Università degli Studi di Pavia
Dipartimento di Fisica

Istituto Nazionale di Fisica Nucleare

Dottorato di Ricerca in Fisica – XXXII Ciclo

Perturbative and nonperturbative QCD regimes in transverse-momentum dependent observables

Fulvio Piacenza

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: whiteboard in our office.

Perturbative and nonperturbative QCD regimes in transverse-momentum dependent observables
Fulvio Piacenza
PhD thesis - University of Pavia
Pavia, Italy, September 2019

La voce sottile della verità è la stessa dei sentimenti puri.

Contents

Chapter 1

Introduction

Exactly 50 years before the writing of this thesis, SLAC-MIT breakthrough experiments on high-energy electron-proton scattering [1, 2] showed for the first time that the proton was made of point-like constituents. Our understanding of the strong force (i.e., the force which binds together quarks and gluons) has made major steps since then, thanks to the joint efforts of an ever-growing community of researchers–and PhD students–around the world. Quantum ChromoDynamics (QCD) is one of the building blocks of the Standard Model, and precise QCD predictions are crucial for physics studies at hadron colliders. However, despite the many achievements of physics in the last decades, we cannot say yet that we have a true comprehension of visible matter. We lack a proof for color confinement, i.e., for the fact that colored particles are never observed alone, and we do not have a clear understanding of how the physical properties of hadrons emerge from the interaction of their constituents. For instance, we are not yet able to say how the spin of the proton is generated from the spin of quarks and gluons and from their orbital angular momentum, nor we clearly know how the nucleon constituents are distributed in position and momentum.

The reason for our knowledge gaps is that QCD at hadronic scales reaches the region where perturbation theory breaks down, due to the strength of the coupling, so that theoretical calculations must resort to nonperturbative techniques. One such technique is the calculation of Green functions on a discretized space-time lattice (lattice QCD), but there is a limited number of observables that can be presently calculated to good accuracy with this method. Besides lattice calculations, the information we presently have on hadron physics comes from experimental measurements. Factorization theorems allow us to reliably separate cross sections into perturbative parts, which describe the actual high energy scattering of quarks and gluons involved in the collision, and nonperturbative matrix elements, which give the momentum and space distribution of these partons[1] inside the nucleon. The most well-known example of nonperturbative functions extracted in this way is given by Parton Distribution Functions (PDFs). In a simplified picture, these give the probability distribution for a parton to carry a certain fraction x of the nucleon momentum, in a frame where this momentum is light-like. For PDFs a direct calculation on the lattice—hence with Euclidean time—is not possible, since this involves matrix elements of bilocal operators separated along the light-cone. Only recently, pioneering methods to circumvent this problem have been proposed [3]. Commonly extracted from the simultaneous analysis of different experiments, PDFs have reached an accuracy of few % over a wide kinematic range and are continuously updated, also thanks to LHC data (see, e.g., [4–6]).

A study restricted to the longitudinal (light-like) momentum component of partons is however not sufficient to unravel the mysteries of QCD dynamics inside hadrons, as a complete description should include transverse momentum, which is an independent degree of freedom. Moreover, we also need information on spatial coordinates. Indeed, we cannot directly measure a wave-function in experiments, but we can think of reconstructing it from the phase-space distribution of the system, i.e., the analogue of the classical Liouville density for the case of a quantum system. This kind of distributions are known as Wigner distributions. Due to the uncertainty principle, they cannot be interpreted as probability distributions, and are not even positive definite. In any case, they can be used to calculate any expectation value of a system. Thus, to have the maximal information on its state, we would have to measure the distribution of a parton in both momentum and space coordinates.

Generalized Transverse Momentum Distributions (GMTDs), which can be thought of as the Fourier-transforms of Wigner distributions, may potentially have a connection with experiments. However, devising suitable observables for the extraction of GTMDs is complicated, and only recently some proposals have been presented (see, e.g., [7, 8] and references therein). In the meantime, Transverse Momentum Distributions (TMDs), and Generalized Parton Distributions (GPDs),[2] which correspond to projections of GTMDs, have been extensively studied on the formal side, and some extractions from data have been performed. The uncertainties on this functions, though, are still large and their key properties are often not yet well determined.

Of the above functions, in this thesis we will concentrate on TMDs, describing the probability distribution for partonic transverse momentum inside a nucleon. Intrinsic transverse momentum received a lot of attention in recent years, due

---

[1] "Parton" is used to generally denote quarks and gluons, without distinction.

[2] Generalized Parton Distributions (GPDs) describe the distribution of partons in the impact parameter sapce as well as the longitudinal momentum x. See [9] for a theoretical review.

to its connection to the big puzzles of the strong force, but its potential role in experimental observables was conjectured since the early days of QCD [10, 11]: it was known that the transverse momentum (with respect to the beam axis) of particles produced in hadronic collisions could originate either from the motion of partons inside the nucleon or from the recoil against QCD radiation during the hard scattering. The first clear evidences for intrinsic transverse momentum were shown by Drell-Yan[3] cross sections measured by hadronic accelerators at Fermilab and CERN in the late '70s – early '80s [13, 14]. The observed rates displayed a gaussian-like dependence on q_T,[4] for $q_T \lesssim 1$ GeV, whose width was compatible with typical momentum scales of quarks inside the proton (see, e.g., [15]). However, theoretically sound studies had to wait for the improvement of perturbative QCD techniques: QCD predicts a strong enhancement of gluon emission at low q_T, requiring the summation of infinitely many emissions and the inclusion of a nonperturbative component to obtain a reliable result. A formalism for the resummation of large contributions in transverse-momentum dependent Drell-Yan cross sections was developed by Collins, Soper and Sterman (CSS) [16], and made possible a consistent inclusion of both perturbative and nonperturbative contributions in the same formula. This allowed to successfully describe Z-production data measured at Tevatron during the '90s and early 2000s, together with earlier data at low energy [17–20]. Later, the formalism was extended [21] to describe the transverse momentum of single hadrons produced in deep inelastic scattering (a process commonly denoted as Semi Inclusive DIS, or SIDIS), allowing for the description of collider data from DESY [22], as well as fixed target data from both DESY and CERN (see, e.g., [23–25]). On the theoretical side, there has been recently important progress, with the formal proof that CSS cross sections can be interpreted in a factorized form in terms of transverse momentum dependent parton distribution functions (TMDs) [26–28].

