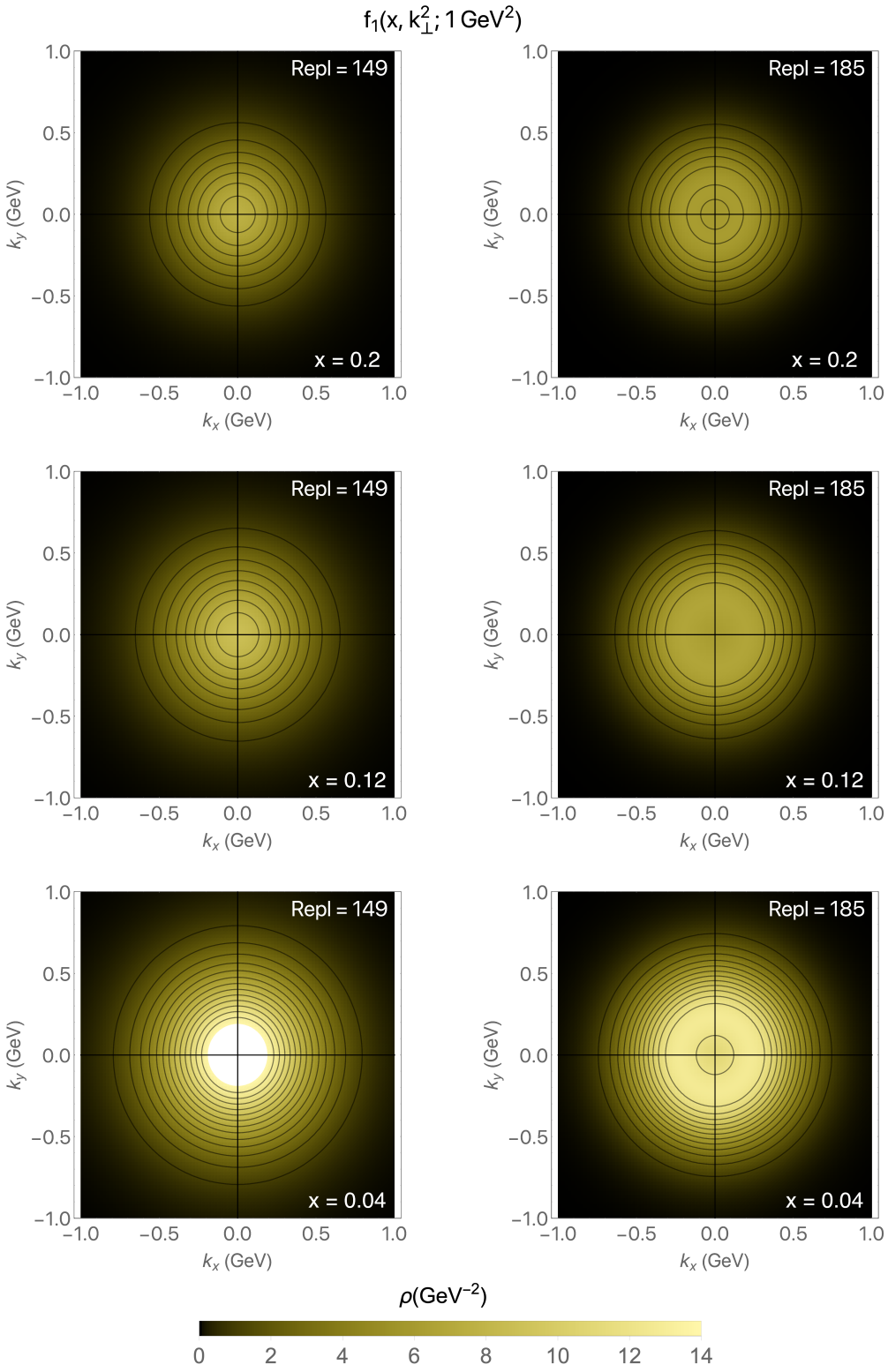


Figure 3.17: Density plots corresponding to Fig. 3.16 for different replica parameters.

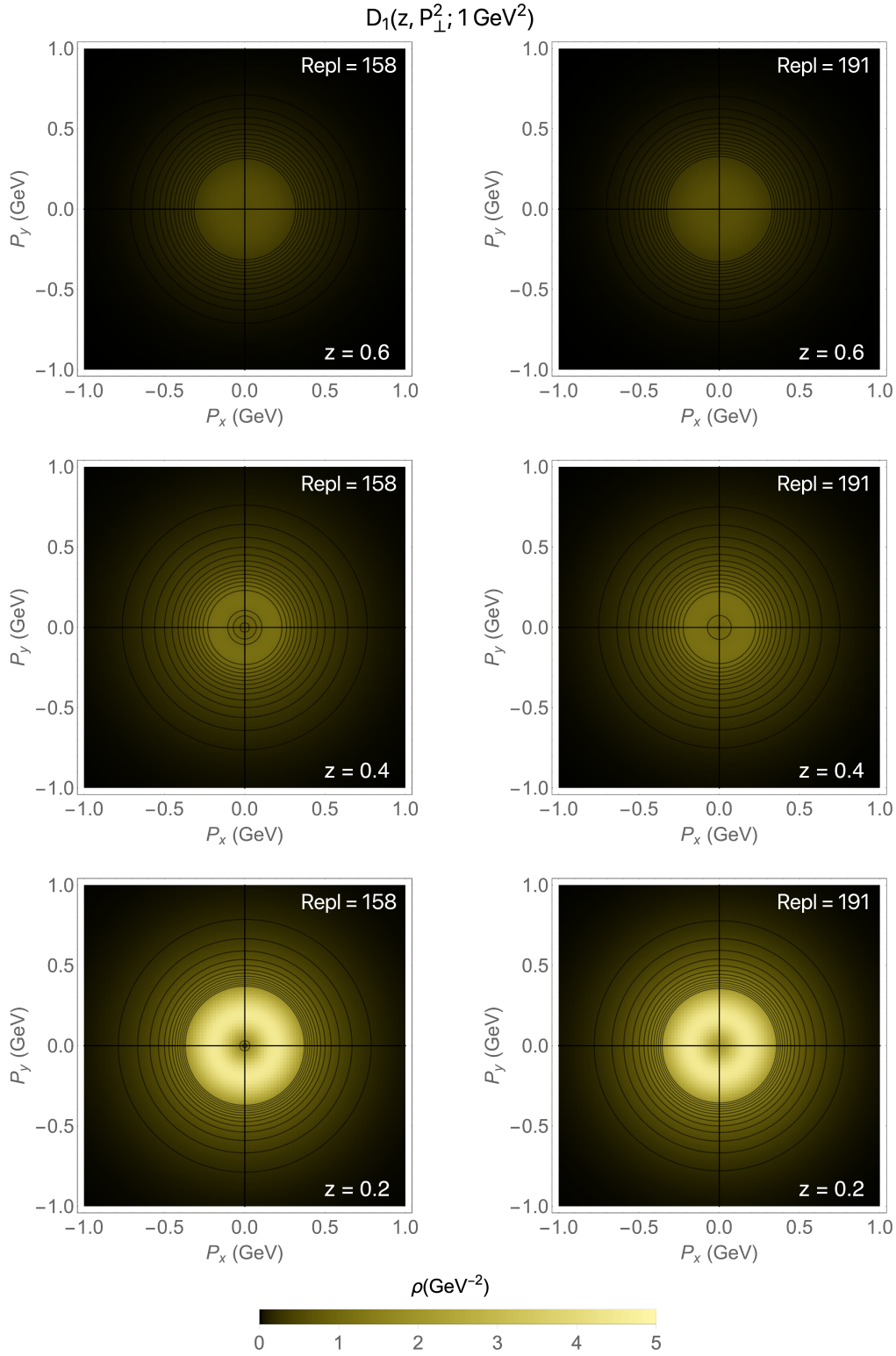


Figure 3.18: Density plots for the TMD fragmentation function in momentum space, calculated at $Q^2 = 1$ GeV^2 for the up quark. The specific replica parameters used for the calculation are indicated in each panel.

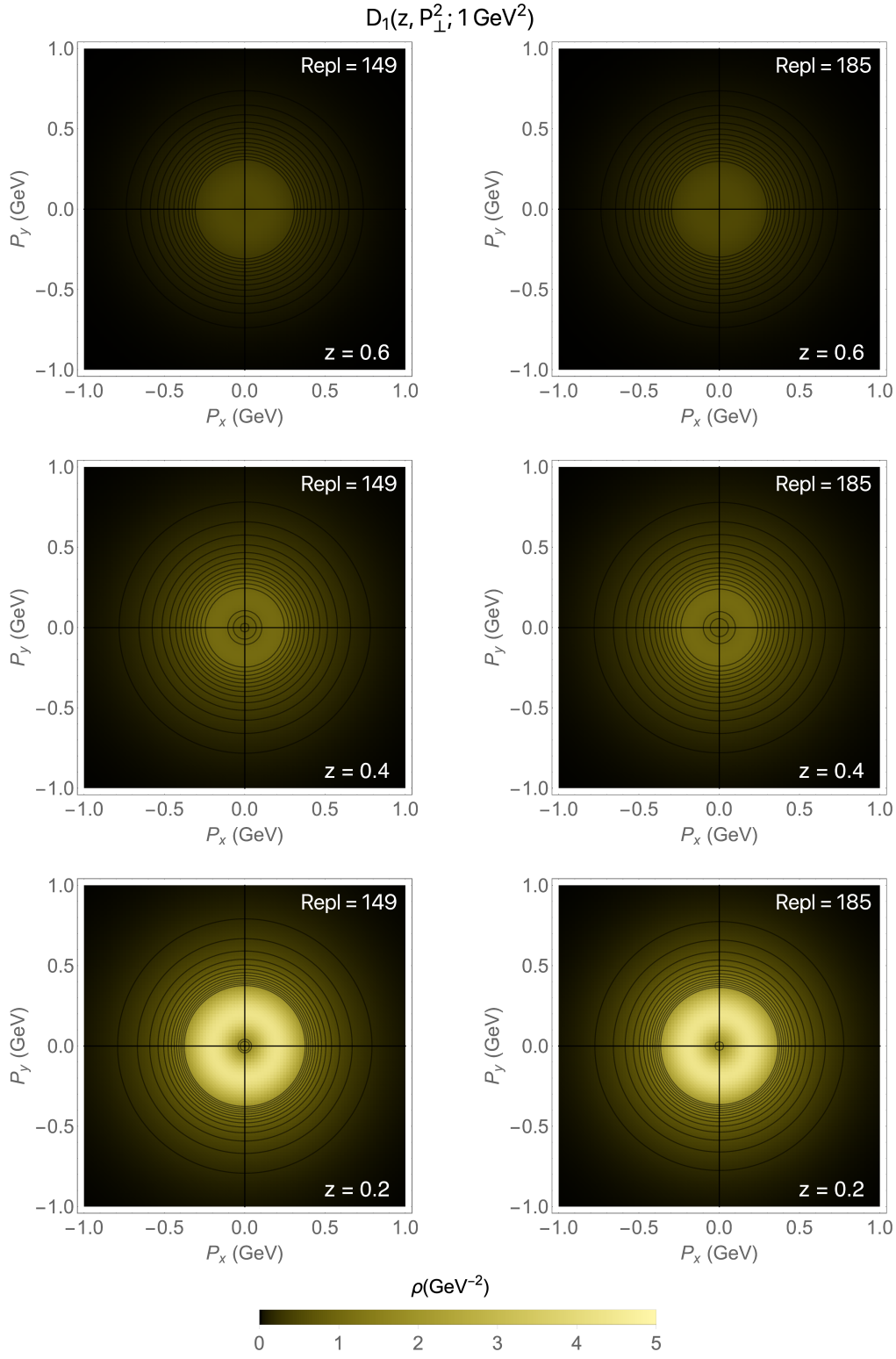


Figure 3.19: Density plots corresponding to Fig. 3.18 for different replica parameters.

Chapter 4

Phenomenology of Sivers Function

4.1 Introduction

The study of the internal structure of nucleons become even more challenging when we consider also the polarization of hadrons and partons. After having completed the extraction of unpolarized TMDs in the previous chapter, we now focus on the analysis of the TMD Sivers distribution function, related to the number density of unpolarized partons inside a transversely polarized nucleon. An accurate determination of this TMD distribution could help clarify the relation between the motion of quarks and the nucleon spin.

As in the previous analysis, we are not able to determine the Sivers function only through analytical calculation, but we have to extract it from experimental data. In particular, we will study the single-spin asymmetry $A_{UT}^{\sin(\phi_h-\phi_S)}$ measured in SIDIS. We will include measurements taken at HERMES at COMPASS in 2009 and 2017, and at JLab. This extraction is connected with our previous determination of unpolarized TMDs, because in the expression used to calculate the azimuthal asymmetry both TMD PDFs and FFs appear. Our analysis, therefore, wants to achieve a determination of the Sivers function, using for the first time unpolarized TMDs directly obtained from experimental data, with the inclusion of the complete TMD evolution formalism.

4.2 Formalism

The Sivers distribution function can be determined through its contributions to the cross section of the semi-inclusive deep inelastic scattering process. To access the Sivers function, we perform a fit on SIDIS measurements where the nucleon is transversely polarized with respect to the virtual photon. The corresponding cross-section can be expressed in a model-independent way by a set of structure functions [46]. In particular, we focus on the measurements

of the Sivers asymmetry, defined through a $\sin(\phi_h - \phi_S)$ modulation of the differential cross section:

$$\frac{d\sigma}{dx\,dy\,dz\,d\phi_S\,d\phi_h d\boldsymbol{P}_{hT}^2} = \frac{\alpha^2}{xyQ^2}\frac{y^2}{2(1-\varepsilon)}\left(1+\frac{\gamma^2}{2x}\right)\Big\{F_{UU,T} + \varepsilon F_{UU,L} + \sin(\phi_h-\phi_S)|\boldsymbol{S}_T|\left[F_{UT,T}^{\sin(\phi_h-\phi_S)} + \varepsilon F_{UT,L}^{\sin(\phi_h-\phi_S)}\right] + \dots\Big\} \quad (4.1)$$

where $\boldsymbol{S}_T$ is the covariant spin vector, P_{hT} is the transverse component of the hadron momentum in the final state and ϕ_S, ϕ_h are their respective azimuthal angles. The dots in the last term indicates higher order contributions. Isolating the terms relevant to the $\sin(\phi_h - \phi_S)$ modulation, the asymmetry becomes

$$A_{UT}^{\sin(\phi_h-\phi_S)} = 2\frac{\int d\phi_S d\phi_h [d\sigma^\uparrow - d\sigma^\downarrow]\sin(\phi_h-\phi_S)}{\int d\phi_S d\phi_h [d\sigma^\uparrow + d\sigma^\downarrow]} \quad (4.2)$$

$$= \frac{F_{UT,T}^{\sin(\phi_h-\phi_S)} + \varepsilon F_{UT,L}}{F_{UU,T} + \varepsilon F_{UU,L}} \quad (4.3)$$

In this analysis we neglect the lepton mass and rely on the one-photon exchange approximation. Moreover, we will consider only the terms at order α_S^0, consequently the structure functions can be written in terms of convolutions of TMDs, in the following way [47]:

$$F_{UU,T} = \mathcal{C}\left[f_1 D_1\right], \quad (4.4)$$

$$F_{UU,L} = \mathcal{O}\left(\frac{M^2}{Q^2}, \frac{P_{hT}^2}{Q^2}\right) = 0, \quad (4.5)$$

$$F_{UT,T}^{\sin(\phi_h-\phi_S)} = \mathcal{C}\left[-\frac{\hat{\boldsymbol{h}}\cdot\boldsymbol{k}_\perp}{M} f_{1T}^\perp D_1\right] \quad (4.6)$$

$$F_{UU,L}^{\sin(\phi_h-\phi_S)} = 0 \quad (4.7)$$

where we introduced the normalized vector $\hat{\boldsymbol{h}} = \boldsymbol{P}_{hT}/|\boldsymbol{P}_{hT}|$ and denoted with $\mathcal{C}$ the transverse-momentum convolution of TMDs. Another consequence of including only LO contributions, is that the effects of the collinear matching term $Y_{UU,T}$ are not included. The Sivers asymmetry measured experimentally can be expressed as this level of accuracy as the ratio between a Sivers function and the corresponding unpolarized TMD distribution, both convoluted with an unpolarized fragmentation function, as in

$$A_{UT}^{\sin(\phi_h-\phi_S)} = \frac{\sigma_0(x,y,Q^2)\mathcal{C}\left[-\frac{\hat{\boldsymbol{h}}\cdot\boldsymbol{k}_\perp}{M} f_{1T}^\perp D_1\right]}{\sigma_0(x,y,Q^2)\mathcal{C}\left[f_1 D_1\right]} \quad (4.8)$$

where we have defined

$$\sigma_0(x,y,Q^2) = \frac{1}{xyQ^2}\frac{1}{1+\gamma^2}\left(1-y+\frac{1}{2}y^2+\frac{1}{4}\gamma^2 y^2\right)\left(1+\frac{\gamma^2}{2x}\right). \quad (4.9)$$

This term is present both in the numerator and denominator, but might not cancel if we take into account the integral of the cross section over at least one of its variable x, y, Q^2.

We can write the structure functions $F_{UU,T}$ and $F_{UU,T}^{\sin(\phi_h-\phi_S)}$ in terms of the Fourier transform of the TMDs in coordinate ξ_T-space, using the following representation of the delta function:

$$\delta^2(z\boldsymbol{k}_\perp + \boldsymbol{P}_\perp - \boldsymbol{P}_{hT}) = \frac{1}{z^2}\int \frac{d^2\xi_T}{(2\pi)^2} e^{i\xi_T\cdot(\boldsymbol{k}_\perp + \boldsymbol{P}_\perp/z - \boldsymbol{P}_{hT}/z)} \tag{4.10}$$

and the Bessel functions J_0 and J_1, defined in general as

$$J_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} d\theta e^{-in\theta + ix\sin\theta}\,. \tag{4.11}$$

We recall the expression for the unpolarized TMD distribution and fragmentation functions, written in ξ_T-space as

$$\tilde{f}_1^a(x,\xi_T^2;Q^2) = \frac{1}{2\pi}\int d^2\boldsymbol{k}_\perp e^{i\xi_T\cdot\boldsymbol{k}_\perp} f_1^a(x,k_\perp^2;Q^2) = \tag{4.12}$$
$$\int_0^\infty d|\boldsymbol{k}_\perp||\boldsymbol{k}_\perp| J_0(\xi_T|\boldsymbol{k}_\perp|) f_1^a(x,k_\perp^2;Q^2)\,,$$

$$\tilde{D}_1^{a\to h}(z,\xi_T^2;Q^2) = \frac{1}{2\pi}\int \frac{d^2\boldsymbol{P}_\perp}{z^2} e^{i\xi_T\cdot\boldsymbol{P}_\perp/z} D_1^{a\to h}(z,P_\perp^2;Q^2) = \tag{4.13}$$
$$\int_0^\infty \frac{d|\boldsymbol{P}_\perp|}{z^2}|\boldsymbol{P}_\perp| J_0(\xi_T\frac{|\boldsymbol{P}_\perp|}{z}) D_1^{a\to h}(z,\boldsymbol{P}_\perp^2;Q^2),$$

$$\tag{4.14}$$

and we introduce the general expression for the derivatives of the Sivers function:

$$\tilde{f}_{1T}^{\perp(n)a}(x,\xi_T^2;Q^2) = n!\left(-\frac{-2}{M^2}\partial_{\xi_T^2}\right)^n \tilde{f}_{1T}^{\perp a}(x,\xi_T^2;Q^2) = \tag{4.15}$$
$$\frac{n!}{(M^2)^n}\int_0^\infty d|\boldsymbol{k}_\perp||\boldsymbol{k}_\perp|\left(\frac{|\boldsymbol{k}_\perp|}{\xi_T}\right)^n J_n(\xi_T|\boldsymbol{k}_\perp|)\tilde{f}_{1T}^{\perp a}(x,\xi_T^2;Q^2)\,.$$

The structure functions can be then defined as convolutions of these TMDs in ξ_T-space as

$$F_{UU,T}(x,z,\boldsymbol{P}_{hT}^2,Q^2) = \tag{4.16}$$
$$2\pi\sum_a e_a^2 x\int_0^\infty d\xi_T\xi_T J_0(\xi_T|\boldsymbol{P}_{hT}|/z)\tilde{f}_1^a(x,\xi_T^2;Q^2)\tilde{D}_1^{a\to h}(z,\xi_T^2;Q^2)\,,$$

$$F_{UT,T}^{\sin(\phi_h-\phi_S)}(x,z,\boldsymbol{P}_{hT}^2,Q^2) = \tag{4.17}$$
$$-2\pi M\sum_a e_a^2 x\int_0^\infty d\xi_T\xi_T^2 J_1(\xi_T|\boldsymbol{P}_{hT}|/z)\tilde{f}_{1T}^{\perp(1)a}(x,\xi_T;Q^2)\tilde{D}_1^{a\to h}(z,\xi_T^2;Q^2)\,.$$

Substituting these Fourier transforms for the structure function we obtain the following expression for the Sivers asymmetry:

$$A_{UT}^{\sin(\phi_h-\phi_S)} = -\frac{M\sigma_0(x,y,Q^2)\sum_a e_a^2 x}{\sigma_0(x,y,Q^2)\sum_a e_a^2 x} \times \frac{\int_0^\infty d\xi_T\, \xi_T^2 J_1(\xi_T|\boldsymbol{P}_{hT}|/z)\tilde{f}_{1T}^{\perp(1)a}(x,\xi_T;Q^2)\tilde{D}_1^{a\to h}(z,\xi_T^2;Q^2)}{\int_0^\infty d\xi_T\, \xi_T J_0(\xi_T|\boldsymbol{P}_{hT}|/z)\tilde{f}_1^a(x,\xi_T^2;Q^2)\tilde{D}_1^{a\to h}(z,\xi_T^2;Q^2)}, \quad (4.18)$$

which will allow to extract the Sivers distribution from experimental data, after an appropriate parametrization.

4.2.1 Parametrization of $f_{1T}^{\perp(1)}$

The azimuthal asymmetry $A_{UT}^{\sin(\phi_h-\phi_S)}$ is now expressed in terms of TMD distributions. However, to complete the framework necessary to the extraction of the Sivers distribution, we need to take another step and choose a reasonable parametrization for $f_{1T}^{\perp}(x,k_\perp^2)$ which describes in an appropriate way its dependence on x and the transverse momentum. The Sivers function can be expressed in term of its first moment, integrated in $k_\perp^2$, and a nonperturbative part:

$$f_{1T}^{\perp}(x,k_\perp^2) = f_{1T}^{\perp(1)}(x) f_{1TNP}^{\perp}(x,k_\perp^2) \quad (4.19)$$

This way to express the Sivers function is particularly useful when considering TMD evolution, because the evolution equations are usually applied to the first moment, as seen in Eqs. (2.152) and (2.153). The last term in Eq. (4.19), even if its definition should be arbitrary in general, is constrained by the positivity bound [145]:

$$\left(f_{1T}^{\perp(1)}(x,k_\perp^2)\right)^2 \leq \frac{\boldsymbol{P}_{hT}^2}{4M^2} f_1^2(x,k_\perp^2) \quad (4.20)$$

We have chosen to implement the unpolarized TMDs obtained from our global fit in the definition of Eq. (4.18), therefore, to respect the positivity bound, we define the nonperturbative $f_{1TNP}^{\perp}(x,k_\perp^2)$ as

$$f_{1TNP}^{\perp}(x,k_\perp^2) = \frac{1}{\pi K_f}\frac{(1+\lambda_S k_\perp^2)}{(M_1^2+\lambda_S M_1^4)} e^{-k_\perp^2/M_1^2} f_{1NP}(x,k_\perp^2) \quad (4.21)$$

where the definition $f_{1NP}(x,k_\perp^2)$ is taken from Eq. (3.21) and K_f is a normalization factor, which will be defined later.

The function $f_{1T}^{\perp}(x,k_\perp^2)$ integrated over the transverse momentum should reduce to $f_{1T}^{\perp}(x)$ or, equivalently, its nonperturbative part, which depends on $k_\perp$, should be normalized to unity. Then, it is necessary to integrate this

quantity in order to completely define $f_{1T}^{\perp}$:

$$\begin{aligned} K_f &\equiv \int d^2k_\perp \frac{k_\perp^2}{2M^2} f_{1TNP}^{\perp}(x,k_\perp^2) \\ &= \pi \int dk_\perp^2 \frac{k_\perp^2}{2M^2} \frac{(1+\lambda_S k_\perp^2)}{(M_1^2+\lambda M_1^4)} e^{-k_\perp^2/M_1^2} \frac{1}{\pi} \frac{(1+\lambda k_\perp^2)}{(g_{1a}+\lambda g_{1a}^2)} e^{-k_\perp^2/g_{1a}} \\ &= \frac{g_{1a} M_1^2}{2\pi M^2 (g_{1a}\lambda+1)\left(g_{1a}+M_1^2\right)^4 (\lambda_S M_1^2+1)} \Big[2g_{1a}\lambda M_1^2 (3g_{1a}\lambda_S M_1^2 \\ &\quad + g_{1a} + M_1^2) + \left(g_{1a}+M_1^2\right)\left(2g_{1a}\lambda_S M_1^2 + g_{1a} + M_1^2\right) \Big] \end{aligned} \tag{4.22}$$

The x-dependence of $f_{1T}^{\perp(1)}(x)$ is chosen to be flexible enough to eventually describe a sign change in the distribution for different parton flavors and longitudinal momentum fractions. Taking inspiration from [146] we define the function

$$g_{Siv}^a(x) = N_{Siv}^a \frac{\mathcal{G}_{siv}^a(x)}{max_x[|\mathcal{G}_{siv}^a(x)|]} \tag{4.23}$$

$$\mathcal{G}_{Siv}^a(x) = x^{\alpha_a}(1-x)^{\beta_a}\left[1 + A_a T_1(x) + B_a T_2(x)\right] \tag{4.24}$$

where $T_n(x)$ are Chebyshev polynomials of order n and N_{Siv}^a is a normalization parameter. We divide $\mathcal{G}_{Siv}^a(x)$ by a fixed parameter, equal to its maximum absolute value, to automatically satisfy the positivity bound at the starting energy scale. Then the complete expression of $f_{1T}^{\perp}$ integrated over $k_\perp^2$ reads

$$f_{1T}^{\perp(1)_a}(x) = \frac{N_{Siv}^a}{G_{max}^a} x^{\alpha_a}(1-x)^{\beta_a}\left[1 + A_a T_1(x) + B_a T_2(x)\right] f_1(x,Q^2) \tag{4.25}$$

Finally, having defined all the necessary terms, we can define the complete parametrization $f_{1T}^{\perp}(x,k_\perp^2)$ used in our analysis:

$$f_{1T}^{\perp a}(x,k_\perp^2) = \frac{1}{K_f} \frac{(1+\lambda_S k_\perp^2)}{\pi(M_1^2+\lambda_S M_1^4)}\, e^{-k_\perp^2/M_1^2} \frac{(1+\lambda k_\perp^2)}{\pi(g_{1a}+\lambda g_{1a}^2)}\, e^{-k_\perp^2/g_{1a}} g_{Siv}^a(x) f_1^a(x)\ . \tag{4.26}$$

To deal with energy scale evolution we want to implement the CSS formalism; it is then more convenient to express the first moment of the Sivers distribution in ξ_T-space. Following Eq. (4.15) we have

$$\begin{aligned} \tilde{f}_{1T}^{\perp(1)}(x,\xi_T^2;Q^2) &= \frac{1}{32\pi\left(g_1+M_1^2\right)^2} \\ &\quad \exp\left(-\frac{\xi_T^2 g_1 M_1^2}{4\left(g_1+M_1^2\right)}\right) \\ &\times \frac{}{(2g_1 M_1^2(3g_1\lambda\lambda_S M_1^2 + g_1\lambda_S + g_1\lambda + (1/2(M_1^2))g_1 + \lambda_S M_1^2 + \lambda M_1^2 + 1 + M_1^2))} \end{aligned} \tag{4.27}$$

$$
\begin{aligned}
&\times \Big(\xi_T^4 g_1^4 \lambda\lambda_S M_1^8 - 24\xi_T^2 g_1^4 \lambda\lambda_S M_1^6 - 4\xi_T^2 g_1^4 \lambda_S M_1^4 - 4\xi_T^2 g_1^4 \lambda M_1^4 \\
&\quad - 24\xi_T^2 g_1^3 \lambda\lambda_S M_1^8 - 8\xi_T^2 g_1^3 \lambda_S M_1^6 - 8\xi_T^2 g_1^3 \lambda M_1^6 \\
&\quad - 4\xi_T^2 g_1^2 \lambda_S M_1^8 - 4\xi_T^2 g_1^2 \lambda M_1^8 + 96 g_1^4 \lambda\lambda_S M_1^4 + 32 g_1^4 \lambda_S M_1^2 + 32 g_1^4 \lambda M_1^2 + 16 g_1^4 \\
&\quad + 192 g_1^3 \lambda\lambda_S M_1^6 + 96 g_1^3 \lambda_S M_1^4 + 96 g_1^3 \lambda M_1^4 + 64 g_1^3 M_1^2 + 96 g_1^2 \lambda\lambda_S M_1^8 + 96 g_1^2 \lambda_S M_1^6 \\
&\quad + 96 g_1^2 \lambda M_1^6 + 96 g_1^2 M_1^4 + 32 g_1 \lambda_S M_1^8 + 32 g_1 \lambda M_1^8 + 64 g_1 M_1^6 + 16 M_1^8 \Big) \\
&\times g_{Siv}(x) f_1(x; Q^2)
\end{aligned}
$$

Without the inclusion of λ_S the previous expression assumes a simpler form:

$$
\begin{aligned}
f_{1T}^{\perp(1)} =& \frac{1}{2\pi} \left(1 - \frac{\lambda g_1^2 M_1^4 \xi_T^2}{4\left(g_1 + M_1^2\right)\left(g_1 + M_1^2 + 2\lambda g_1 M_1^2\right)} \right) \\
&\times e^{-\frac{\xi_T^2 g_1 M_1^2}{4\left(g1 + M_1^2\right)}} g_{siv}(x) f_1(x, Q^2)
\end{aligned} \tag{4.28}
$$

If we consider the distributions of valence and sea quarks as different contributions, we have to redefine our function as

$$
\begin{aligned}
\tilde{f}_{1T}^{\perp(1)q}(x, \xi_T^2; Q^2) &= \tilde{f}_{1T}^{\perp(1)q}(x, \xi_T^2; Q^2) + \tilde{f}_{1T}^{\perp(1)\bar{q}}(x, \xi_T^2; Q^2) \\
&= \widetilde{f}_{1T\mathrm{norm}}^{\perp(1)q}(x, \xi_T^2; Q^2) \\
&\times \left\{ g_{Siv}^{q_v}(x) \Big[f_1^q(x; Q^2) - f_1^{\bar{q}}(x; Q^2) \Big] + g_{Siv}^{\bar{q}}(x) \Big[f_1^{\bar{q}}(x; Q^2) \Big] \right\}
\end{aligned} \tag{4.29}
$$

where we defined $\widetilde{f}_{1T\mathrm{norm}}^{\perp(1)q}(x, \xi_T^2; Q^2)$ as Eq. (4.28) without the x-dependence expressed by $g_{Siv}(x)$. We use the same nonperturbative contribution for all flavors, because the parameter g_1 is taken from our unpolarized TMD fit which, for the moment, does not take into account flavor dependence.