Despite these successes, major questions remain open. Present data seem not to allow us to clearly determine the evolution of TMDs with respect to the exchanged momentum Q^2, nor their variation with the longitudinal momentum x, and quark flavor. Even their shape in transverse momentum has not been sharply established: while the works cited above agreed on a gaussian distribution, studies of newer data seem to require different shapes [25, 29, 30]. In this thesis, we present a TMD study in Drell-Yan, including the most recent data, released from LHC. To date, the only nonperturbative studies making use of these data are [29, 30], where however theoretical predictions for some experiments are renormalized to

---

[3] Drell-Yan is the production of a lepton-antilepton pair in the collision of two hadrons, via the electroweak annihilation of a quark and antiquark [12]. Although the term is commonly referred to the case where the intermediate state is a virtual photon, in the following we will extend it to Z-boson production.

[4] q_T here denotes the transverse momentum of the lepton pair.

the integral of the experimental cross section. At the same time, in studies like [31] a fair description of LHC data is obtained with the sole use of perturbative resummation. At variance with [29, 30], in this work we compare to LHC data without the use of normalization factors, thus making a study of both the shape and absolute value of cross sections. This more direct comparison puts in evidence the problem posed by modern data, especially those from ATLAS: since their uncertainties reach the few ‰ level, it becomes contradictory trying to fit them without taking into account also the uncertainties of theoretical formulas, which are often larger. For this reason, we introduce collinear PDF uncertainties in the calculation of the χ^2.

For what concerns Semi Inclusive Deep Inelastic Scattering, will show for the first time a phenomenological study of TMDs beyond zeroth order in pQCD, highlighting the crucial effects of first order corrections: these are such to apparently prevent any description of data beyond leading order. We provide a theoretical explanation for this discrepancy, essentially based on an analysis of what is in principle the kinematic region for the proper application of the TMD formalism. The argument is supported by a study of the most recent data available.

TMDs are the theoretical tool for describing low values of transverse momentum. A complete description of data requires the transition to ordinary ("collinear") pQCD at higher q_T, where the observed transverse momentum is generated by hard gluon emission. The understanding of the whole q_T spectrum of data is essential for a solid determination of the respective importance of perturbative and nonperturbative physics. However, the matching of the TMD and collinear formalisms is known to present problems in practical applications [32]. What is more, an analysis of SIDIS data in the framework of collinear pQCD has recently shown serious discrepancies [33], which implies that we do not have presently a good understanding of high-q_T data in the low energy regimes of SIDIS experiments. An analogous analysis for collinear predictions in low-energy Drell-Yan is addressed in this thesis, finding a similar issue in the description of data. The impact of possible perturbative and non perturbative corrections to the collinear-factorized formulas is analyzed. The size of the discrepancies and the similarities in the kinematics seem to point at a common origin of the problems in SIDIS and Drell-Yan, and it is crucial to understand them before attempting a description of the whole q_T spectrum.

Transverse-momentum dependent cross sections acquire a particular interest when some of the particles involved are polarized. Indeed, asymmetries that arise as a consequence of polarization can provide precious information on nonperturbative dynamics. One famous example are the asymmetries generated by the Sivers TMD [34, 35], which are due to the correlation between the spin of a transversely polarized nucleon and the momentum of the parton. The Sivers effect would not exist without orbital motion of partons, and in model studies it can even be connected to the contribution of orbital angular momentum to the proton spin [36]. In this thesis we will deal with different but still very interesting asymmetries, i.e., those arising in SIDIS when the target nucleon is polarized in the longitudinal direction. These were the first single-spin asymmetries measured in SIDIS [37–39], and have been studied in the TMD framework, although their description is not unambigous [40–44] . However, a perturbative calculation for the matching at higher q_T was lacking. Therefore, we provide 1-loop results in collinear pQCD and phenomenological predictions for a future Electron-Ion Collider, as well as a comparison to the above data. The study of single-spin asymmetries in pQCD is also interesting by itself, as they are "T-odd," i.e., odd under naive time reversal, which corresponds to the application of time reversal without interchanging initial and final states in the scattering process. Perturbative T-odd effects are pure loop effects, as they arise from the imaginary parts of absorptive amplitudes. As such they have been long searched for in the past, but a clear comparison to data has never been possible due to the small size of these phenomena.

The outline of the thesis is as follows: in Ch. 2 we introduce the kinematics of Semi Inclusive Deep Inelastic Scattering and Drell-Yan, and the general decomposition of cross sections in terms of hadronic structure functions. In Ch. 3 we summarize the theory of Transverse Momentum Distributions, presenting the framework used in Ch. 4, where we report our phenomenological studies on TMDs in SIDIS and Drell-Yan. In Ch. 5 we show the analysis of collinear factorization in Drell-Yan at low energy at high transverse momentum In Ch. 6 we present the results on spin asymmetries in SIDIS, providing our analytic calculation for the involved structure functions, and numerical predictions for an EIC. We also show a comparison to available HERMES data, to study the role of collinear-factorized theory at low energies.

Chapter 2

Transverse-Momentum-Dependent cross sections: Semi-Inclusive Deep Inelastic Scattering and Drell-Yan

In this Chapter, we summarize some of the most relevant aspects in the theoretical description of Drell-Yan and Semi-Inclusive Deep Inelastic Scattering: after describing the kinematics of the processes, we review their factorization properties, in particular for cross sections that are differential with respect to the transverse momentum.

2.1 Deep Inelastic Scattering

In Deep Inelastic Scattering (DIS) a lepton scatters off a nucleon, the nucleon is destroyed, and, in the most inclusive case, only the outgoing lepton is observed:

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

(particles momenta are indicated in parentheses, and N denotes a generic nucleon). At leading order in the elctroweak coupling, the interaction is described by the exchange of a single virtual photon or weak boson (although for simplicity the formulae shown throughout the Chapter are for the case of virtual photon).

2.1.1 Kinematics

In this thesis, we will consider a frame of reference where the boson momentum q^μ is directed along the $+\hat{z}$ direction and the nucleon momentum P^μ along $-\hat{z}$. This is opposite to the most common convention in the literature on inclusive DIS, but is in accordance with the usual convention for Semi-Inclusive DIS. Moreover, if not specified otherwise, we will use the so-called infinite momentum frame, i.e., the frame where the boson has only spatial components. In light-cone coordinates:

$$\begin{aligned} P^\mu &= \left[\frac{Q}{x_B\sqrt{2}}, \frac{x_B M^2}{Q\sqrt{2}}, \mathbf{0}\right], \\ q^\mu &= \left[-\frac{Q}{\sqrt{2}}, \frac{Q}{\sqrt{2}}, \mathbf{0}\right]. \end{aligned} \tag{2.2}$$

In our calculations and in the following sections we will neglect hadron masses,[1] so that the nucleon momentum will be light-like. The cross section is expressed in terms of the following invariants:

- the modulus of the boson virtuality

  $$Q^2 = -q^2, \tag{2.3}$$

  which is required to be large, $Q^2 \gg \Lambda^2_{QCD}$, for the applicability of asymptotic freedom and perturbative QCD;

- the squared center-of-mass energy of the lepton and the nucleon

  $$s = (l + P)^2 \simeq 2P \cdot l; \tag{2.4}$$

- the invariant mass of the hadronic final state after interaction with the lepton

  $$W^2 = (q + P)^2; \tag{2.5}$$

- the Bjorken variable x_B , which in the parton model corresponds to the fraction of momentum of the struck quark with respect to the nucleon,

  $$x_B = \frac{Q^2}{2P \cdot q}; \tag{2.6}$$

- the“inelasticity” y of the process, which in the nucleon rest frame corresponds to the fraction of energy exchanged by the lepton,

  $$y = \frac{P \cdot q}{P \cdot l} = \frac{Q^2}{2x_B s}. \tag{2.7}$$

---

[1] See, for instance, [45, 46] for studies of hadron mass corrections.

2.1.2 Hadronic matrix elements and Parton Distribution Functions

The cross section of process (2.1) can be expressed as

$$\mathrm{d}\sigma = \frac{1}{2s}\left|\bar{\mathcal{M}}\right|^2 \mathrm{d(Lips)}, \tag{2.8}$$

where $2s$ is the flux factor, d(Lips) is the Lorentz-invariant phase space

$$\mathrm{d(Lips)} = (2\pi)^4\delta^4\left(l + P - l' - P_X\right)\frac{\mathrm{d}^3P_X}{(2\pi)^3 2P_X^0}\frac{\mathrm{d}^3l'}{(2\pi)^3 2l'^0}, \tag{2.9}$$

while $\mathcal{M}$ is the Feynman amplitude for lepton-nucleon inelastic scattering

$$i\mathcal{M} = \bar{u}\left(l', \lambda'\right)\gamma^\mu u(l, \lambda)\frac{e^2}{Q^2}\left\langle P_X, S_X\left|J_\mu(0)\right|P, S\right\rangle. \tag{2.10}$$

For simplicity, we will now consider the case of unpolarized scattering, and average (sum) over the spin of initial (final) state particles:

$$\left|\bar{\mathcal{M}}\right|^2 = \frac{1}{4}\sum_{\lambda,\lambda',S,S_X}|\mathcal{M}|^2. \tag{2.11}$$

The matrix element in (2.11) is commonly written as

$$\left|\bar{\mathcal{M}}\right|^2\frac{\mathrm{d}^3P_X}{(2\pi)^3 2P_X^0} = \frac{e^4}{Q^4}L_{\mu\nu}W^{\mu\nu}, \tag{2.12}$$

where $L_{\mu\nu}$ is the (spin averaged) lepton tensor, which only depends on l, l':

$$\begin{aligned}L_{\mu\nu}\left(l, l'\right) &= \frac{1}{2}\sum_{\lambda,\lambda'}\left(\bar{u}\left(l', \lambda'\right)\right)\gamma_\mu u\left(l, \lambda\right))^*\left(\bar{u}\left(l', \lambda'\right)\right)\gamma_\nu u\left(l, \lambda\right))\\&= -Q^2 g_{\mu\nu} + 2\left(l_\mu l'_\nu + l'_\mu l_\nu\right),\end{aligned} \tag{2.13}$$

and $W^{\mu\nu}$ is the hadron tensor, which is expressed in terms of hadronic matrix elements as[2]

$$\begin{aligned}W^{\mu\nu}(q, P) =& \frac{1}{2\pi}\frac{1}{2}\sum_S\sum_X\int\frac{\mathrm{d}^3P_X}{(2\pi)^3P_X^0}(2\pi)^4\delta^{(4)}\left(q + P - P_X\right)\\&\left\langle P, S\left|J^\mu(0)\right|P_X\right\rangle\left\langle P_X\left|J^\nu(0)\right|P, S\right\rangle.\end{aligned} \tag{2.14}$$

---

[2] Summation over spin in the final state is intended.

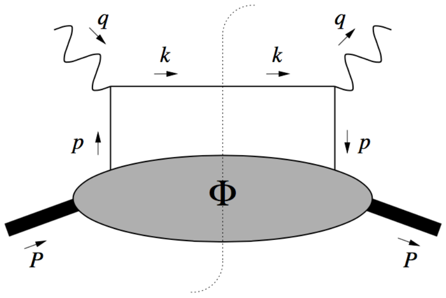


Figure 2.1: The Feynman amplitude and its complex conjugate (cut diagram) for DIS in the parton model.

The "deep inelastic" regime is defined by Q^2, $P \cdot q \to \infty$, with x_B kept fixed. Considering the infinite momentum frame, relativistic time dilation and length contraction in this regime are such that the interaction takes place on a much shorter time scale than the one of the (nonperturbative) inner dynamics of the nucleon, and the scattering involves a single quark, which can be effectively considered free during the interaction. This means that the final state X in (2.14) is given by an on-shell quark with momentum $k^\mu = p^\mu + q^\mu$ and a remainder R (Fig. 2.1), so that the hadron tensor can be rewritten as

$$\begin{aligned} W^{\mu\nu}(q,P) = &\sum_q e_q^2 \frac{1}{2} \sum_S \sum_R \int \frac{\mathrm{d}^3 \boldsymbol{P}_R}{(2\pi)^3 2P_R^0} \int \mathrm{d}^4 p\, \delta\left((p+q)^2 - m^2\right) \\ &\times \theta\left(p^0 + q^0 - m\right) \delta^{(4)}\left(P - p - P_R\right) \langle P,S|\bar{\psi}_I^q(0)|R\rangle\langle R|\psi_J^q(0)|P,S\rangle \\ &\times \gamma^\mu_{IK}(\not{p} + \not{q} + m)_{KL}\gamma^\nu_{LJ}, \end{aligned} \tag{2.15}$$

where ψ_I^q is the Dirac component I of a quark field with flavor q, and e_q is the fractional charge of the quark.