4.2.2 Choices for TMD evolution of the Sivers function

Until this point we have not yet included in our Sivers distribution the modification due to TMD evolution. To be coherent with our choices for the unpolarized fit we follow a similar method based on CSS formalism [147, 25]. The first moment of the Sivers distribution evolved to a certain energy scale Q^2 can be expressed in ξ_T-space as

$$
\begin{aligned}
\widetilde{f}_{1T}^{\perp(1)}(x, \xi_T^2; Q^2) = & \sum_{i=q,\bar{q},g} \left(C_{a/i} \otimes f_1^i \right)(x, \bar{\xi}_*, \mu_b^2) \\
&\times e^{S(\mu_b^2, Q^2)} \left(\frac{Q^2}{\mu_b^2} \right)^{-K(\bar{\xi}_*; \mu_b)} \left(\frac{Q^2}{Q_0^2} \right)^{g_K(\xi_T)} \widetilde{f}_{1T\mathrm{norm}}^{\perp(1)a}(x, \xi_T^2) g_{Siv}(x) \,.
\end{aligned} \tag{4.30}
$$

where we have included the nonperturbative part of Eq. (4.28), that is

$$\widetilde{f}_{1T\,\mathrm{NP}}^{\perp(1)a}(x,\xi_T^2) = \frac{1}{2\pi M^4}\left(1-\frac{\lambda g_1^2 M_1^4 \xi_T^2}{4(g_1+M_1^2)(g_1+M_1^2+2\lambda g_1 M_1^2)}\right) e^{-\frac{\xi_T^2 g_1 M_1^2}{4(g1+M_1^2)}} g_{Siv}(x)\,. \quad (4.31)$$

We recall the definitions, first introduced in the formalism for the evolved unpolarized TMDs, of the scale μ_b and its parameters:

$$\mu_b = \frac{2e^{-\gamma_e}}{\bar{\xi}_*}, \qquad \bar{\xi}_* = \xi_{\max}\left(\frac{1-e^{-\xi_T^4/\xi_{\max}^4}}{1-e^{-\xi_T^4/\xi_{\min}^4}}\right)^{(1/4)} \quad (4.32)$$

where the values of $\xi_{\max}$ and $\xi_{\min}$ are defined in Eq. (3.14).

We choose to describe our asymmetry with NLL accuracy and to consider only the leading term in the expansion of the Wilson coefficients $\mathcal{C}$. As a consequence the expression for the evolved Sivers distribution is reduced to

$$\widetilde{f}_{1T}^{\perp(1)a}(x,\xi_T^2;Q^2) = f_1^a(x;\mu_b^2)\; e^{S(\mu_b^2,Q^2)}\; e^{g_K(\xi_T)\ln(Q^2/Q_0^2)}\; \widetilde{f}_{1T\,\mathrm{NP}}^{\perp(1)a}(x,\xi_T^2)\,. \quad (4.33)$$

The explicit functional form for the Sudakov exponent S and the nonperturbative Sudakov factor g_K have been already defined in Eqs. (2.130) and (3.18). Including in Eq. (4.18) the functional form of Eq. (4.33), together with the evolved unpolarized TMD $\widetilde{f}_1^a(x,\xi_T^2;Q^2)$ and $\widetilde{D}_1^{a\to h}(z,\xi_T^2;Q^2)$ in Eq. (4.18) we obtain the complete expression for the Sivers asymmetry evolved at a certain scale Q and match the experimental measurements taken at different energy values.

4.3 Data Sets

An accurate extraction of the Sivers function with TMD evolution requires the inclusion of asymmetry measurements taken by different experimental collaboration which cover different ranges of kinematic variables. This way we can check the universality of the Sivers formalism and observe the effect produced by different type of targets and final-state hadrons, while at the same time studying the contributions of TMD evolution. In our fit we studied SIDIS Sivers asymmetries on different targets published by the Hermes, Compass and JLab collaborations.

Hermes measured azimuthal single-spin asymmetries during the 2002–2005 run [148], using a transversely polarized gaseous hydrogen target internal to the 27.6 GeV HERA lepton storage ring. Scattered leptons and coincident hadrons were identified using a spectrometer and charged hadrons were additionally separated using a ring-imaging Cherenkov detector (RICH) enabling the possibility to distinguish $\pi^\pm$, π^0, $K^\pm$ contributions in the final measurements. The measured events were preliminarily selected applying the requirements $Q^2 > 1$ GeV^2, $W^2 > 10$ GeV^2, $0.1 < y < 0.95$ and $0.023 < x < 0.4$,

while final-state hadrons were accepted if they satisfied $0.2 < z < 0.7$. The asymmetries for pions and charged kaons are presented in the published data tables as projections of x, z, and $P_{h\perp}$. They are all positive except in the π^- channel, where they are consistent with zero.

The COMPASS collaboration published multiple measurements of the SIDIS asymmetries relative to different runnning periods of the experiment and with different binning criteria. The first ones that we included in our analysis are the data collected from 2002 to 2004 [149], using a longitudinally polarized μ^+ beam scattered off a polarized deuteron target, consisting in two oppositely polarized cells containing solid ^{6}LiD. The measurements are selected using the following cuts: $Q^2 > 1$ GeV^2, $W > 5$ GeV, $0.1 < y < 0.9$. Finally, the data are presented as function of x, z or $P_{h\perp}$ separated in $\pi^\pm$, K^0_S, $K^\pm$ channels. Both pion and kaon asymmetries measured on the deuteron target are rather small and consistent with zero.

The second group of COMPASS SIDIS data that we have considered was taken in 2010 [150] using the same muon beam line with longitudinal polarization and a NH_3 proton target with transverse polarization. These data, already presented in Ref. [151], have been recently presented with a different binning in Q^2. They were reorganized in order to be compatible with Drell–Yan data collected in 2015 using a similar transversely polarized target, in order to allow an easier comparison of Sivers distributions obtained from SIDIS data with those obtained from Drell–Yan data, in future global analyses. Therefore, it is useful to rearrange the SIDIS data using the same Q^2 ranges that appear in DY measurements (subsequent comments relate to the di–muon process):

i) 1 GeV $<$ Q $<$ 2 GeV: low di-muon mass range with many background processes

ii) 2 GeV $<$ Q $<$ 2.5 GeV: intermediate mass range

iii) 2.5 GeV $<$ Q $<$ 4 GeV: mass range of J/ψ

iv) 4 GeV $<$ Q $<$ 9 GeV: high mass range with suppressed background effects.

The lower energy range has a much higher yield.

The requirements for the selection of the events are $Q^2 > 1$ GeV^2, $W^2 > 10$ GeV^2, and $0.1 < y < 0.9$. The particle identification in this case was limited to negatively and positively charged final-state hadrons and, as in the previous case, the data are showed as functions of x, z or $P_{h\perp}$. The mean Sivers asymmetry is positive for all the channels, except for the lowest Q^2 range of the negatively charged hadrons, where it is compatible with zero. For positively charged hadrons, the mean Sivers asymmetry is positive for all four Q^2 ranges, while for negatively charged hadrons it is consistent with zero in the lowest and positive for the other three Q^2-ranges.

Together with proton and deuteron target measurements, we included also single-spin asymmetries measured at Jefferson Lab from semi-inclusive DIS on

	HERMES	COMPASS'09	COMPASS'17	JLab
Reference	[148]	[149]	[150]	[152]
Data points	95	88	111	6
Target	p [H]	D [^{6}LiD]	p [NH_3]	n [^{3}He]
Q^2 range [GeV^2]	1.28 - 6.63	1.24 - 31.4	1.27 - 37.7	1.38 - 2.68
x range	0.035 - 0.275	0.006 - 0.286	0.006 - 0.638	0.15 - 0.35
y range	0.043 - 0.717	0.19 - 0.66	0.11 - 0.77	0.70 - 0.81
z range	0.233 - 0.663	0.233 - 0.879	0.14 - 0.74	0.50 - 0.58
P_{hT} range [GeV]	0.11 - 0.972	0.154 - 1.57	0.21 - 1.33	0.24 - 0.44

Table 4.1: Target type and kinematic ranges of Q^2, x, y , z and P_{hT} for SSA data. The data point numbers refer to the ones that satisfy the kinematic cuts of the fit.

a transversely polarized ^{3}He neutron target in the 2008–2009 run [152]. SIDIS events selections required a four- momentum transfer squared $Q^2 > 1$ GeV^2, an hadronic final-state invariant mass $W > 2.3$ GeV, and a mass of undetected final-state particles $W' > 1.6$ GeV. The data are separated into positive and negative final-state π and presented as a function of x only, for a total of 8 data points. JLab data suggest negative π^+ Sivers moments, while the π^- moments are close to zero.

For the selection of data included in our study we applied the same criteria used for the unpolarized TMD fit, i.e., $Q^2 > 1.4$ GeV^2, $0.20 < z < 0.74$ and $P_{hT} < \min[0.2Q, 0.7Qz] + 0.5$ GeV. After applying these kinematic cuts we have a total of 300 data points, that is 95 from HERMES , 88 from COMPASS'09 , 111 from COMPASS'17 , and 6 from JLab. The characteristics of the included data points are summarized in Tab. 4.1 .

4.4 Results

The purpose of our fit is the extraction of Sivers distribution from the complete set of currently available SIDIS asymmetries measurements, using a consistent TMD formalism for the treatment of energy scale evolution.

Our previous determination of the unpolarized TMDs gives us an interesting opportunity to include in the model for $A_{UT}^{\sin(\phi_h-\phi_S)}$ a parametrization for the unpolarized distribution that is not simply based on reasonable assumptions, but instead has been extracted directly from experimental data [153]. Having a good knowledge of the unpolarized TMD distributions is essential to consistently study the Sivers function, considering their interplay in the convolutions that appear in the definition of the azimuthal asymmetry $A_{UT}^{\sin(\phi_h-\phi_S)}$. More-

over, using our unpolarized TMD parametrization could be useful to check if it is actually able to describe the azimuthal asymmetries or if it has any pathological behavior that was not clearly evident in the description of the unpolarized data.

The procedure used for the best fit of the Sivers asymmetries is analogous to the replica method described in Section 3.4. We created $\mathcal{M}$ different replicas of the original data, shifting them by a Gaussian noise and maintaining the same original variance for every single data point. When calculating the theoretical predictions for $A_{UT}^{\sin(\phi_h-\phi_S)}$, we need to choose a suitable parametrization for the unpolarized TMDs. For every replica r of the asymmetries we have chosen the set of parameters extracted from the corresponding replica r of unpolarized data. Considering this, and following the procedure applied for the unpolarized extraction, we create a total of 200 replicas of the original A_{UT} data. For the minimization process we adopted, where possible, the same choices used in our previous work, in order to be consistent with the framework where our description of the TMD PDFs and FFs is valid. Therefore, for the calculation of collinear PDFs we kept the GJR08FFnloE set [106] and for the collinear FFs we again used the DSS14 NLO set for pions [107] and the DSS07 NLO set for kaons [108]. Each minimization started from a different set of initial parameters chosen in a reasonable interval, to explore the parameters space without being too much constrained by their initial choice, while at the same time avoiding area of no physical significance.

As said in the previous section, the Compass and Hermes measurements of the asymmetry are presented as function of x, z and P_{hT}. However, these three groups of data refer to the same measurements, only projected on different observables. Therefore we decided to fit only one of these projections in order to avoid considering fully correlated measurements. We chose to analyze the x sets of data, given that we are mainly interested in the x-dependence of the Sivers function.

Adopting this configuration for our minimization, we considered 118 data points of the data sets projected on x which, after being reduced by 14 free parameters, gives a total number of degrees of freedom equal to 104. We obtained a good agreement between the experimental measurements of A_{UT} as a function of x and our theoretical prediction, with an overall value of $\chi^2/\text{d.o.f.} = 1.06 \pm 0.12$. The uncertainties are estimated from the 68% confidence level obtained through the replica methodology. The global results of our minimization are summarized in Table 4.2 and the histogram of the distribution of χ^2/d.o.f. values obtained from the replica methodology is shown in Fig. 4.1.

In Table 4.3 we report the value of total χ^2 and the number of data points analyzed, distinguished according to their experimental collaboration. Instead, Table 4.4 present the same quantities, separated with respect to the detected hadron in the final state. We observe that our parametrization is able to describe very well the x-group data of Compass'17, even if they have smaller

Points	Parameters	χ^2	χ^2/d.o.f.
118	14	110.19 ± 10.84	1.06 ± 0.10

Table 4.2: Number of included data points, of free parameters and values of χ^2. The difference between the number of data points and free parameters gives the total degrees of freedom.

	HERMES	COMPASS'09	COMPASS'17	JLab
χ^2	47.60 ± 7.29	30.10 ± 4.75	31.10 ± 5.98	5.01 ± 1.54
Points	30	32	50	6

Table 4.3: Values of obtained total χ^2 and uncertainties, and corresponding number of data points, separated according to their experimental collaborations.

	π^+	π^0	π^-	K^+	K^-
28.80 ± 5.05	5.68 ± 0.30	44.40 ± 2.80	13.30 ± 2.82	20.70 ± 3.21	
Points	42	6	42	14	14

Table 4.4: Values of obtained total χ^2 and uncertainties, and corresponding number of data points, separated according to the detected hadron in the final state.

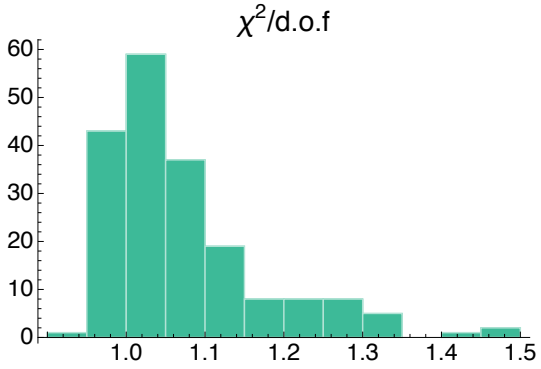


Figure 4.1: Global distribution of χ^2/d.o.f. values obtained from the minimization of 200 replicas.

uncertainties compared with the other data sets. This could be probably due to their Q^2 binning, which is much finer compared to the other data sets considered. Another possible reason is that the COMPASS'17 data do not identify the type of hadron in the final state, only their charge; they could be more compatible with our extraction of unpolarized TMDs, which did not take into account flavor dependence and was based mostly on COMPASS data without final-state particle identification. Also in the cases of COMPASS'09 and JLab measurements we obtain a good agreement with our theoretical prediction. However, in these cases we have larger uncertainties compared to the ones of COMPASS'17.

The HERMES asymmetries data set is, in general, well described by our extracted functional form. Their agreement however is slightly worse compared with the other experiments and its χ^2 value is larger than the one obtained for the global fit. Comparing the distributions of the HERMES data points with the rest of the data sets, we observe that it is less regular and it has smaller uncertainties, resulting then more challenging to describe accurately.

Our extraction of the Sivers distribution function presents larger uncertainties compared with similar results obtained by other collaborations. There are different factors that contribute to produce larger error bands. First of all, we adopted a functional form more flexible than in the past, to reduce the bias connected to its choice. This aspect is emphasized by the implementation of the replica methodology to determine error bands, as it allows the already flexible function to fit different realistic outcomes of the experiments.
Another aspect to take into account is that we are using unpolarized TMDs which in turn have been extracted from experimental measurements. Therefore, we are introducing in the final outcome also these additional uncertainties. We believe it is important to include them, as the determinations of unpolarized TMDs are still affected by relatively large uncertainties.
Last but not least, in our fit we consider only the x-projection of the SIDIS $A_{UT}^{\sin(\phi_h-\phi_S)}$ data. Past studies included also the z and P_{hT} projections, which

are however fully correlated. This is equivalent to artificially increase the number of data, duplicating the actual experimental measurements, and probably producing smaller error bands as a consequence. Therefore, we believe that our recent error estimates more accurately reflect the actual knowledge of the Sivers function.

The inclusion of TMD evolution contributions in the study of the Sivers asymmetry was not dictated only by the necessity to give a good description of the experimental data. In fact, repeating the analysis using the same framework but excluding the terms related to TMD evolution, we are able to obtain a similar outcome, with $\chi^2 = 1.07 \pm 0.17$. The main reason to include these contributions is that we are using unpolarized TMDs extracted from experimental data and we want to be coherent with the formulas used to obtain them. In that case TMD evolution has been proved to play an important role, as the data cover a wide range in Q^2. Using this parametrization to describe the asymmetries, without including all the necessary ingredients, would introduce some inconsistencies in our analysis. We observe that, even if the terms related to TMD evolution do not produce a significant effect due to the limited range energy range, in principle their exclusion is not justified by theoretical considerations, but only as a way to simplify the formulas used. In any case, we want to test the validity of the complete TMD formalism, including evolution, and its ability to reproduce the quantities experimentally observed. Moreover, the study of the Sivers distribution can be extended to A_N asymmetries measured on DY processes. These measurements cover a wider energy interval and TMD evolution could play an important role, thus, it makes sense to already include these contributions, in view of future analyses which may include DY data.