It is convenient to Fourier transform the four-momentum conservation delta function in the second line of (2.15) and use it to perform a translation of the quark field:

$$\begin{aligned} W^{\mu\nu}(q,P) = &\sum_q e_q^2 \frac{1}{2} \sum_S \int \mathrm{d}^4 p\, \delta\left((p+q)^2 - m^2\right) \theta\left(p^0 + q^0 - m\right) \\ &\times \int \frac{\mathrm{d}^4 \xi}{(2\pi)^4} \mathrm{e}^{-ip\cdot\xi} \langle P,S|\bar{\psi}_I^q(\xi)\psi_J^q(0)|P,S\rangle \\ &\times \gamma^\mu_{IK}(\not{p} + \not{q} + m)_{KL}\gamma^\nu_{LJ}, \end{aligned} \tag{2.16}$$

where we made use of the completeness of the remainder states $|R\rangle$. We thus have an expression for the hadron tensor in terms of a quark-quark correlator

$$\Phi(p, P, S) = \int \frac{\mathrm{d}^4\xi}{(2\pi)^4} \mathrm{e}^{-ip\cdot\xi} \langle P, S|\bar{\psi}_I^q(\xi)\psi_J^q(0)|P, S\rangle. \tag{2.17}$$

Interesting results can already be obtained by performing an expansion of (2.16) in powers of $1/Q$, and retaining only the leading terms. In this way one selects the leading part of the quark quark correlator, obtaining the so-called leading twist Parton Distribution Functions (PDFs), that are the most well-known among the PDFs, and have a probabilistic interpretation. To apply these approximations, it is necessary to consider what is the behavior of the components of particles momenta. Using light-cone coordinates, the initial quark momentum can be parametrized as

$$p^\mu = \left[xP^+, \frac{p^2 + |\boldsymbol{p}_T|^2}{2xP^+}, \boldsymbol{p}_T \right], \tag{2.18}$$

where the quark virtuality p^2 and its transverse momentum $|\boldsymbol{p}_T|^2$ are assumed to be small compared to the hard scale Q^2. Without showing all the details, we only notice that the leading component of $p^\mu + q^\mu$ is the plus component, so that in third line of (2.16) we can retain just the term $\gamma^\mu\gamma^-\gamma^\nu$, while the delta function in the first line becomes proportional to $\delta(x - x_B)$, which eliminates the integral over the p^- component. In this way, the final result becomes

$$W^{\mu\nu}(q, P) \simeq \sum_q e_q^2 \frac{1}{2} \operatorname{Tr}\left[\Phi^q\left(x_B\right)\gamma^\mu\gamma^-\gamma^\nu\right], \tag{2.19}$$

where we have introduced the integrated (and spin averaged) correlation function

$$\begin{aligned} \Phi_{JI}(x) &= \frac{1}{2}\sum_S \int \mathrm{d}^2\boldsymbol{p}_T \mathrm{d}p^+ \Phi_{JI}\left(p, P, S\right)\big|_{p^-=xP^-} \\ &= \int \frac{\mathrm{d}\xi^+}{2\pi} \mathrm{e}^{-\mathrm{i}p\cdot\xi} \left\langle P, S\left|\overline{\psi}_I(\xi)\psi_J(0)\right| P, S\right\rangle\Big|_{\xi^-=\xi_T=0} \end{aligned} . \tag{2.20}$$

From the above expression, recalling the definition of the (collinear) parton distribution function f_1,

$$f_1(x) \propto \frac{1}{2}\sum_S \int_{-\infty}^{+\infty} \frac{\mathrm{d}\xi_+}{2\pi} e^{-ixP_-\xi_+} \left\langle P, S\left|\overline{\psi}\left(\xi_+\right)\gamma_-\psi(0)\right| P, S\right\rangle, \tag{2.21}$$

it is possible to show that the DIS cross section has the form

$$\frac{\mathrm{d}\sigma}{\mathrm{d}x_B\mathrm{d}Q^2}(P, q) = \int_0^1 \frac{\mathrm{d}x}{x}\frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}x\mathrm{d}Q^2}(p, q) f_1\left(\frac{x_B}{x}\right), \tag{2.22}$$

where $\mathrm{d}\hat{\sigma}(p,q)$ is the cross section for the scattering off a quark with momentum p.

Up to this point, we considered the scattering at zeroth order in QCD. It can be shown, through a so-called factorization proof, that the form in (2.22) for the cross section also holds in full QCD. Indeed, the partonic cross section $\mathrm{d}\hat{\sigma}$ can be calculated from Feynman diagrams up to the desired order in α_s. Infrared divergences that are not canceled in the sum of virtual and real contributions can be factorized out of the partonic cross section through the introduction of a factorization scale μ_F. In rough terms, only particles with a virtuality greater than μ_F^2 will appear in the partonic scattering amplitude, all the others being considered as contributing to the nonperturbative dynamics embodied by the PDFs. On the other hand, ultraviolet divergences will require the introduction of a renormalization scale μ. Setting for simplicity $\mu_F = \mu$, we can write the generalization of (2.22) as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}x_B \mathrm{d}Q^2}(x_B, Q^2) = \int_0^1 \frac{\mathrm{d}x}{x} \frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}x \mathrm{d}Q^2}(x, Q^2/\mu^2, \alpha_s(\mu^2)) f_1\left(\frac{x_B}{x}, \mu^2\right), \tag{2.23}$$

where we have re-expressed the kinematic dependence in terms of x_B and Q^2. The PDF now depends on the scale μ, and its evolution can be calculated in perturbative QCD, from the Renormalization Group Equation (RGE).[3] Although there exist processes for which factorization is known to be violated [48, 49], all the observables we will consider in this thesis can be factorized as a convolution of short-distance partonic scattering and Parton Distribution Functions.

A remark should be made concerning QCD corrections and the quark-quark correlator in (2.17) to (2.21): as it stands, the correlator is not gauge invariant, as can be seen by applying a local gauge transformation to the quark fields. However this is just an artifact of adopting a pure parton model picture. When including QCD corrections, final state (or initial state, in the case of Drell-Yan) gluon radiation must be considered. This radiation factorizes, i.e., it can be included inside the definition of Parton Distribution Functions, and contributions from any number of gluons sum up to give a path-ordered exponential of integrals of gluon fields. This provides precisely the gauge link (or Wilson line) $\mathcal{L}$ necessary to connect the two quark fields at different space-time positions ξ_1, ξ_2 in the correlator, thus ensuring gauge invariance:[4]

$$\mathcal{L}[\xi_1, \xi_2] = \mathcal{P} \exp\left(ig \int_{\xi_1}^{\xi_2} dz_\mu A^\mu(z) \right). \tag{2.24}$$

In the case of collinear PDFs, (2.21), the space-time separation of the fields is along a light-cone direction, and the gauge link is set to 1 by choosing a proper light-cone gauge, $n \cdot A = 0$.

---

[3] See, e.g., [47]

[4] A detailed treatment can be found in [26]

2.2 Semi-Inclusive Deep Inelastic Scattering

Particularly interesting observables in DIS can be obtained when including information on a hadron h detected in the final state:

$$\ell(l) + N(P) \to \ell(l') + h(P_h) + X. \tag{2.25}$$

This is denoted as Semi-Inclusive Deep Inelastic Scattering (SIDIS). We will assume that the final hadron is generated by the hadronization of a quark or gluon coming from the hard partonic scattering involving the virtual photon (current fragmentation), and not from the target debris (target fragmentation). The kinematics for SIDIS is shown in Fig. 2.2. The lepton momenta define the $\hat{x}$-$\hat{z}$ plane, with respect to which we define the azimuthal angle of the outgoing hadron, ϕ_h, and, in the case of a transversely polarized target, the angle of the trasnverse spin ϕ_S. $P_{h\perp}$ is defined as the component of the outgoing hadron momentum transverse to the $\hat{z}$ axis.[5] The fragmentation variable z is instead defined in relation to the longitudinal part of P_h^μ:

$$z = \frac{P \cdot P_h}{P \cdot q}. \tag{2.26}$$

In the parton model, z corresponds to the fractional momentum of the detected hadron with respect to the quark which originated the fragmentation process.

The most simple observable that can be devised in SIDIS is obtained by integrating the events over the azimuthal angle ϕ and the transverse momentum $P_{h\perp}$, so that the cross section will be differential in x, Q^2 and z (we consider for the moment unpolarized scattering). To reduce systematic uncertainties, it is convenient to normalize by the total number of DIS events with the given x and Q^2, obtaining what we will refer to as *integrated multiplicity* (in the following, we will rewrite the Bjorken variable x_B as x):

$$M^{\mathrm{h}}\left(x, z, Q^2\right) = \frac{\mathrm{d}\sigma^{\mathrm{h}}/\left(\mathrm{d}x\mathrm{d}Q^2\mathrm{d}z\right)}{\mathrm{d}\sigma^{\mathrm{DIS}}/\left(\mathrm{d}x\mathrm{d}Q^2\right)}. \tag{2.27}$$

This observable is an important source of data for the extraction of Fragmentation Functions,

$$D_1^{q/h}\left(z, Q^2\right), \tag{2.28}$$

which give the probability for the quark of flavor q to fragment into a hadron h carrying fractional momentum z. Indeed, the factorized formula for the numerator

---

[5] In the following chapters we will use also the notation P_{hT}, when referring to the modulus of the transverse components of P_h^μ.

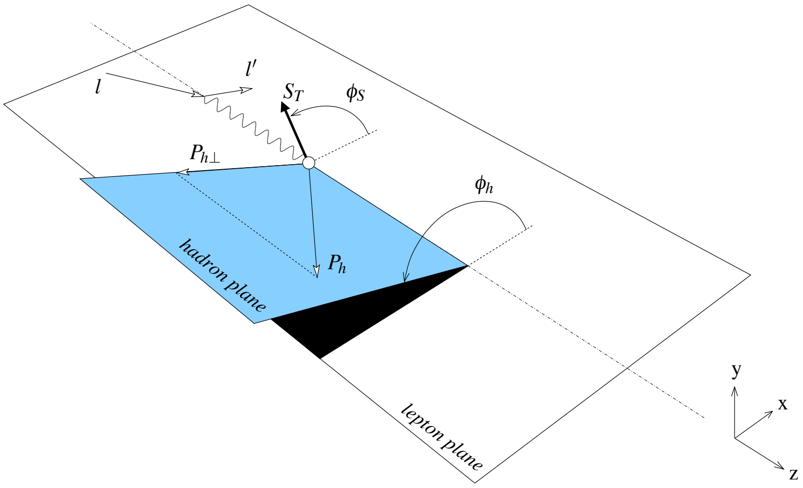


Figure 2.2: The kinematics for Semi-Inclusive Deep Inelastic Scattering.

of (2.27) reads

$$\begin{aligned}\frac{\mathrm{d}\sigma^h}{\mathrm{d}x\mathrm{d}Q^2\mathrm{d}z} = \sum_{a,b}\int &\mathrm{d}\hat{x}\mathrm{d}\hat{z}\mathrm{d}\xi\mathrm{d}\eta\delta(x-\xi\hat{x})\delta(z-\eta\hat{z}) \\ &\times\frac{\mathrm{d}\hat{\sigma}_{ab}}{\mathrm{d}\hat{x}\mathrm{d}Q^2\mathrm{d}\hat{z}} f_1^a\left(\xi,Q^2\right) D_1^{b/h}\left(\eta,Q^2\right),\end{aligned} \tag{2.29}$$

where as before $\mathrm{d}\hat{\sigma}$ is the partonic cross section calculated perturbatively.

2.2.1 Structure functions in SIDIS

We will now consider the fully differential cross section in SIDIS. Also in this case, the cross section is expressed in terms of lepton and hadron tensor

$$\frac{\mathrm{d}\sigma^{\mathrm{h}}}{\mathrm{d}x\mathrm{d}Q^2\mathrm{d}z\mathrm{d}P_{h\perp}^2\mathrm{d}\phi\mathrm{d}\phi_S} \propto L^{\mu\nu}W_{\mu\nu}, \tag{2.30}$$

where we allow for lepton polarization:

$$L_{\mu\nu}(l,l',\lambda_e) = -Q^2 g_{\mu\nu} + 2\left(l_\mu l'_\nu + l'_\mu l_\nu\right) + 2i\lambda_e\epsilon_{\mu\nu\rho\sigma}l^\rho l'^\sigma, \tag{2.31}$$

with $\lambda_e = +(-)1$ corresponding to a right (left) handed letpon beam. The hadron tensor will now include information on the detected hadron:

$$\begin{aligned} W^{\mu\nu}(q, P, P_h, S) = \sum_X \int \frac{d^3\boldsymbol{P}_X}{(2\pi)^3 2P_X^0} \delta^{(4)}(q + P - P_X - P_h) \\ \times \langle P, S | J^\mu(0) | P_h, X \rangle \langle P_h, X | J^\nu(0) | P, S \rangle \, . \end{aligned} \tag{2.32}$$

To understand the structure of the ϕ and y dependence of the cross section, it is useful to write it in terms of boson-nucleon scattering amplitudes. The longitudinal and transverse boson polarization states can be defined respectively as

$$\begin{aligned} \epsilon_0^\mu &= \frac{1}{Q}\left(q^\mu + \frac{Q^2}{P \cdot q} P^\mu\right), \\ \epsilon_{+1}^\mu &= \frac{1}{\sqrt{2}}(0, 1, +i, 0), \quad \epsilon_{-1}^\mu = \frac{1}{\sqrt{2}}(0, 1, -i, 0). \end{aligned} \tag{2.33}$$

Let us write

$$L^{\mu\nu} W_{\mu\nu} = j^\mu (j^\nu)^* W_{\mu\nu}, \tag{2.34}$$

j^μ being the lepton current. We can then expand this current onto the helicity basis in (2.33):

$$j^\mu = \frac{1}{\sqrt{2}}\left((1-y)^{+\frac{1}{2}\lambda_e} \epsilon_{-1}^\mu + (1-y)^{-\frac{1}{2}\lambda_e} \epsilon_{+1}^\mu\right) + \epsilon_0^\mu . \tag{2.35}$$