The agreement of theoretical predictions and experimental data can be observed more in detail in the figures contained in Section 4.6. In Fig. 4.5 are shown the asymmetries $A_{UT}^{\sin(\phi_h - \phi_S)}$ measured at Hermes, in Fig. 4.6 the COMPASS'09 data, in Fig. 4.7 and Fig. 4.8 the COMPASS'17 data for positive and negative final-state hadrons respectively and, finally, in Fig. 4.9 we have the results for JLab measurements. All plots presents the Sivers asymmetry as a function of x,z and ¶$_{hT}$, except for the Jlab data where only the x bin is present. The colored data points represent the experimental data with the corresponding uncertainties, while the grey bands denote the 68% confidence level of the 200 replicas.

In Table 4.5 we present the best fit parameters extracted from our fit. They are flavor dependent and have different values for the contributions of valence u and d quarks and for the sea quarks s. In particular, we noticed that the β parameters are not well determined by the minimization, and we decided to assign them a fixed value in a reasonable interval, determined through preliminary fits with β as the only free parameter. In the second line of Table 4.5 we show the best fit parameters obtained using the replica 105 parametrization, used as a reference to explore more in detail the results of the unpolarized fit.

	M_1	λ_S	α_d	α_u	α_s
All Repl.	0.55 ± 0.55	-9.23 ± 10.10	0.33 ± 0.70	-0.01 ± 0.30	0.52 ± 0.66
Repl. 105	1.30×10^{-6}	-1.97	1.24	0.02	0.22
β_d (r)	β_u (r)	β_s (r)	A_d	A_u	A_s
0.50 ± 0.33	0.48 ± 0.34	0.49 ± 0.33	10.80 ± 19.30	2.47 ± 8.97	-0.01 ± 6.55
0.85	0.99	0.20	12.40	-0.55	-0.20
B_d	B_u	B_s	Nx_d	Nx_u	Nx_s
1.31 ± 3.62	-2.25 ± 6.12	0.15 ± 2.96	0.11 ± 0.56	-0.02 ± 0.12	-0.05 ± 0.27
6.49	1.10	1.01	-1.00	0.05	-1.00

Table 4.5: Values of best fit parameters for Sivers distribution. The values in the second row refers to the best fit parameters obtained from replica 105. The symbol (r) denotes that the parameter is not free, but randomly chosen.

The plots in Fig. 4.2 shows the values of $x f_{1T}^{\perp(1)}(x, Q^2)$ for the valence and sea quarks are shown as a function of x. The first moment of the Sivers distribution are evaluated at $Q_0 = 1$ GeV and in the case of the sea quarks, we show only the strange quark case, as an example. The bands represents the 68% confidence level obtained through the replica methodology, calculating $x f_{1T}^{\perp(1)}(x, Q^2)$ for each group of replica parameters, and excluding for each value of x the results outside the 16^{th} and 84^{th} percentile.

The distribution for the up quark is negative, while for the down quark is positive and they have a similar magnitude. The contribution of the s quark is instead close to zero. The values that we found for the first moment of the Sivers distributions are compatible with previous studies of the Sivers function, that will be discussed in the next section.

4.5 Comparison with previous extractions of the Sivers function.

4.5.1 2011 Pavia group

The Pavia group carried out a previous study on the Sivers function [23], considering also its relation with the quark angular momentum. The group of SIDIS asymmetries data included in this study differs from the current analysis for the important addition of COMPASS proton data. Moreover, it did not consider TMD evolution effects. For the parametrization of unpolarized TMDs, it used a Gaussian model:

$$f_1^a\left(x, k_\perp^2; Q_0^2\right) = \frac{f_1^a\left(x; Q_0^2\right)}{\pi \left\langle k_\perp^2 \right\rangle} e^{-k_\perp^2/\langle k_\perp^2 \rangle}, \quad D_1^a\left(z, P_\perp^2; Q_0^2\right) = \frac{D_1^a\left(z; Q_0^2\right)}{\pi \left\langle P_\perp^2 \right\rangle} e^{-P_\perp^2/\langle P_\perp^2 \rangle} \tag{4.34}$$

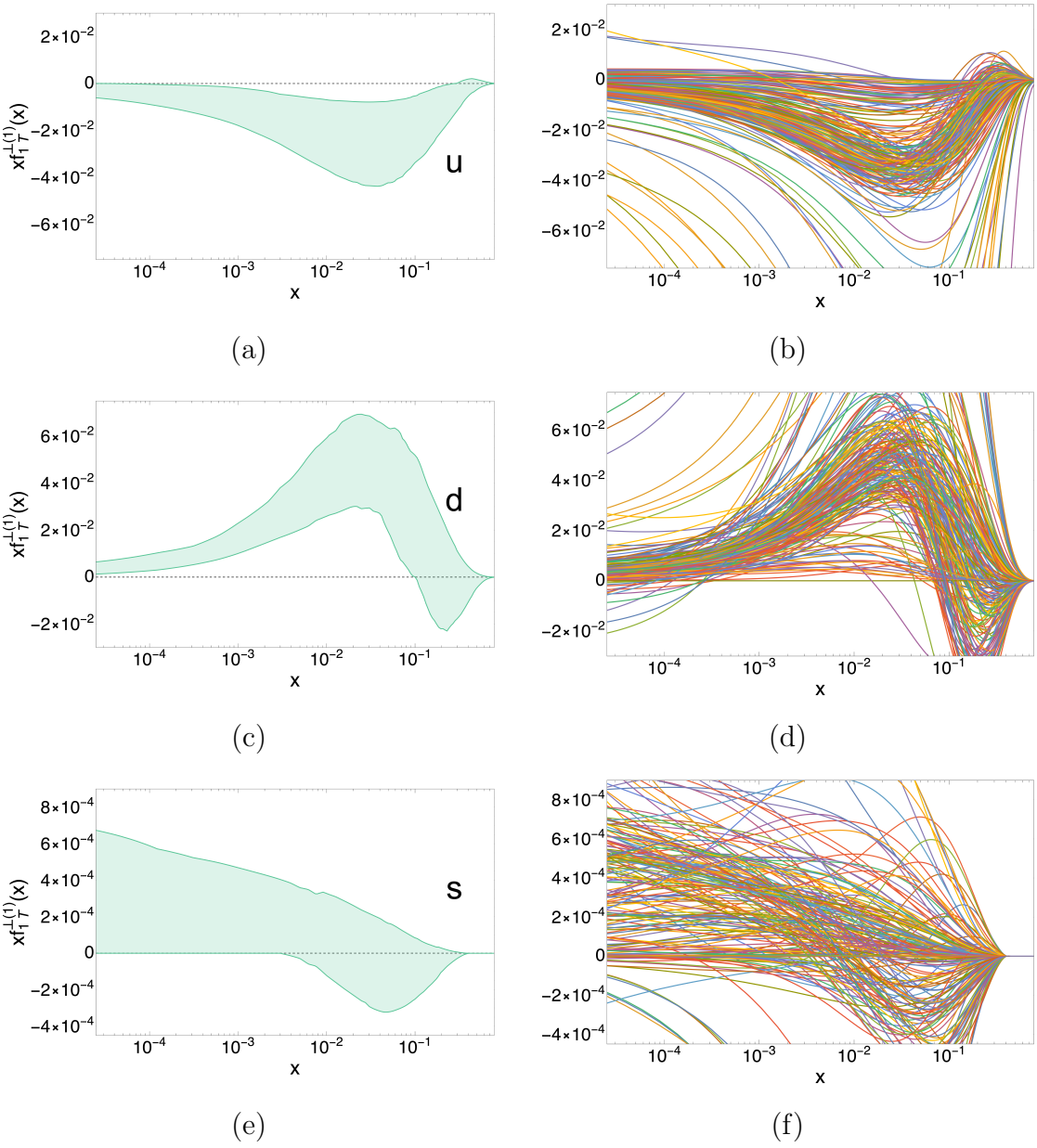


Figure 4.2: The function $x f_{1T}^{\perp(1)}$ as a function of x calculated for the up (a,b), down (c,d) and strange (e,f) quark at the scale $Q_0 = 1$ GeV. The uncertainty bands shown in (a),(c) and (e) are created using the 68% C.L. obtained with the replica methodology. In the panels (b),(d) and (f) we show all the curves corresponding to the 200 replica fits.

where the definitions of the variables is presented more in detail in the referenced paper.

The Sivers function, instead, was parametrized as function of its zeroth moment in the following way:

$$f_{1T}^{\perp a}\left(x,k_{\perp}^{2};Q_{0}^{2}\right)=f_{1T}^{\perp(0)a}\left(x;Q_{0}^{2}\right)\frac{M_{1}^{2}+\langle k_{\perp}^{2}\rangle}{\pi M_{1}^{2}\langle k_{\perp}^{2}\rangle}e^{-k_{\perp}^{2}/M_{1}^{2}}e^{-k_{\perp}^{2}/\langle k_{\perp}^{2}\rangle} \tag{4.35}$$

and, for the valence quark:

$$f_{1T}^{\perp(0)q_{v}}\left(x;Q_{0}^{2}\right)=C^{q_{v}}\sqrt{2e}\frac{MM_{1}}{M_{1}^{2}+\langle k_{\perp}^{2}\rangle}\frac{1-x/\alpha^{q_{v}}}{|\alpha^{q_{v}}-1|}(1-x)f_{1}^{q_{v}}(x;Q_{0}^{2}) \tag{4.36}$$

where M_1, α^{q_v},C^{q_v} were free parameters and the model distinguished up, down and sea quarks contributions.

Moreover, to explore the relation between the quark angular momentum and the azimuthal asymmetries, the function $f_{1T}^{(0)a}(x,Q^2)$ is connected to the generalized parton distribution E^a, at $Q = 1$ GeV, through a certain lensing function in the following way:

$$f_{1T}^{\perp(0)a}\left(x;Q_{L}^{2}\right)=-L(x)E^{a}\left(x,0,0;Q_{L}^{2}\right)\,, \tag{4.37}$$

taking inspiration from spectator model results and theoretical consideration. The total longitudinal angular momentum J^a can be computed from the GPDs H_a and E_a, creating a relation with the Sivers function and thus imposing an additional constraint to its extraction, while at the same time opening a possible way to determine quark angular momentum. This model gave a good agreement between the experimental values and the theoretical predictions with a $\chi^2/d.o.f$ around 1.3 and gives a good estimate of the quark angular momentum.

As a note, we briefly show how to pass from the functional forms of this original paper to our the current framework. The first important step is to Fourier transform the Gaussian models for TMDs from the momentum space to the configuration one, as discussed in Section 4.2.

We can express $f_{1T}^{\perp}(x,Q_0^2)$ as a function of its first moment, in a similar way to our recent fit:

$$f_{1T}^{\perp}(x,Q_{0}^{2})=\frac{2M^{2}\left(\langle k_{\perp}^{2}\rangle+M_{1}^{2}\right)^{2}}{\pi\left(\langle k_{\perp}^{2}\rangle M_{1}^{2}\right)^{2}}f_{1T}^{\perp(1)}(x,Q_{0}^{2})e^{-k_{\perp}^{2}/M_{1}^{2}}e^{-k_{\perp}^{2}/\langle k_{\perp}^{2}\rangle} \tag{4.38}$$

The first moment $\tilde{f}_{1T}^{\perp(1)a}$, if we include the contributions of TMD evolution, can be written as

$$\tilde{f}_{1T}^{\perp(1)a}(x,\xi_{T}^{2};Q_{0}^{2})=\frac{\langle k_{\perp}^{2}\rangle M_{1}^{2}}{4\pi M^{2}(\langle k_{\perp}^{2}\rangle+M_{1}^{2})}e^{-\frac{\xi_{T}^{2}\langle k_{\perp}^{2}\rangle M_{1}^{2}}{4(\langle k_{\perp}^{2}\rangle+M_{1}^{2})}}f_{1T}^{\perp(0)a}(x;\mu_{b}^{2})\;e^{g_{K}(\xi_{T})\ln(Q^{2}/Q_{0}^{2})}\,. \tag{4.39}$$

Inserting this distribution in $A_{UT}^{\sin(\phi_h-\phi_S)}$ we can reproduce the results of the original paper.

4.5.2 EIKV collaboration

In 2014 the Echevarria-Ibildi-Kang-Vitev collaboration [25] performed a fit of the Sivers asymmetry in SIDIS, focusing in particular on the different terms that contribute to the TMD evolution, such as the nonperturbative Sudakov factor. The minimization considers the data sets included in the Pavia 2011 fit with the addition of COMPASS data with proton target. The proton COMPASS data, however, are from a different data set than the one included in our current fit, and refers to a previous publication without a fine binning in Q^2.

To describe the TMD PDF and FF, a simple Gaussian functional form was chosen, with their widths related to the transverse momenta. To take into account the flavor dependence, the contributions for the quarks u, d, $\bar{u}$, $\bar{d}$, and for the sea quark s were distinguished.

The authors used the model obtained from this global fit to make predictions for the Sivers asymmetry in DY lepton pair and W/Z production at the characteristic energies of experimental measurements at Tevatron and LHC, to test the sign change of the Sivers functions between SIDIS and DY processes and constrain the sea quark Sivers functions.

4.5.3 Torino-Cagliari collaboration.

In 2018 a collaboration between the research groups of Torino and Cagliari published an extraction of the Sivers function from azimuthal asymmetries measured in polarized SIDIS [153]. At the moment this is the most updated published analysis of the Sivers asymmetries and can be considered the continuation of their previous studies of the Sivers distribution from SIDIS [154] and W/Z production data [26].

This analysis and, later, our work included the same data sets taken from HERMES COMPASS and JLab, given that they are the most recent data available for polarized semi-inclusive DIS. They excluded the measurements for the K^- channel because the sea contributions, which drives the kaons data, was not separated from the valence contribution. Another important difference is that the Torino-Cagliari group included together in the minimization all the projections of the data on x, z, P_{hT}.

To address the Q^2 dependence of the Sivers function the authors considered three different approaches, one without any evolution, one adopting a collinear twist-3 evolution, and the last one using a TMD-like approach [24]. For each one of these cases they found a good agreement with the data, with a χ^2/d.o.f of 0.99, 0.94 and 0.99 respectively.

A crucial feature of the framework adopted for our fit, which emerges from this comparison and constitutes an interesting improvement, is the inclusion of unpolarized TMDs extracted directly from unpolarized data. The Torino-Cagliari fit, instead, adopted for the unpolarized TMDs an arbitrary reasonable parametrization, not related to an extraction form experimental data. Moreover, in our minimization for the scale dependence we fully included the

standard formalism for the TMD evolution. With these choices we found results comparable with the one published by the Torino-Cagliari group. The comparison between the two studies will be discussed in the next section.

4.5.4 Graphical comparison of different extraction.

We present the comparison between our extraction of the first moment of the Sivers function with the results obtained by other research groups, in Fig. 4.3 for the up quarks and in Fig. 4.4 for the down quarks. We denote with PV11 the result of the Pavia group in 2011 [23], with TC18 the *reference fit* of Torino-Cagliari [153] results and with PV18 our most recent results. The data were presented calculated for different values of Q^2, the curves of TC18 were calculated at $Q^2 = 1.2$ GeV^2, the results of Pavia at $Q^2 = 1$ GeV^2, while the curve corresponding to EIKV were presented in the original publication at $Q^2 = 2.4$ GeV^2. In our plot, however we calculated again the EIKV curve at $Q^2 = 1$ GeV^2 adding also the estimate for the error bands, which were not present originally. In some cases we had to rearrange the final results, so that they referred to the same quantity $x f_{1T}^{\perp(1)}$; for example in the case of TC18 the final results were presented in terms of $\Delta^N f^{(1)}_{q/p^\uparrow}(x) = -f_{1T}^{q\perp(1)}(x)$, while EIKV discussed the Qiu-Sterman function $T_{q,F}(x,x,Q)$.

From the plots, we can conclude that our current analysis is in good agreement with the previous extraction of the Sivers function from similar data sets of polarized SIDIS.

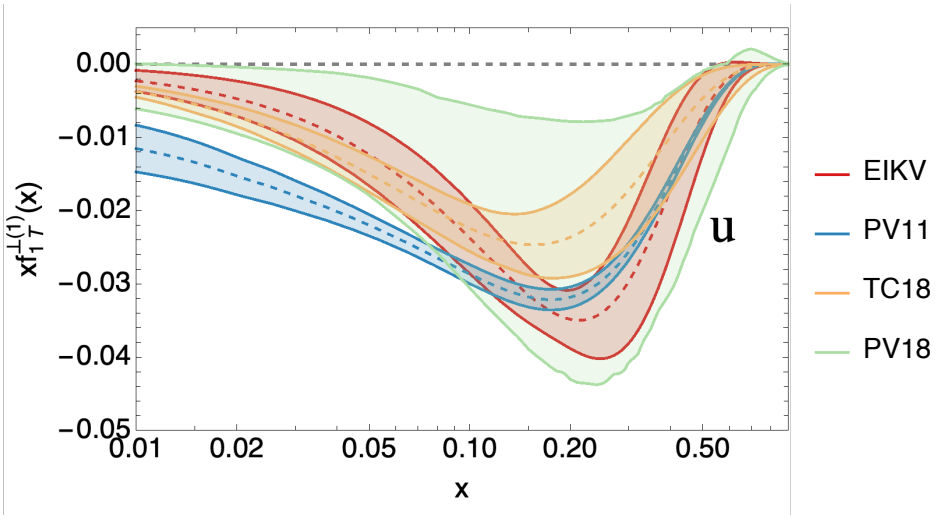


Figure 4.3: Comparison of different extractions of the first moment of Sivers distribution as a function of x for up quark.

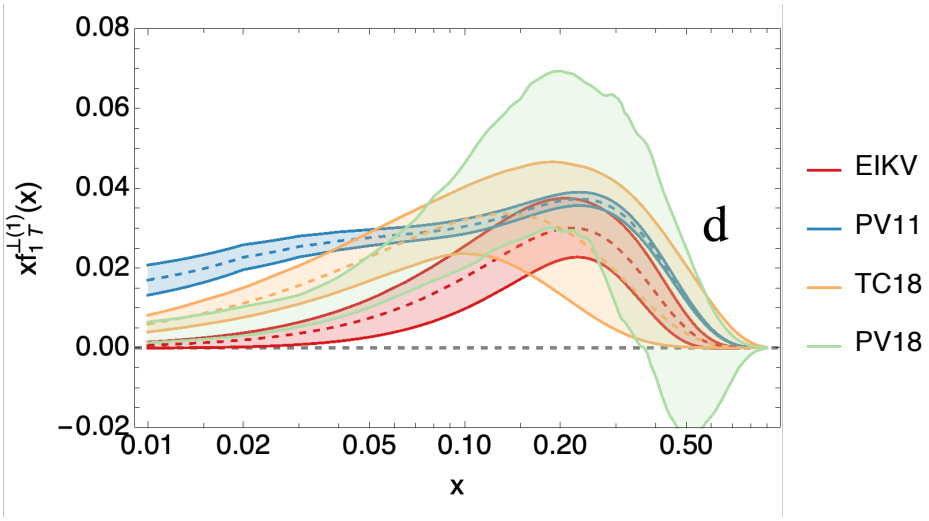


Figure 4.4: Comparison of different extractions of the first moment of Sivers distribution as a function of x for down quark.

4.6 Appendix: Plots of best fit results for Sivers asymmetries.

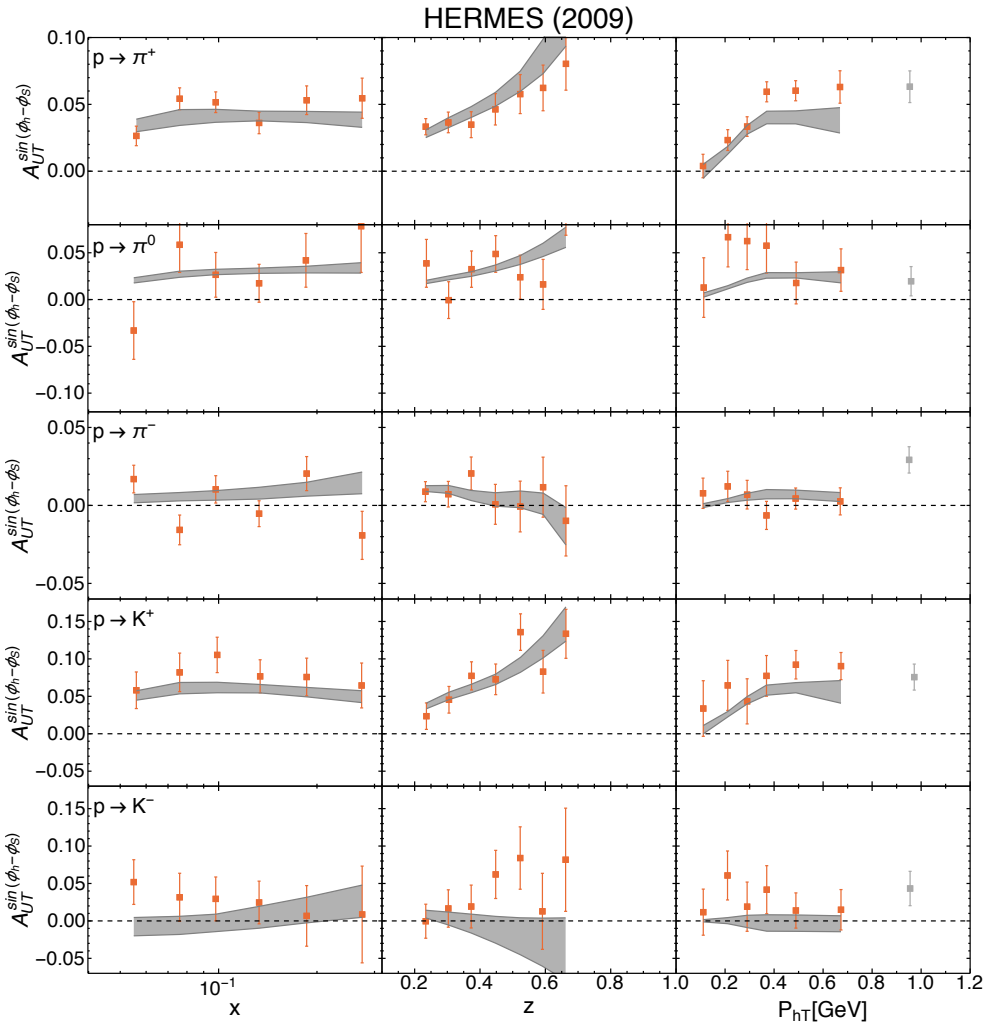


Figure 4.5: HERMES Sivers asymmetries from SIDIS off a proton target (H) with production of π^+, π^0, π^-, K^+, K^- in the final state, presented as function of x, z, P_{hT}.

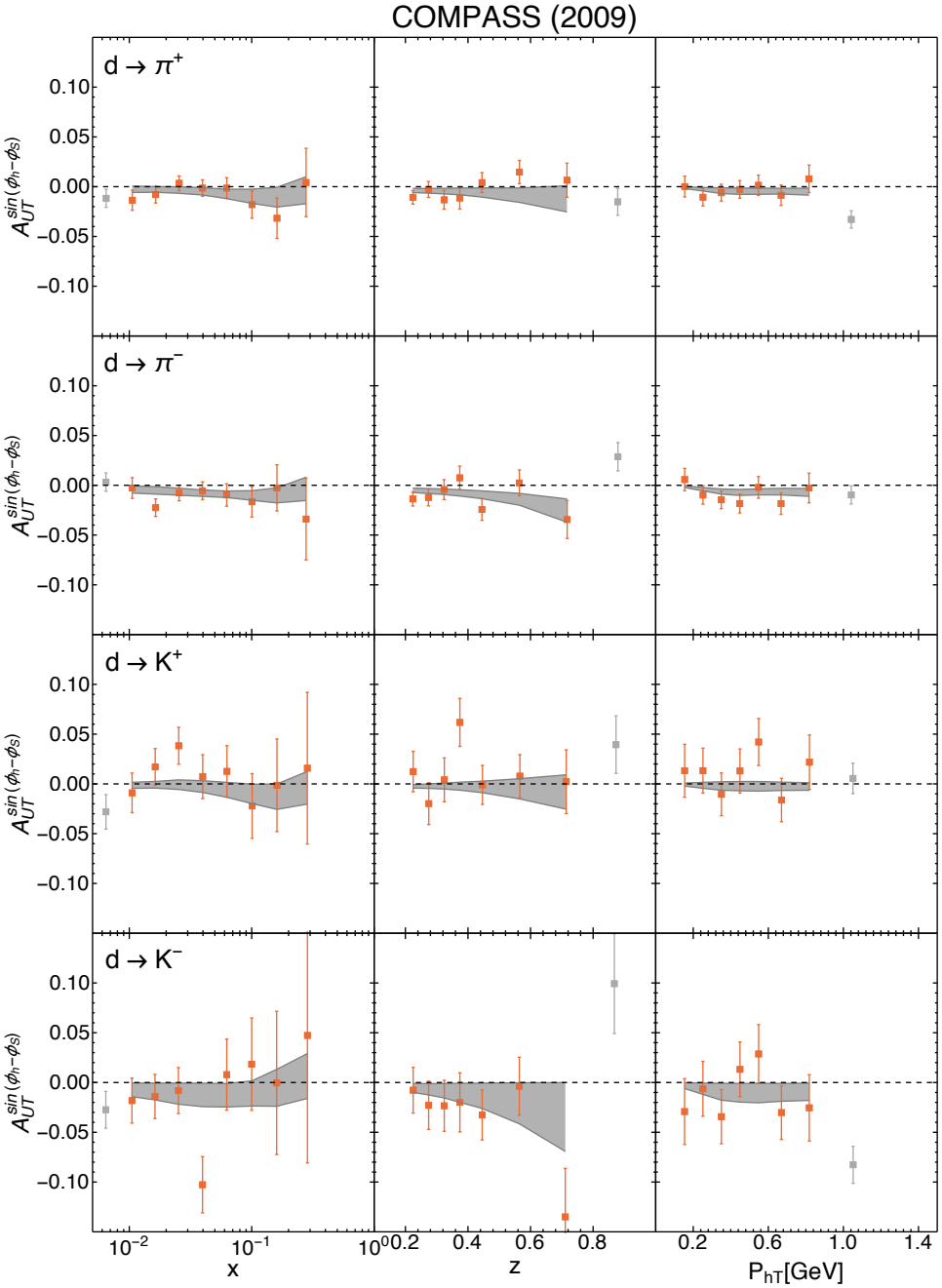


Figure 4.6: COMPASS'09 Sivers asymmetries from SIDIS off a deuteron target (^{6}LiD) with production of π^+, π^-, K^+, K^- in the final state, presented as function of x, z, P_{hT}.

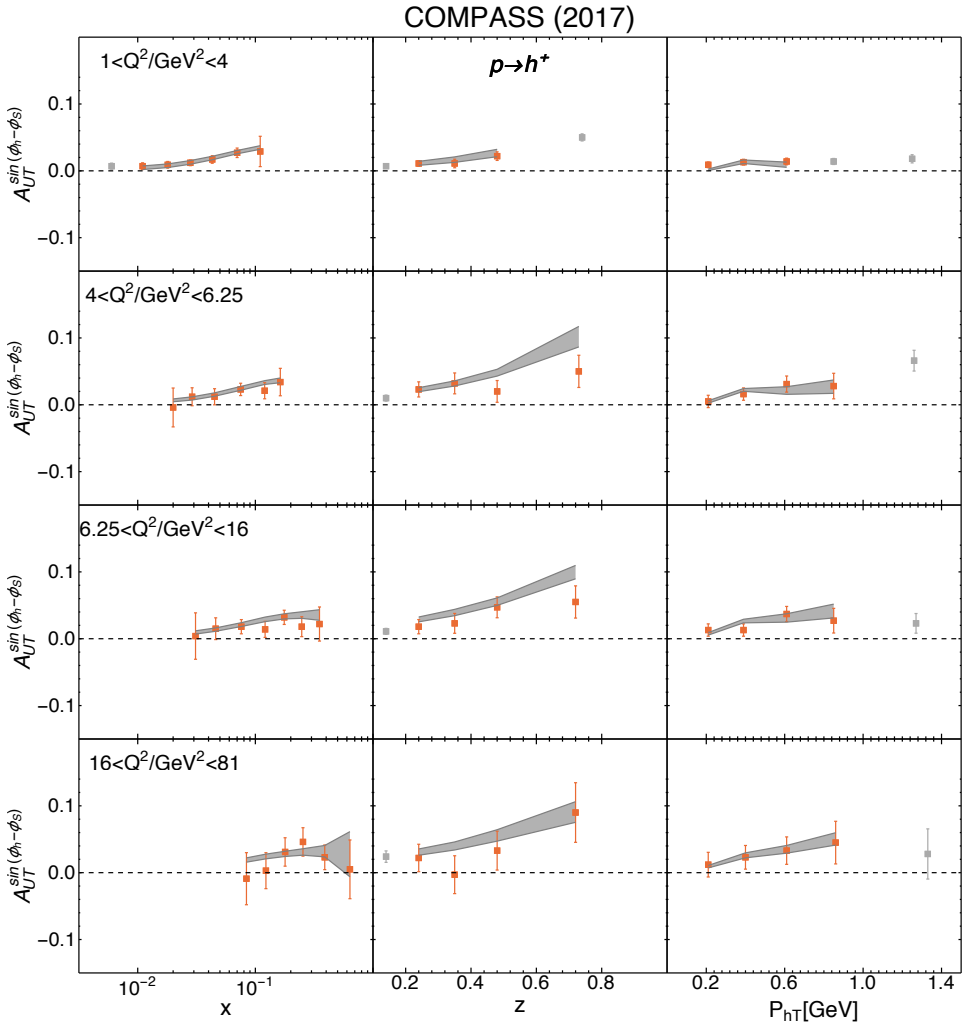


Figure 4.7: COMPASS'17 Sivers asymmetries from SIDIS off a proton target (NH_3) with production of positive hadrons h^+, presented as function of x, z, P_{hT} and divided in four different Q^2 bins.

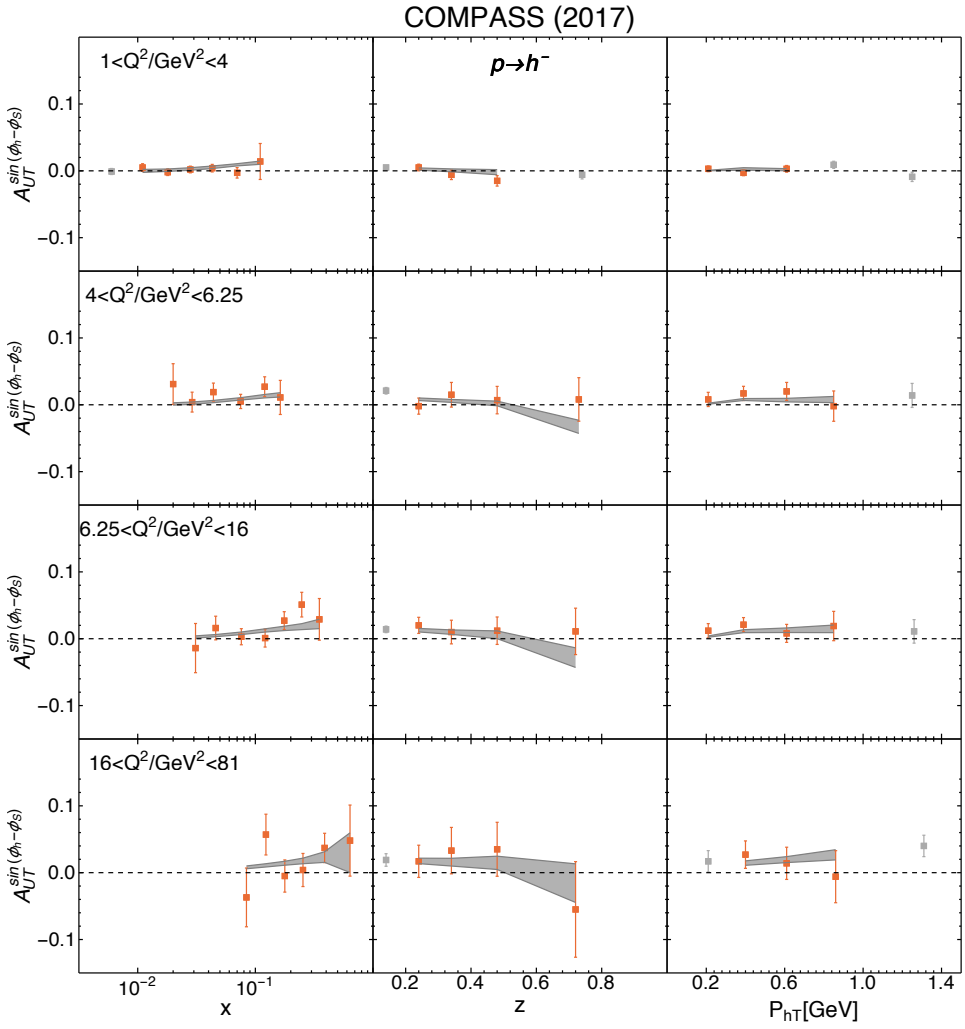


Figure 4.8: COMPASS'17 Sivers asymmetries from SIDIS off a proton target (NH_3) with production of negative hadrons h^-, presented as function of x, z, P_{hT} and divided in four different Q^2 bins.