Considering first the case of unpolarized target, we can derive the azimuthal dependence of the cross section by taking the hadron plane coincident with the lepton plane, and then rotating the lepton system by an angle $-\phi$ around the $\hat{z}$ axis, which is equivalent to rotating the hadron system by an angle ϕ. The rotated lepton current will be

$$j^\mu(-\phi) = \frac{1}{\sqrt{2}}\left((1-y)^{+\frac{1}{2}\lambda_e} \mathrm{e}^{-i\phi} \epsilon_{-1}^\mu + (1-y)^{-\frac{1}{2}\lambda_e} \mathrm{e}^{i\phi} \epsilon_{+1}^\mu\right) + \epsilon_0^\mu . \tag{2.36}$$

Inserting this into (2.34), we obtain

$$\begin{aligned}
(1-y)L^{\mu\nu}W_{\mu\nu} = {} & \frac{1}{2}\left[1+(1-y)^2\right](W_{--}+W_{++}) \\
& -\lambda_e\frac{1}{2}\left[1-(1-y)^2\right]\left(\underset{\cdots\cdots\cdots}{W_{--}-W_{++}}\right)+2(1-y)W_{00} \\
& +2(1-y)\left[\operatorname{Re}W_{-+}\cos 2\phi-\underline{\underset{\cdots\cdots}{\operatorname{Im}W_{-+}}}\sin 2\phi\right] \\
& +\sqrt{2}(1-y)^{1/2}(2-y)\left[\operatorname{Re}(W_{-0}+W_{+0})\cos\phi+\right. \\
& \left.-\underline{\underset{\cdots\cdots\cdots\cdots\cdots}{\operatorname{Im}(W_{-0}-W_{+0})}}\sin\phi\right]-\lambda_e\sqrt{2}y(1-y)^{1/2} \\
& \times\left[\underset{\cdots\cdots\cdots\cdots\cdots}{\operatorname{Re}(W_{-0}-W_{+0})}\cos\phi-\underline{\operatorname{Im}(W_{-0}+W_{+0})}\sin\phi\right],
\end{aligned} \tag{2.37}$$

where

$$W_{ab}=\epsilon_a^{\mu}\left(\epsilon_b^{\nu}\right)^* W_{\mu\nu}, \qquad a,b=+,-,0. \tag{2.38}$$

We see that modulations in the azimuthal angle of the produced hadron are due to the quantum interference of different polarization states. Terms underlined with dots are P-odd and hence only present in the case of weak interactions. Terms underlined with a continuous line are instead odd under naive time-reversal (T-odd), which means time reversal without the interchange of initial and final states, i.e., reversal of the spatial components of momenta and spins. $W_{\mu\nu}$ is intended here to be the spin averaged version of the hadron tensor. The following notation is commonly used for the transverse and longitudinal parts:

$$W_T=\frac{1}{2}(W_{++}+W_{--}), \qquad W_L=W_{00}. \tag{2.39}$$

If we allow for nucleon polarization, defining the target spin three-vector in our reference frame as

$$\boldsymbol{S}=\begin{pmatrix} S_T\cos\phi_S \\ S_T\sin\phi_S \\ -S_L \end{pmatrix}, \tag{2.40}$$

then the following terms have to be added to (2.37):

$$
\begin{aligned}
&- S_L \left[2(1-y)\sin(2\phi)\underline{\mathrm{Im}\, W^{++}_{+-}} + \sqrt{2}(1-y)^{1/2}(2-y)\sin\phi\, \underline{\mathrm{Im}\left(W^{++}_{+0} - W^{--}_{+0}\right)}\right] \\
&+ S_L \lambda_e \left[y\,(2-y)\,\frac{1}{2}\left(W^{++}_{++} - W^{--}_{++}\right) - \sqrt{2}y(1-y)^{1/2}\cos\phi\, \mathrm{Re}\left(W^{++}_{+0} - W^{--}_{+0}\right)\right] \\
&- S_T \left[\sin(\phi - \phi_S)\, \underline{\mathrm{Im}\left(W^{+-}_{++} + (2-y)W^{+-}_{00}\right)} + \frac{(2-y)}{2}\sin(\phi + \phi_S)\, \underline{\mathrm{Im}\, W^{+-}_{+-}}\right. \\
&\qquad + \frac{(2-y)}{2}\sin(3\phi - \phi_S)\, \underline{\mathrm{Im}\, W^{-+}_{+-}} + \sqrt{2}(1-y)^{1/2}(2-y)\sin\phi_S \underline{\mathrm{Im}\, W^{+-}_{+0}} \\
&\qquad \left. + \sqrt{2}(1-y)^{1/2}(2-y)\sin(2\phi - \phi_S)\, \underline{\mathrm{Im}\, W^{-+}_{+0}}\right] \\
&+ S_T \lambda_e \left[y\,(2-y)\cos(\phi - \phi_S)\, \mathrm{Re}\, W^{+-}_{++}\right. \\
&- \sqrt{2}y(1-y)^{1/2}\cos\phi_S\, \mathrm{Re}\, W^{+-}_{+0} - \sqrt{2}y(1-y)^{1/2}\cos(2\phi - \phi_S)\, \mathrm{Re}\, W^{-+}_{+0}\left.\right],
\end{aligned}
\tag{2.41}
$$

where the upper ± labels in the hadron tensor have been introduced to denote the helicity states of the nucleon in the amplitude and its complex conjugate, and we have now excluded for brevity P-violating terms. A different notation can be found in the literature ([50]), where the structure functions are labeled by the polarization of the external particles and the type of the azimuthal modulation. For instance, one has

$$
\begin{aligned}
F_{UU,T} &= \frac{1}{2}\left(F^{++}_{++} + F^{--}_{++}\right), \qquad F_{UU,L} = F^{++}_{00} \\
F^{\cos\phi_h}_{UU} &= -\frac{1}{\sqrt{2}}\,\mathrm{Re}\left(F^{++}_{+0} + F^{--}_{+0}\right), \\
F^{\sin(\phi_h - \phi_S)}_{UT,T} &= -\,\mathrm{Im}\, F^{+-}_{++},
\end{aligned}
\tag{2.42}
$$

where the letters in the subscript represent, respectively, the lepton, target, and boson polarization (U = unpolarized, T = transverse, L = longitudinal), so that, e.g., $F^{\sin(\phi_h - \phi_S)}_{UT,T}$ is the structure function multiplying $\sin(\phi_h - \phi_S)$, when considering an unpolarized lepton, transversely polarized target, and averaging over the transverse polarization states of the boson. The functions F^{ij}_{mn} are closely related to the hadronic amplitudes W^{ij}_{mn} above. They are defined as

$$
F^{ij}_{mn}\left(x, Q^2, z, P^2_{h\perp}\right) = \frac{Q^2(1-x)}{4\pi^3\alpha}\,\frac{d\sigma^{ij}_{mn}}{dz dP^2_{h\perp}},
\tag{2.43}
$$