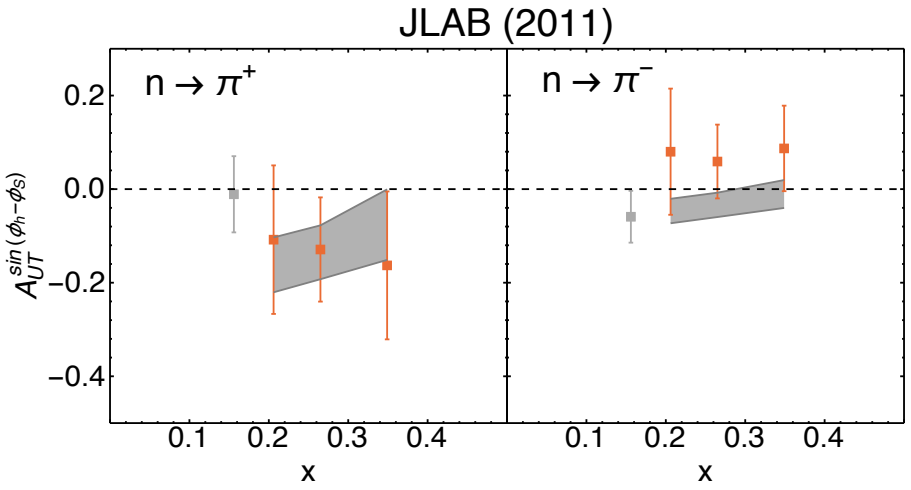


Figure 4.9: JLab Sivers asymmetries from SIDIS off a deuteron target (^{6}LiD) with production of positive and negative π in the final state, presented as function of x.

Chapter 5

Conclusions

In this thesis we presented a phenomenological study of unpolarized TMD parton distribution and fragmentation functions and, subsequently, of the Sivers distribution function. These two extraction are tightly connected, as the extraction of the Sivers function requires an accurate knowledge of the unpolarized TMDs. In both cases we paid particular attention to their Q^2-dependence, through the inclusion of TMD evolution contributions at NLL accuracy. Another interesting aspect of our minimization is the implementation of the replica methodology, a Monte Carlo approach to estimate the uncertainties of the functions obtained.

The extraction of the unpolarized TMDs can be considered as the first attempt at a global fit, in the context of TMD factorization and including also TMD evolution. Then, we effectively proved that, using the same TMD PDFs and FFs, it is possible to simultaneously describe data taken from SIDIS, Drell–Yan and Z boson production collected by different experiments. For the first time a phenomenological study tested the universality property of TMDs; using the same unpolarized TMDs we achieved a good description of a wide set of measurements covering different processes in different kinematic regions. Moreover, considering data at different values of Q^2 we were able to test the validity of the TMD evolution formalism. We extracted unpolarized TMDs using 8059 data points with 11 free parameters using a replica methodology. We selected data with $Q^2 > 1.4$ GeV^2 and $0.2 < z < 0.74$. We restricted our fit to the small transverse momentum region, selecting the maximum value of transverse momentum on the basis of phenomenological considerations (see Sec. 3.3). With these choices, we included regions where TMD factorization could be questioned, but we checked that our results describe very well the regions where TMD factorization is supposed to hold. Our fit is performed assuming that the intrinsic transverse momentum dependence of TMD PDFs and FFs can be parametrized by a normalized linear combination of a Gaussian and a weighted Gaussian, without considering an eventual dependence on parton flavors. We considered that the widths of the Gaussians depend on the longitudinal momenta. For the nonperturbative component of TMD evolution,

we adopted the choice most often used in the literature (see Sec. 3.2.3).

In general, the parametrization obtained from our fit gives a good description of the different experimental data sets, with an average χ^2/d.o.f. equal to 1.55 ± 0.05, including 8059 points with 11 free parameters. We have observed that restricting the kinematic region considered by excluding regions where the TMD factorization could be challenged, there is an even better match between our predictions and data, with an improved value of χ^2/d.o.f. of 1.02. Most of the discrepancies between experimental data and theory comes from the normalization and not from the transverse momentum shape. The results presented in this thesis, therefore, represent the most comprehensive determination of unpolarized TMDs available at the moment.

The three-dimensional internal structure of the nucleons in momentum space can be described through the accurate determination of TMDs. Using the results obtained from our fit, we were able to visualize the partonic 3D structure and to explore how it is modified by different values of x and Q^2, and what are its uncertainties. With our analysis we demonstrated that it is possible to reconstruct a map of hadrons in three dimensions through the study of high-energy scattering observables, which is currently one of the main purposes behind the multiple efforts to determine TMDs.

Having completed the determination of unpolarized TMDs, we used the obtained results to extract the Sivers distribution function, following a similar procedure. We considered the most updated measurements of polarized SIDIS azimuthal asymmetries, including data sets taken from Compass , Hermes , and JLab experiments. At variance with previous extractions, we limited our fit to the x projections, excluding the z and P_{hT} projection, in order to avoid examining fully correlated data.

We found a good agreement between our calculations of the azimuthal asymmetries and the data, with an average value of χ^2/d.o.f. $= 1.06 \pm 0.12$, including 118 data points and 14 free parameters. We were able to reproduce consistently also the curves for the data presented as a function of z and P_{hT} even if they were not fitted directly, thanks to the fact that we could rely on unpolarized TMD FF extracted from data. In this analysis we were able to distinguish the contributions of different flavors, obtaining different parameters for the u, d and sea quarks.

The analysis presented in this thesis is the first extraction of the Sivers function that implements in the description of azimuthal asymmetries an accurate functional form for unpolarized TMD PDF and FFs, obtained directly from experimental data. At the same time, our fit includes a coherent TMD evolution formalism to describe effects due to the hard scale variation. These features represent a significant progress in the procedure to determine the Sivers distribution.

Testing the formalism of TMD factorization and understanding the structure of unpolarized TMDs is only the first crucial step in the exploration of the 3D proton structure in momentum space and this work opens the way to

global determinations of TMDs. Building on this, we can proceed to deepen our understanding of hadron structure via polarized structure function and asymmetries (see, e.g., Refs. [155, 156]) and, at the same time, to test the impact of hadron structure in precision measurements at high-energies, such as at the LHC. A detailed mapping of hadron structure is essential to interpret data from hadronic collisions, which are among the most powerful tools to look for footprints of new physics.

5.1 Outlook

In order to overcome some limitations of the extraction presented in this thesis, we plan to improve some aspects of the procedure that we used. To increase the level of accuracy of our model, we should include also NLO contributions to the formalism adopted to describe the TMDs and their evolution. In future studies, different functional forms for all the nonperturbative ingredients should be explored, including also a possible flavor dependence of the intrinsic transverse momenta. A more precise analysis from the perturbative point of view is also needed, which should in principle make it possible to relax the tension in the normalization and to describe data at higher transverse momenta. Moreover, the description at low transverse momentum should be properly matched to the collinear fixed-order calculations at high transverse momentum.

Together with an improved theoretical framework, in order to better understand the formalism more experimental data is needed. It would be particularly useful to extend the coverage in x, z, rapidity, and Q^2. The 12 GeV physics program at Jefferson Lab [157] will be very important to constrain TMD distributions at large x. Additional data from SIDIS (at Compass, at a future Electron-Ion Collider), Drell–Yan (at Compass, at Fermilab, and at A Fixed-Target Experiment at the LHC [158]), Z/W production (at LHC, RHIC) will be very important. Measurements related to unpolarized TMD FFs at e^+e^- colliders (at Belle-II, BES-III, at a future International Linear Collider) will be invaluable, since they are presently missing.

Moreover, it would be useful to fit a wider group of data sets, in order to extend the coverage in rapidity, x, z and Q^2 . Currently, the data sets that could be included in our analysis, in addition to the ones already considered, are the recently published Compass SIDIS measurements with improved statistics, and the data from Z boson production at LHCb, CMS and ATLAS. In particular, the LHC data allow to cover a larger interval of Q^2, which could be useful to study more in detail the contributions related to TMD evolution. Additional data from Drell–Yan (at Compass, at Fermilab, and at A Fixed-Target Experiment at the LHC [158]), Z/W production (at LHC, RHIC) will be very important for the determination of TMDs. This updated extraction would probably allow us to study also the flavor dependence of unpolarized TMDs, which produced no particular effects in the current analysis. Otherwise, we

will have to include COMPASS proton data to achieve flavor separation.

In the next years, we expect further data that will be useful for the determination of TMDs. Our global fit is currently missing the independent determination of TMD fragmentation functions, which will be possible with the inclusion of data from e^+e^- annihilation that should become available from the Belle collaboration [159, 160]. Moreover, the 12 GeV physics program at Jefferson Lab [157] has started to collect a large amount of data after the upgrade, which will be available in the near future. These measurements will be very important to constrain TMD distributions at large x.

In the longer run, the LHC beam can also be used in fixed-target mode, either with gaseous or solid targets [161, 162], which would allow to probe TMD PDFs at highest energies in fixed-target kinematics (e.g., including the large-x region). The target will be unpolarized, but there are plans also to build a polarized target.

Finally, in the US there are plans to build a new Electron-Ion Collider (EIC) [163], which has been defined as one of the highest priorities for the Nuclear Physics community in the US and has been recently endorsed by the National Science Academy. This collider will be able to collide electrons with protons and light nuclei, and it will cover the kinematic area of low-x and high Q^2, currently unexplored. The data collected by this experiment will provide significant information for the mapping of the nucleon internal structure and the investigation of the relation between spin and partons.

Our work focused on quark TMDs. We remark that at present almost nothing is known experimentally about gluon TMDs [19, 164], because they typically require higher-energy scattering processes and they are harder to isolate as compared to quark distributions. Several promising measurements have been proposed in order to extract both the unpolarized and linearly polarized gluon TMDs inside an unpolarized proton. The cleanest possibility would be to look at dijet and heavy quark pair production in electron-proton collisions at a future EIC [165, 166]. Other proposals include isolated photon-pair production at RHIC [167] and quarkonium production at the LHC [168, 169, 170, 171].

As we mentioned earlier, an accurate description of the flavor dependence of the unpolarized TMDs will be necessary for precise measurements at high energies, such as the determination of the W boson mass.

With regards to the study of the Sivers function, there are many aspects that will require further attention. Having a detailed parametrization for the polarized TMDs is crucial for investigating the relation between the spin of the nucleons and the motion of their partons. It would be interesting to study the connection of the updated Sivers distribution function with the orbital angular momentum of the proton, improving for example the results obtained in the Pavia 2011 analysis.

Moreover, the A_N asymmetry measured in Drell–Yan processes presently have large errors. In the future, following the general idea of a global fit, we plan to include them in the extraction of the Sivers function.

In order to better evaluate the stability of our extraction, it could be useful to explore alternative functional forms for the Sivers function and study more in detail the flavor dependence, increasing eventually the number of flavor described by different parameters.

As we said this work proceeds in parallel with the study of the unpolarized TMDs, therefore, once we improve the unpolarized extraction, we plan to repeat the analysis of the Sivers function, including the updated results and eventually higher-order contributions.

Having a detailed parametrization for the polarized TMDs is crucial for investigating the relation between the spin of the nucleons and the motion of their partons. This updated extraction of the Sivers distribution, therefore, will be useful to study the orbital angular momentum of protons, improving for example the results obtained in the Pavia 2011 analysis.

The study of the internal structure of the hadron is in fact an ongoing process, where the progress in the experimental observations of high-energy processes stimulates new developments in their theoretical description and vice versa, leading with each step to a deeper understanding of the phenomena in particle physics.

Acknowledgments

I would like to thank all the people whose support support has been fundamental during the years of my doctorate. I am deeply grateful to Alessandro Bacchetta for being a guide and inspiration during my studies and for having always found the time and patience to deal with my questions and doubts, despite the never-ending list of deadlines and responsibilities. Thanks to him, it has been possible to create an active and diverse group of research which shares the same passion for the investigation of the fascinating landscape which appears at subnuclear dimensions.

I want to express my gratitude to the component of our research group. I would like to thank Barbara Pasquini who, since the Quantum Mechanics course, has given me fundamental advices, always supported by her extensive knowledge and rare humanity. I want to acknowledge the support of Marco Radici who often gave me a concrete solution when I was stuck in a seemingly overwhelming problem, and helped me fill the gap in my knowledge of the wine culture during travels and conferences.

I want to say thanks to another valued travel and office companion, Luca Mantovani, with whom I tried to bring some Italian stereotypes all over the world, from to US to Russia. Moreover, he introduced me to the other components of the Beto group and to their entertaining discussions and exhausting volleyball trainings. During these years I spent many moments with my PhD colleagues, Fulvio Piacenza and Chiara Bissolotti, which I would like to thank for having shared the difficulties and hopes of our studies and for having helped me to broaden my knowledge discussing the questions which arises everyday from our research.

My analysis would probably be still incomplete without the fundamental contribution of Cristian Pisano. I would also like to thank Giuseppe Bozzi, for his useful insights and refreshing humor, Pieter Taels, for the interesting discussions and for keeping alive in our group the ancient Belgian tradition of appreciating fine beer and celebrating Sinterklaas, and Miguel Echevarria for his important suggestions which helped me during the last year of my doctorate.

I am greatly thankful to Piet Mulders for having offered me the possibility

to work with his research group in Amsterdam. This offered me the occasion to work in a really stimulating environment with great people, Sabrina Cotogno and Valerio Bertone. I really appreciated the lengthy discussion about physics with the group of italian PhD students, and the even longer ones about politics.

Moreover, I want to thank Alberto Accardi and Harut Avakian for the interesting discussions during their multiple visits in Pavia, which I'm sure will continue during the next stage of my research at JLab. A special thanks to Andrea Signori, which followed this path before me and managed to find the right words to encourage me at every important step. Finally, I would like to express my gratitude to the referees, Francesco Murgia and Gunar Schnell, for taking time to evaluate my thesis and for their accurate remarks.

However, not only physicists have been important for the achievement of my doctorate goals. First of all, I am grateful to my family who greatly supported my during all my life and particularly during my education, giving me the means and the right advices to help me become who I am today.

I would like to thank you also my lifelong group of friends which became my second family in Pavia. In particular, I want say thanks to Mosco, whose everlasting research for a deeper understanding of the physical world is able to give a clearer vision to the people around him, usually promptly blurred by a pint of excellent beer; to Bos, whose kindness and incredible skills always create the perfect atmosphere in our group; to Maffo, a good comrade in everyday life and during our long motorbike travels; to Rumi, who always reminds us of the spiritual side of the existence, not opposed but coexistent to its physical description; and to Zenk, who under the exterior veil of calmness hides a creative restlessness that is always fascinating to witness.

Together we created Sborgomeo, a house who welcomed during these years many amazing people, interested in studying as much as in having a good time, such as Sus, Pol, Andre, Ades, Tobias, Camille, Margaux, Antonio, Paolo, Chiara, Fede who helped me broaden my horizons by sharing with me slices of their life.

I want to thank also the Physics students who became my friends, Marco, Mehdi, Luca, Euge, Alberto and Edo and the medical students with whom I spent many beautiful moments. Finally, I want to acknowledge the importance of my friends in Bergamo, I'm really glad that during these years, coming back from Pavia, I was sure to be welcomed by a group of amazing people, always with the same warmness and joy.

Bibliography

[1] H. Gegier and E. Marsden. On a diffuse reflection of the α-particles. *Proc. R. Soc. Lond. A*, 82(557):495–500, 1909.

[2] R. W. McAllister and R. Hofstadter. Elastic scattering of 188-MeV electrons from the proton and the alpha particle. *Phys. Rev.*, 102:851–856, May 1956.

[3] F. Bumiller, M. Croissiaux, E. Dally, and R. Hofstadter. Electromagnetic form factors of the proton. *Phys. Rev.*, 124:1623–1631, Dec 1961.

[4] E. D. Bloom et al. High-Energy Inelastic e p Scattering at 6-Degrees and 10-Degrees. *Phys. Rev. Lett.*, 23:930–934, 1969.

[5] M. Riordan. The discovery of quarks. *Science*, 256(5061):1287–1293, 1992.

[6] M. Gell-Mann and Y. Neemam. *The Eightfold way: a review with a collection of reprints*, 1964.

[7] M. Gell-Mann. A Schematic Model of Baryons and Mesons. *Phys. Lett.*, 8:214–215, 1964.

[8] J. D. Bjorken and E. A. Paschos. Inelastic electron-proton and γ-proton scattering and the structure of the nucleon. *Phys. Rev.*, 185:1975–1982, Sep 1969.

[9] R. Feynman, R. Field, and G. Fox. Correlations among particles and jets produced with large transverse momenta. *Nuclear Physics B*, 128(1):1 – 65, 1977.

[10] R. P. Feynman, R. D. Field, and G. C. Fox. Quantum-chromodynamic approach for the large-transverse-momentum production of particles and jets. *Phys. Rev. D*, 18:3320–3343, Nov 1978.

[11] 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.*, B206:364, 1988. [,340(1987)].

[12] A. Bacchetta, G. Bozzi, M. Radici, M. Ritzmann, and A. Signori. Effect of flavor-dependent partonic transverse momentum on the determination of the w boson mass in hadronic collisions. *Physics Letters B*, 2018.

[13] A. Bacchetta and M. Contalbrigo. The proton in 3D. *Il Nuovo Saggiatore*, 28:16–27, 2012.

[14] J. C. Collins and D. E. Soper. Back-To-Back Jets in QCD. *Nucl. Phys.*, B193:381, 1981. [Erratum: Nucl. Phys.B213,545(1983)].

[15] P. J. Mulders and R. D. Tangerman. The Complete tree level result up to order 1/Q for polarized deep inelastic leptoproduction. *Nucl. Phys.*, B461:197–237, 1996. [Erratum: Nucl. Phys.B484,538(1997)].

[16] A. M. Kotzinian and P. J. Mulders. Longitudinal quark polarization in transversely polarized nucleons. *Phys. Rev.*, D54:1229–1232, 1996.

[17] D. Boer and P. J. Mulders. Time-reversal odd distribution functions in leptoproduction. *Phys. Rev.*, D57:5780–5786, 1998.

[18] A. Bacchetta and P. J. Mulders. Deep inelastic leptoproduction of spin-one hadrons. *Phys. Rev.*, D62:114004, 2000.

[19] P. J. Mulders and J. Rodrigues. Transverse momentum dependence in gluon distribution and fragmentation functions. *Phys. Rev.*, D63:094021, 2001.

[20] D. Boer, S. Cotogno, T. van Daal, P. J. Mulders, A. Signori, and Y.-J. Zhou. Gluon and Wilson loop TMDs for hadrons of spin ≤ 1. *JHEP*, 10:013, 2016.

[21] J. C. Collins, D. E. Soper, and G. F. Sterman. Factorization of Hard Processes in QCD. *Adv. Ser. Direct. High Energy Phys.*, 5:1–91, 1989.

[22] J. Collins. *Foundations of perturbative QCD*. Cambridge University Press, 2013.

[23] A. Bacchetta and M. Radici. Constraining quark angular momentum through semi-inclusive measurements. *Phys. Rev. Lett.*, 107:212001, 2011.

[24] M. Anselmino, M. Boglione, and S. Melis. A Strategy towards the extraction of the Sivers function with TMD evolution. *Phys. Rev.*, D86:014028, 2012.

[25] M. G. Echevarria, A. Idilbi, Z.-B. Kang, and I. Vitev. QCD Evolution of the Sivers Asymmetry. *Phys. Rev.*, D89:074013, 2014.

[26] 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.

[27] Z. Lu and I. Schmidt. Updating Boer-Mulders functions from unpolarized pd and pp Drell-Yan data. *Phys. Rev.*, D81:034023, 2010.

[28] V. Barone, M. Boglione, J. O. Gonzalez Hernandez, and S. Melis. Phenomenological analysis of azimuthal asymmetries in unpolarized semi-inclusive deep inelastic scattering. *Phys. Rev.*, D91(7):074019, 2015.

[29] C. Lefky and A. Prokudin. Extraction of the distribution function $h_{1T}^{\perp}$ from experimental data. *Phys. Rev.*, D91(3):034010, 2015.

[30] 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.*, D87:094019, 2013.

[31] Z.-B. Kang, A. Prokudin, P. Sun, and F. Yuan. Extraction of Quark Transversity Distribution and Collins Fragmentation Functions with QCD Evolution. *Phys. Rev.*, D93(1):014009, 2016.

[32] D. Sivers. Single-spin production asymmetries from the hard scattering of pointlike constituents. *Phys. Rev. D*, 41:83–90, Jan 1990.

[33] E. C. Aschenauer, U. D'Alesio, and F. Murgia. TMDs and SSAs in hadronic interactions. *The European Physical Journal A*, 52(6):156, Jun 2016.

[34] U. D'Alesio and F. Murgia. Azimuthal and single spin asymmetries in hard scattering processes. *Progress in Particle and Nuclear Physics*, 61(2):394 – 454, 2008.

[35] X. Ji, J.-W. Qiu, W. Vogelsang, and F. Yuan. Unified picture for single transverse-spin asymmetries in hard-scattering processes. *Phys. Rev. Lett.*, 97:082002, Aug 2006.

[36] D. Sivers. Hard-scattering scaling laws for single-spin production asymmetries. *Phys. Rev. D*, 43:261–263, Jan 1991.

[37] J. Collins. Fragmentation of transversely polarized quarks probed in transverse momentum distributions. *Nuclear Physics B*, 396(1):161 – 182, 1993.

[38] S. J. Brodsky, D. S. Hwang, and I. Schmidt. Final-state interactions and single-spin asymmetries in semi-inclusive deep inelastic scattering. *Physics Letters B*, 530(1):99 – 107, 2002.

[39] S. J. Brodsky, D. S. Hwang, and I. Schmidt. Initial-state interactions and single-spin asymmetries in drell–yan processes. *Nuclear Physics B*, 642(1):344 – 356, 2002.

[40] J. C. Collins. Leading twist single transverse-spin asymmetries: Drell-Yan and deep inelastic scattering. *Phys. Lett.*, B536:43–48, 2002.

[41] A. Bacchetta, P. J. Mulders, and F. Pijlman. New observables in longitudinal single-spin asymmetries in semi-inclusive DIS. *Phys. Lett.*, B595:309–317, 2004.

[42] R. L. Jaffe. Can transversity be measured? In *Deep inelastic scattering off polarized targets: Theory meets experiment. Physics with polarized protons at HERA. Proceedings, Workshops, SPIN'97, Zeuthen, Germany, September 1-5, 1997 and Hamburg, Germany, March-September 1997*, pages 167–180, 1997.

[43] R. L. Jaffe. Spin, twist and hadron structure in deep inelastic processes. In *The spin structure of the nucleon. Proceedings, International School of Nucleon Structure, 1st Course, Erice, Italy, August 3-10, 1995*, pages 42–129, 1996.

[44] A. Bacchetta, U. D'Alesio, M. Diehl, and C. A. Miller. Single-spin asymmetries: The Trento conventions. *Phys. Rev.*, D70:117504, 2004.

[45] H. H. Matevosyan, W. Bentz, I. C. Cloet, and A. W. Thomas. Transverse Momentum Dependent Fragmentation and Quark Distribution Functions from the NJL-jet Model. *Phys. Rev.*, D85:014021, 2012.

[46] A. Bacchetta, M. Diehl, K. Goeke, A. Metz, P. J. Mulders, and M. Schlegel. Semi-inclusive deep inelastic scattering at small transverse momentum. *JHEP*, 02:093, 2007.

[47] A. Bacchetta, D. Boer, M. Diehl, and P. J. Mulders. Matches and mismatches in the descriptions of semi-inclusive processes at low and high transverse momentum. *JHEP*, 08:023, 2008.

[48] D. Boer, L. Gamberg, B. Musch, and A. Prokudin. Bessel-Weighted Asymmetries in Semi Inclusive Deep Inelastic Scattering. *JHEP*, 10:021, 2011.

[49] S. Arnold, A. Metz, and M. Schlegel. Dilepton production from polarized hadron hadron collisions. *Phys. Rev.*, D79:034005, 2009.

[50] J. C. Collins and D. E. Soper. Angular Distribution of Dileptons in High-Energy Hadron Collisions. *Phys. Rev.*, D16:2219, 1977.

[51] D. Boer, R. Jakob, and P. J. Mulders. Angular dependences in electroweak semi-inclusive leptoproduction. *Nucl. Phys.*, B564:471–485, 2000.

[52] A. Kotzinian. New quark distributions and semiinclusive electroproduction on the polarized nucleons. *Nucl. Phys.*, B441:234–248, 1995.

[53] C. J. Bomhof, P. J. Mulders, and F. Pijlman. Gauge link structure in quark-quark correlators in hard processes. *Phys. Lett.*, B596:277–286, 2004.

[54] C. J. Bomhof, P. J. Mulders, and F. Pijlman. The Construction of gauge-links in arbitrary hard processes. *Eur. Phys. J.*, C47:147–162, 2006.

[55] J. C. Collins, D. E. Soper, and G. F. Sterman. Transverse Momentum Distribution in Drell-Yan Pair and W and Z Boson Production. *Nucl. Phys.*, B250:199–224, 1985.

[56] M. Boglione, J. Collins, L. Gamberg, J. O. Gonzalez-Hernandez, T. C. Rogers, and N. Sato. Kinematics of Current Region Fragmentation in Semi-Inclusive Deeply Inelastic Scattering. *Phys. Lett.*, B766:245–253, 2017.

[57] Y. Koike, J. Nagashima, and W. Vogelsang. Resummation for polarized semi-inclusive deep-inelastic scattering at small transverse momentum. *Nucl. Phys.*, B744:59–79, 2006.

[58] A. Idilbi, X. Ji, J.-P. Ma, and F. Yuan. Collins-Soper equation for the energy evolution of transverse-momentum and spin dependent parton distributions. *Phys. Rev.*, D70:074021, 2004.

[59] S. M. Aybat and T. C. Rogers. TMD Parton Distribution and Fragmentation Functions with QCD Evolution. *Phys. Rev.*, D83:114042, 2011.

[60] A. Bacchetta, M. G. Echevarria, P. J. G. Mulders, M. Radici, and A. Signori. Effects of TMD evolution and partonic flavor on $e^+ e^-$ annihilation into hadrons. *JHEP*, 11:076, 2015.

[61] M. Srednicki. *Quantum field theory*. Cambridge University Press, 2007.

[62] C. T. H. Davies and W. J. Stirling. Nonleading Corrections to the Drell-Yan Cross-Section at Small Transverse Momentum. *Nucl. Phys.*, B244:337–348, 1984.

[63] W. Zimmermann. Normal products and the short distance expansion in the perturbation theory of renormalizable interactions. In P. Breitenlohner and D. Maison, editors, *Quantum Field Theory*, pages 278–309, Berlin, Heidelberg, 2000. Springer Berlin Heidelberg.

[64] M. G. Echevarria, A. Idilbi, and I. Scimemi. Unified treatment of the QCD evolution of all (un-)polarized transverse momentum dependent functions: Collins function as a study case. *Phys. Rev.*, D90(1):014003, 2014.

[65] J.-w. Qiu and G. F. Sterman. Single transverse spin asymmetries in hadronic pion production. *Phys. Rev.*, D59:014004, 1999.

[66] Z.-B. Kang, J.-W. Qiu, W. Vogelsang, and F. Yuan. An Observation Concerning the Process Dependence of the Sivers Functions. *Phys. Rev.*, D83:094001, 2011.

[67] Z.-B. Kang and J.-W. Qiu. QCD evolution of naive-time-reversal-odd parton distribution functions. *Phys. Lett.*, B713:273–276, 2012.

[68] S. M. Aybat, J. C. Collins, J.-W. Qiu, and T. C. Rogers. The QCD Evolution of the Sivers Function. *Phys. Rev.*, D85:034043, 2012.

[69] Z.-B. Kang and J.-W. Qiu. Evolution of twist-3 multi-parton correlation functions relevant to single transverse-spin asymmetry. *Phys. Rev.*, D79:016003, 2009.

[70] A. Bacchetta, F. Delcarro, C. Pisano, M. Radici, and A. Signori. Extraction of partonic transverse momentum distributions from semi-inclusive deep-inelastic scattering, Drell-Yan and Z-boson production. *JHEP*, 06:081, 2017.

[71] A. Signori. *Flavor and Evolution Effects in TMD Phenomenology*. PhD thesis, Vrije U., Amsterdam, 2016.

[72] A. Signori, A. Bacchetta, M. Radici, and G. Schnell. Investigations into the flavor dependence of partonic transverse momentum. *JHEP*, 11:194, 2013.

[73] T. C. Rogers. An Overview of Transverse Momentum Dependent Factorization and Evolution. *Eur. Phys. J.*, A52(6):153, 2016.

[74] J. Collins, L. Gamberg, A. Prokudin, T. C. Rogers, N. Sato, and B. Wang. Relating Transverse Momentum Dependent and Collinear Factorization Theorems in a Generalized Formalism. *Phys. Rev.*, D94(3):034014, 2016.

[75] D. Boer et al. Gluons and the quark sea at high energies: Distributions, polarization, tomography. 2011.

[76] S. Chekanov et al. Measurement of the Longitudinal Proton Structure Function at HERA. *Phys. Lett.*, B682:8–22, 2009.

[77] V. Andreev et al. Measurement of inclusive ep cross sections at high Q^2 at $\sqrt{s} = 225$ and 252 GeV and of the longitudinal proton structure function F_L at HERA. *Eur. Phys. J.*, C74(4):2814, 2014.

[78] M. Arneodo et al. Accurate measurement of F2(d) / F2(p) and R**d - R**p. *Nucl. Phys.*, B487:3–26, 1997.

[79] L. W. Whitlow, E. M. Riordan, S. Dasu, S. 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.*, B282:475–482, 1992.

[80] X. Ji and F. Yuan. Parton distributions in light cone gauge: Where are the final state interactions? *Phys. Lett.*, B543:66–72, 2002.

[81] X. Ji, J.-p. Ma, and F. Yuan. QCD factorization for semi-inclusive deep-inelastic scattering at low transverse momentum. *Phys. Rev.*, D71:034005, 2005.

[82] M. G. Echevarria, A. Idilbi, and I. Scimemi. Factorization Theorem For Drell-Yan At Low q_T And Transverse Momentum Distributions On-The-Light-Cone. *JHEP*, 07:002, 2012.

[83] M. G. Echevarria, A. Idilbi, A. Schäfer, and I. Scimemi. Model-Independent Evolution of Transverse Momentum Dependent Distribution Functions (TMDs) at NNLL. *Eur. Phys. J.*, C73(12):2636, 2013.

[84] J. C. Collins and T. C. Rogers. Equality of Two Definitions for Transverse Momentum Dependent Parton Distribution Functions. *Phys. Rev.*, D87(3):034018, 2013.

[85] E. Moffat, W. Melnitchouk, T. C. Rogers, and N. Sato. What are the low-Q and large-x boundaries of collinear QCD factorization theorems? 2017.

[86] M. G. Echevarria, I. Scimemi, and A. Vladimirov. Unpolarized Transverse Momentum Dependent Parton Distribution and Fragmentation Functions at next-to-next-to-leading order. *JHEP*, 09:004, 2016.