$d\sigma^{ij}_{mn}$ being boson-nucleon scattering amplitudes

In order to give accurate QCD predictions for the SIDIS structure functions, it is crucial to distinguish between a region of high transverse momentum, where collinear factorization is applicable, and the TMD factorization region at low transverse momentum. To this end, it is necessary to introduce the invariant[6]

$$q_T^2 = -q_T^\mu q_{T\mu}, \tag{2.44}$$

where q_T^μ is the transverse component of the boson momentum with respect to both the nucleon and the final hadron:

$$q_T^\mu = q^\mu + \xi P^\mu - \frac{1}{z} P_h^\mu, \tag{2.45}$$

with

$$\xi \equiv -\frac{P_h \cdot q}{P \cdot P_h}. \tag{2.46}$$

In the frame we are considering, the Minkowski components of P_h can be expressed in terms of q_T as

$$P_h^\mu = \frac{zQ}{2}\left(1 - \frac{q_T^2}{Q^2}, \frac{2q_T}{Q}\cos\phi, \frac{2q_T}{Q}\sin\phi, 1 + \frac{q_T^2}{Q^2}\right). \tag{2.47}$$

Hence we see that there is a simple relation between the transverse momentum of the detected hadron in our frame and that of the boson w.r.t. the hadrons:

$$P_{h\perp} \equiv \sqrt{P_{hx}^2 + P_{xy}^2} = z q_T. \tag{2.48}$$

In a covariant way, the transverse momentum of the final hadron is expressed as

$$P_{h\perp}^\mu = P_h^\mu - z\left(2x - \xi\right) P^\mu - z q^\mu. \tag{2.49}$$

The relevance of the variable q_T is due to the fact that it is directly related to the kinematics of the partonic process: relatively high values of q_T originate from the radiation of hard gluons, while a low q_T calculation requires a resummation of the logarithmically enhanced soft gluon radiation and the inclusion of a model for intrinsic transverse momentum. In the following sections we will describe the two regimes. The discussion can be generalized to other processes, like, e.g., Drell-Yan and $e^+e^- \longrightarrow h_1 h_2 X$.

2.2.2 High transverse momentum: Collinear factorization

When q_T is high compared to hadronic scales, $q_T \gg M$, possible effects of intrinsic transverse momentum can be neglected, and the observed transverse

---

[6] Where not specified otherwise, we will use a_T to denote the modulus of the transverse component of a^μ: $\mathrm{a}_T \equiv |\mathbf{a}_T|$

momentum can be ascribed entirely to the recoil of the scattered partons against hard QCD radiation, which can be calculated perturbatively. In this case, the collinear factorization formula applies, and the cross section is a convolution of a partonic cross section $\mathrm{d}\hat{\sigma}_{ab}$ and collinear PDFs and FFs:

$$\begin{aligned}\frac{\mathrm{d}\sigma}{\mathrm{d}x\mathrm{d}Q^2\mathrm{d}z\mathrm{d}q_T\mathrm{d}\phi} = {} & \sum_{a,b}\int \mathrm{d}\hat{x}\mathrm{d}\hat{z}\mathrm{d}\xi\mathrm{d}\eta\,\delta(x-\xi\hat{x})\delta(z-\eta\hat{z}) \\ & \times \frac{\mathrm{d}\hat{\sigma}_{ab}}{\mathrm{d}\hat{x}\mathrm{d}Q^2\mathrm{d}\hat{z}\mathrm{d}q_T\mathrm{d}\phi} f_1^a\left(\xi, Q^2\right) D_1^{b/h}\left(\eta, Q^2\right).\end{aligned} \tag{2.50}$$

The partonic momenta are taken to be collinear to the external hadrons: $p_a^\mu = \xi P^\mu$, $p_b^\mu = \frac{1}{\eta}P_h^\mu$. Considering the leading order calculation, the partonic process is of the form $\gamma^* + q(\bar{q}) \longrightarrow q(\bar{q}) + g$ or $\gamma^* + g \longrightarrow q + \bar{q}$, and the structure functions have the following form:

$$\begin{aligned}F_{UU,T} = {} & \frac{1}{Q^2}\frac{\alpha_s}{(2\pi z)^2}\sum_a x e_a^2 \int_x^1 \frac{d\hat{x}}{\hat{x}} \int_z^1 \frac{d\hat{z}}{\hat{z}} \delta\left(\frac{q_T^2}{Q^2} - \frac{(1-\hat{x})(1-\hat{z})}{\hat{x}\hat{z}}\right) \\ & \times f_1^a\left(\frac{x}{\hat{x}}\right) D_1^b\left(\frac{z}{\hat{z}}\right) C_{UU,T}^{ab}\left(\hat{x}, \hat{z}\right),\end{aligned} \tag{2.51}$$

where we took $F_{UU,T}$ for the sake of example. The δ-function in (2.51) ensures the on-shellness of the unobserved particle in the (LO) partonic cross section. In the limit of small transverse momentum, $q_T \ll Q$, one can approximate it as (see, e.g., [51, 52])

$$\begin{aligned}\delta\left(\frac{q_T^2}{Q^2} - \frac{(1-\hat{x})(1-\hat{z})}{\hat{x}\hat{z}}\right) = {} & \delta(1-\hat{x})\delta(1-\hat{z})\ln\frac{Q^2}{q_T^2} + \frac{\hat{x}}{(1-\hat{x})_+}\delta(1-\hat{z}) \\ & + \frac{\hat{z}}{(1-\hat{z})_+}\delta(1-\hat{x}) + O\left(\frac{q_T^2}{Q^2}\ln\frac{Q^2}{q_T^2}\right),\end{aligned} \tag{2.52}$$

with this approximation, one can obtain the behavior of the structure function in the small-q_T limit, commonly known as the *asymptotic expansion*. For $F_{UU,T}$, this yields

$$\begin{aligned}F_{UU,T} \simeq {} & \frac{1}{q_T^2}\frac{\alpha_s}{2\pi^2 z^2}\sum_a x e_a^2 \Bigg[f_1^a(x) D_1^a(z) C_F \left(2\ln\frac{Q^2}{q_T^2} - 3\right) + f_1^a(x)\Big(D_1^a \otimes P_{qq} \qquad . \\ & + D_1^g \otimes P_{gq}\Big)(z) + \left(P_{qq}\otimes f_1^a + P_{qg}\otimes f_1^g\right)(x) D_1^a(z)]\end{aligned} \tag{2.53}$$

The structure function in this case is divergent for $q_T \longrightarrow 0$, because of the well-known soft and collinear divergences of QCD. The structure of divergences is such that at perturbative order α_s^N, the asymptotic expression for $F_{UU,T}$ will have the form

$$\frac{1}{q_T^2}\sum_{n=1}^{N}\sum_{m=0}^{2n-1} c_{n,m}(x,z)\left(\frac{\alpha_s}{2\pi}\right)^n \ln^m\left(\frac{Q^2}{q_T^2}\right). \tag{2.54}$$

Then, for $q_T \ll Q$, each power of α_s is multiplied by large logarithmic (and inverse power) terms, which make a fixed order perturbative expansion unreliable in this regime. Different structure functions have a different asymptotic behavior for $q_T \ll Q$. For instance, $F_{UU,L}$ is suppressed by a relative factor q_T^2/Q^2 with respect to $F_{UU,T}$, and hence can be neglected in phenomenological studies at low q_T.