[87] D. Boer and W. Vogelsang. Drell-Yan lepton angular distribution at small transverse momentum. *Phys. Rev.*, D74:014004, 2006.

[88] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016.

[89] M. Lambertsen and W. Vogelsang. Drell-Yan lepton angular distributions in perturbative QCD. *Phys. Rev.*, D93(11):114013, 2016.

[90] G. Parisi and R. Petronzio. Small Transverse Momentum Distributions in Hard Processes. *Nucl. Phys.*, B154:427–440, 1979.

[91] G. Altarelli, R. K. Ellis, M. Greco, and G. Martinelli. Vector Boson Production at Colliders: A Theoretical Reappraisal. *Nucl. Phys.*, B246:12–44, 1984.

[92] G. Bozzi, S. Catani, G. Ferrera, D. de Florian, and M. Grazzini. Production of Drell-Yan lepton pairs in hadron collisions: Transverse-momentum resummation at next-to-next-to-leading logarithmic accuracy. *Phys. Lett.*, B696:207–213, 2011.

[93] D. Boer and W. J. den Dunnen. TMD evolution and the Higgs transverse momentum distribution. *Nucl. Phys.*, B886:421–435, 2014.

[94] E. Laenen, G. F. Sterman, and W. Vogelsang. Higher order QCD corrections in prompt photon production. *Phys. Rev. Lett.*, 84:4296–4299, 2000.

[95] J.-w. Qiu and X.-f. Zhang. Role of the nonperturbative input in QCD resummed Drell-Yan Q_T distributions. *Phys. Rev.*, D63:114011, 2001.

[96] S. Frixione, P. Nason, and G. Ridolfi. Problems in the resummation of soft gluon effects in the transverse momentum distributions of massive vector bosons in hadronic collisions. *Nucl. Phys.*, B542:311–328, 1999.

[97] G. Bozzi, S. Catani, D. de Florian, and M. Grazzini. Transverse-momentum resummation and the spectrum of the Higgs boson at the LHC. *Nucl. Phys.*, B737:73–120, 2006.

[98] P. M. Nadolsky, D. R. Stump, and C. P. Yuan. Semiinclusive hadron production at HERA: The Effect of QCD gluon resummation. *Phys. Rev.*, D61:014003, 2000. [Erratum: Phys. Rev.D64,059903(2001)].

[99] F. Landry, R. Brock, P. M. Nadolsky, and C. P. Yuan. Tevatron Run-1 Z boson data and Collins-Soper-Sterman resummation formalism. *Phys. Rev.*, D67:073016, 2003.

[100] A. V. Konychev and P. M. Nadolsky. Universality of the Collins-Soper-Sterman nonperturbative function in gauge boson production. *Phys. Lett.*, B633:710–714, 2006.

[101] C. A. Aidala, B. Field, L. P. Gamberg, and T. C. Rogers. Limits on transverse momentum dependent evolution from semi-inclusive deep inelastic scattering at moderate Q. *Phys. Rev.*, D89(9):094002, 2014.

[102] J. Collins and T. Rogers. Understanding the large-distance behavior of transverse-momentum-dependent parton densities and the Collins-Soper evolution kernel. *Phys. Rev.*, D91(7):074020, 2015.

[103] I. Scimemi and A. Vladimirov. Power corrections and renormalons in Transverse Momentum Distributions. *JHEP*, 03:002, 2017.

[104] U. D'Alesio, M. G. Echevarria, S. Melis, and I. Scimemi. Non-perturbative QCD effects in q_T spectra of Drell-Yan and Z-boson production. *JHEP*, 11:098, 2014.

[105] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr, and G. Watt. LHAPDF6: parton density access in the LHC precision era. *Eur. Phys. J.*, C75:132, 2015.

[106] M. Gluck, P. Jimenez-Delgado, and E. Reya. Dynamical parton distributions of the nucleon and very small-x physics. *Eur. Phys. J.*, C53:355–366, 2008.

[107] D. de Florian, R. Sassot, M. Epele, R. J. Hernández-Pinto, and M. Stratmann. Parton-to-Pion Fragmentation Reloaded. *Phys. Rev.*, D91(1):014035, 2015.

[108] D. de Florian, R. Sassot, and M. Stratmann. Global analysis of fragmentation functions for pions and kaons and their uncertainties. *Phys. Rev.*, D75:114010, 2007.

[109] D. de Florian, M. Epele, R. J. Hernandez-Pinto, R. Sassot, and M. Stratmann. Parton-to-Kaon Fragmentation Revisited. *Phys. Rev.*, D95(9):094019, 2017.

[110] A. Bacchetta, L. P. Gamberg, G. R. Goldstein, and A. Mukherjee. Collins fragmentation function for pions and kaons in a spectator model. *Phys. Lett.*, B659:234–243, 2008.

[111] B. Pasquini, S. Cazzaniga, and S. Boffi. Transverse momentum dependent parton distributions in a light-cone quark model. *Phys. Rev.*, D78:034025, 2008.

[112] H. Avakian, A. V. Efremov, P. Schweitzer, and F. Yuan. The transverse momentum dependent distribution functions in the bag model. *Phys. Rev.*, D81:074035, 2010.

[113] A. Bacchetta, M. Radici, F. Conti, and M. Guagnelli. Weighted azimuthal asymmetries in a diquark spectator model. *Eur. Phys. J.*, A45:373–388, 2010.

[114] M. Burkardt and B. Pasquini. Modelling the nucleon structure. *Eur. Phys. J.*, A52(6):161, 2016.

[115] A. Bacchetta, R. Kundu, A. Metz, and P. J. Mulders. Estimate of the Collins fragmentation function in a chiral invariant approach. *Phys. Rev.*, D65:094021, 2002.

[116] P. Schweitzer, M. Strikman, and C. Weiss. Intrinsic transverse momentum and parton correlations from dynamical chiral symmetry breaking. *JHEP*, 01:163, 2013.

[117] A. Airapetian et al. Multiplicities of charged pions and kaons from semi-inclusive deep-inelastic scattering by the proton and the deuteron. *Phys. Rev.*, D87:074029, 2013.

[118] C. Adolph et al. Hadron Transverse Momentum Distributions in Muon Deep Inelastic Scattering at 160 GeV/c. *Eur. Phys. J.*, C73(8):2531, 2013. [Erratum: Eur. Phys. J.C75,no.2,94(2015)].

[119] M. Anselmino, M. Boglione, J. O. Gonzalez Hernandez, S. Melis, and A. Prokudin. Unpolarised Transverse Momentum Dependent Distribution and Fragmentation Functions from SIDIS Multiplicities. *JHEP*, 04:005, 2014.

[120] M. Aghasyan et al. Transverse-momentum-dependent Multiplicities of Charged Hadrons in Muon-Deuteron Deep Inelastic Scattering. *Phys. Rev.*, D97(3):032006, 2018.

[121] F. Landry, R. Brock, G. Ladinsky, and C. P. Yuan. New fits for the nonperturbative parameters in the CSS resummation formalism. *Phys. Rev.*, D63:013004, 2001.

[122] U. D'Alesio, M. G. Echevarria, S. Melis, and I. Scimemi. Non-perturbative QCD effects in q_T spectra of Drell-Yan and Z-boson production. *JHEP*, 11:098, 2014.

[123] A. S. Ito et al. Measurement of the Continuum of Dimuons Produced in High-Energy Proton - Nucleus Collisions. *Phys. Rev.*, D23:604–633, 1981.

[124] G. Moreno et al. Dimuon production in proton - copper collisions at $\sqrt{s}$ = 38.8-GeV. *Phys. Rev.*, D43:2815–2836, 1991.

[125] T. Affolder et al. The transverse momentum and total cross section of e^+e^- pairs in the Z boson region from $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV. *Phys. Rev. Lett.*, 84:845–850, 2000.

[126] B. Abbott et al. Measurement of the inclusive differential cross section for Z bosons as a function of transverse momentum in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV. *Phys. Rev.*, D61:032004, 2000.

[127] T. Aaltonen et al. Transverse momentum cross section of e^+e^- pairs in the Z-boson region from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev.*, D86:052010, 2012.

[128] V. M. Abazov et al. Measurement of the shape of the boson transverse momentum distribution in $p\bar{p} \to Z/\gamma^* \to e^+e^- + X$ events produced at $\sqrt{s}$=1.96-TeV. *Phys. Rev. Lett.*, 100:102002, 2008.

[129] A. Abulencia et al. Measurements of inclusive W and Z cross sections in p anti-p collisions at s**(1/2) = 1.96-TeV. *J. Phys.*, G34:2457–2544, 2007.

[130] S. Catani, L. Cieri, G. Ferrera, D. de Florian, and M. Grazzini. Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO. *Phys. Rev. Lett.*, 103:082001, 2009.

[131] A. Bacchetta, A. Courtoy, and M. Radici. First extraction of valence transversities in a collinear framework. *JHEP*, 03:119, 2013.

[132] M. Radici, A. Courtoy, A. Bacchetta, and M. Guagnelli. Improved extraction of valence transversity distributions from inclusive dihadron production. *JHEP*, 05:123, 2015.

[133] S. Forte, L. Garrido, J. I. Latorre, and A. Piccione. Neural network parametrization of deep inelastic structure functions. *JHEP*, 05:062, 2002.

[134] R. D. Ball, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, A. Piccione, J. Rojo, and M. Ubiali. A Determination of parton distributions with faithful uncertainty estimation. *Nucl. Phys.*, B809:1–63, 2009. [Erratum: Nucl. Phys.B816,293(2009)].

[135] R. D. Ball, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, J. Rojo, and M. Ubiali. A first unbiased global NLO determination of parton distributions and their uncertainties. *Nucl. Phys.*, B838:136–206, 2010.

[136] M. Epele, R. Llubaroff, R. Sassot, and M. Stratmann. Uncertainties in pion and kaon fragmentation functions. *Phys. Rev.*, D86:074028, 2012.

[137] F. James and M. Roos. Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations. *Comput. Phys. Commun.*, 10:343–367, 1975.

[138] S. Aid et al. Transverse energy and forward jet production in the low x regime at HERA. *Phys. Lett.*, B356:118–128, 1995.

[139] G. A. Ladinsky and C. P. Yuan. The Nonperturbative regime in QCD resummation for gauge boson production at hadron colliders. *Phys. Rev.*, D50:R4239, 1994.

[140] P. Sun, J. Isaacson, C. P. Yuan, and F. Yuan. Nonperturbative functions for SIDIS and Drell–Yan processes. *Int. J. Mod. Phys.*, A33(11):1841006, 2018.

[141] I. Scimemi and A. Vladimirov. Analysis of vector boson production within TMD factorization. *Eur. Phys. J.*, C78(2):89, 2018.

[142] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J.*, C63:189–285, 2009.

[143] J. F. Owens, A. Accardi, and W. Melnitchouk. Global parton distributions with nuclear and finite-Q^2 corrections. *Phys. Rev.*, D87(9):094012, 2013.

[144] P. Schweitzer, T. Teckentrup, and A. Metz. Intrinsic transverse parton momenta in deeply inelastic reactions. *Phys. Rev.*, D81:094019, 2010.

[145] A. Bacchetta, M. Boglione, A. Henneman, and P. J. Mulders. Bounds on transverse momentum dependent distribution and fragmentation functions. *Phys. Rev. Lett.*, 85:712–715, 2000.

[146] M. Radici and A. Bacchetta. First Extraction of Transversity from a Global Analysis of Electron-Proton and Proton-Proton Data. *Phys. Rev. Lett.*, 120(19):192001, 2018.

[147] S. M. Aybat, J. C. Collins, J. W. Qiu, and T. C. Rogers. Qcd evolution of the sivers function. *Phys. Rev. D*, 85:034043, Feb 2012.

[148] A. Airapetian et al. Observation of the Naive-T-odd Sivers Effect in Deep-Inelastic Scattering. *Phys. Rev. Lett.*, 103:152002, 2009.

[149] M. Alekseev et al. Collins and Sivers asymmetries for pions and kaons in muon-deuteron DIS. *Phys. Lett.*, B673:127–135, 2009.

[150] C. Adolph et al. Sivers asymmetry extracted in SIDIS at the hard scales of the Drell–Yan process at COMPASS. *Phys. Lett.*, B770:138–145, 2017.

[151] C. Adolph et al. Collins and Sivers asymmetries in muonproduction of pions and kaons off transversely polarised protons. *Phys. Lett.*, B744:250–259, 2015.

[152] X. Qian et al. Single Spin Asymmetries in Charged Pion Production from Semi-Inclusive Deep Inelastic Scattering on a Transversely Polarized ^{3}He Target. *Phys. Rev. Lett.*, 107:072003, 2011.

[153] M. Boglione, U. D'Alesio, C. Flore, and J. O. Gonzalez-Hernandez. Assessing signals of TMD physics in SIDIS azimuthal asymmetries and in the extraction of the Sivers function. *JHEP*, 07:148, 2018.

[154] M. Anselmino, M. Boglione, U. D'Alesio, A. Kotzinian, S. Melis, F. Murgia, A. Prokudin, and C. Turk. Sivers Effect for Pion and Kaon Production in Semi-Inclusive Deep Inelastic Scattering. *Eur. Phys. J.*, A39:89–100, 2009.

[155] E. C. Aschenauer, U. D'Alesio, and F. Murgia. TMDs and SSAs in hadronic interactions. *Eur. Phys. J.*, A52(6):156, 2016.

[156] M. Boglione and A. Prokudin. Phenomenology of transverse spin: past, present and future. *Eur. Phys. J.*, A52(6):154, 2016.

[157] J. Dudek et al. Physics Opportunities with the 12 GeV Upgrade at Jefferson Lab. *Eur. Phys. J.*, A48:187, 2012.

[158] S. J. Brodsky, F. Fleuret, C. Hadjidakis, and J. P. Lansberg. Physics Opportunities of a Fixed-Target Experiment using the LHC Beams. *Phys. Rept.*, 522:239–255, 2013.

[159] A. Abdesselam et al. Observation of Transverse $\Lambda/\bar{\Lambda}$ Hyperon Polarization in e^+e^- Annihilation at Belle, August 2018. arXiv:1808.05000.

[160] 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.*, D96(3):032005, 2017.

[161] L. Massacrier et al. Physics perspectives with AFTER@LHC (A Fixed Target ExpeRiment at LHC). *EPJ Web Conf.*, 171:10001, 2018.

[162] R. Aaij et al. Measurement of antiproton production in pHe collisions at $\sqrt{s_{NN}} = 110$ GeV, August 2018. arXiv:1808.06127.

[163] A. Accardi et al. Electron Ion Collider: The Next QCD Frontier. *Eur. Phys. J.*, A52(9):268, 2016.

[164] M. G. Echevarria, T. Kasemets, P. J. Mulders, and C. Pisano. QCD evolution of (un)polarized gluon TMDPDFs and the Higgs q_T-distribution. *JHEP*, 07:158, 2015.

[165] D. Boer, S. J. Brodsky, P. J. Mulders, and C. Pisano. Direct Probes of Linearly Polarized Gluons inside Unpolarized Hadrons. *Phys. Rev. Lett.*, 106:132001, 2011.

[166] C. Pisano, D. Boer, S. J. Brodsky, M. G. A. Buffing, and P. J. Mulders. Linear polarization of gluons and photons in unpolarized collider experiments. *JHEP*, 10:024, 2013.

[167] J.-W. Qiu, M. Schlegel, and W. Vogelsang. Probing Gluonic Spin-Orbit Correlations in Photon Pair Production. *Phys. Rev. Lett.*, 107:062001, 2011.

[168] D. Boer and C. Pisano. Polarized gluon studies with charmonium and bottomonium at LHCb and AFTER. *Phys. Rev.*, D86:094007, 2012.

[169] W. J. den Dunnen, J. P. Lansberg, C. Pisano, and M. Schlegel. Accessing the Transverse Dynamics and Polarization of Gluons inside the Proton at the LHC. *Phys. Rev. Lett.*, 112:212001, 2014.

[170] A. Signori. Gluon TMDs in quarkonium production. *Few Body Syst.*, 57(8):651–655, 2016.

[171] J.-P. Lansberg, C. Pisano, and M. Schlegel. Associated production of a dilepton and a $\Upsilon(J/\psi)$ at the LHC as a probe of gluon transverse momentum dependent distributions. *Nucl. Phys.*, B920:192–210, 2017.