A well-known formalism for the all-order resummation of the logarithmic terms for the unpolarized structure function $F_{UU,T}$ was first developed for Drell-Yan in [16], and extended to SIDIS in [52, 53]. In recognition of the work of the authors in [16], this formalism is commonly referred to as CSS resummation. To ensure transverse momentum conservation, the resummation formula is commonly derived in the Fourier-conjugate space with respect to q_T, and has the form[7] [54]

$$\begin{aligned} F_{UU,T} =& \frac{1}{z^2}\sum_a x e_a^2 H\left(Q^2/\mu^2\right)\int \frac{d^2\boldsymbol{b}}{(2\pi)^2} e^{-i\boldsymbol{b}\cdot\boldsymbol{q}_T}\exp[-S(Q,b)] \\ & \times \sum_i \left(C_{ai}^{\text{in}}\otimes f_1^i\right)(x;\mu=b_0/b)\sum_j\left(D_1^j\otimes C_{ja}^{\text{out}}\right)(z;\mu=b_0/b), \end{aligned} \tag{2.55}$$

where $\otimes$ denotes a convolution over x or z. C_{ai}^{in}, C_{ja}^{out}, $H\left(Q^2/\mu^2\right)$ are perturbative coefficients calculated in fixed-order QCD. In particular C_{ai}^{in}, C_{ja}^{out} are the Wilson coefficients of an Operator Product Expansion (OPE) of quark matrix elements at small transverse distance. As such, they are convoluted with the ordinary (collinear) PDFs and FFs. The large transverse momentum logarithms are exponentiated in the Sudakov factor $\exp[-S(Q,b)]$.

Since the Fourier integral in (2.55) extends over the entire $\boldsymbol{b}$ space, the nonperturbative region $b \gtrsim \frac{1}{\Lambda}$ is also involved, and the numerical integration cannot generally be done without a regularization of the perturbative divergences at large b, often done by the introduction of a cutoff function $b^*(b)$, which freezes the value of b after a certain threshold. This is in agreement with the fact that a complete calculation requires the introduction of a nonperturbative function, such to mimic QCD dynamics at large distances. The observed transverse momentum distribution is indeed generated by a convolution of nonperturbative dynamics and the QCD radiation generated in the scattering. Due to the properties of Fourier transforms, a convolution in momentum space translates into a product of factors in the conjugate space. Hence, nonperturbative contributions are introduced as a

---

[7] The resummation formalism is discussed in more detail in Chapt. 3.

multiplicative term inside the Fourier integral of (2.55):

$$
\begin{aligned}
F_{UU,T} =& \frac{1}{z^2} \sum_a x e_a^2 H\left(Q^2/\mu^2\right) \int \frac{d^2\boldsymbol{b}}{(2\pi)^2} e^{-i\boldsymbol{b}\cdot\boldsymbol{q}_T} \exp[-S(Q, b^*(b))] \\
& \times \exp[-S_{NP}(Q, b, x, z)] \\
& \times \sum_i \left(C_{ai}^{\text{in}} \otimes f_1^i\right)(x; \mu = b_0/b^*(b)) \sum_j \left(D_1^j \otimes C_{ja}^{\text{out}}\right)(z; \mu = b_0/b^*(b)),
\end{aligned}
\tag{2.56}
$$

where $S_{NP}(Q, b, x, z)$ is the nonperturbative function to be fitted to data.

2.2.3 Low transverse momentum: TMD factorization

The function S_{NP} in (2.56) mostly affects the shape of the cross section for $q_T \simeq M$, and is intended to mimic the contribution from intrinsic transverse momentum and nonperturbative soft radiation. A natural question is whether theoretical calculations in the regime $M \lesssim q_T \ll Q$, where (2.56) applies, can be re-expressed in terms of three-dimensional nucleon matrix elements, expressing the intrinsic transverse-momentum distribution of partons. Only in recent years there has been a complete explicit proof of Transverse-Momentum-Dependent (TMD) factorization ([26], see also [27, 28]). It turns out that the TMD formalism reproduces the CSS results: the resummation of Sudakov logarithms is obtained in this framework from the RGE of TMD parton distributions. To have an idea of how Transverse Momentum Distributions enter the cross section, it is instructive to look at the parton model expression for the hadron tensor in SIDIS, (2.32). The outgoing hadron h in this framework is originated from the fragmentation of the quark after the absorption of the virtual photon. The expression for $W^{\mu\nu}$ will then have the form [55]

$$
W^{\mu\nu}(q, P, S, P_h) = \sum_q e_q^2 \int \mathrm{d}^4p\,\mathrm{d}^4k\,\delta^{(4)}(p + q - k)\,\mathrm{Tr}\left[\Phi(p, P, S)\gamma^\mu \Delta(k, P_h)\gamma^\nu\right],
\tag{2.57}
$$

where $\Phi(p, P, S)$ is the quark-quark correlator in its unintegrated form, defined in (2.17), while $\Delta(k, P_h)$ is the correlator describing quark fragmentation

$$
\begin{aligned}
\Delta_{KL}(k, P_h) &= \sum_X \int \frac{\mathrm{d}^3\boldsymbol{P}_X}{(2\pi)^3 2P_X^0} \delta^{(4)}(k - P_h - P_X) \\
&\quad \times \langle 0 |\psi_K(0)| P_h, X\rangle \left\langle P_h, X \middle| \overline{\Psi}_L(0) \middle| 0\right\rangle \\
&= \frac{1}{(2\pi)^4} \int \mathrm{d}^4\xi\, \mathrm{e}^{\mathrm{i}k\cdot\xi} \langle 0 |\psi_K(\xi)| P_h\rangle \left\langle P_h \middle| \overline{\psi}_L(0) \middle| 0\right\rangle .
\end{aligned}
\tag{2.58}
$$

The momentum k^μ of the outgoing quark will be connected to that of the detected hadron P_h^μ. This sets a constraint on k^μ, which we must take into account by inserting in (2.57) a $\delta^{(4)}(p+q-k)$ for conservation of 4-momentum. To understand more clearly the effect of the δ-function, we can temporarily set ourselves in a frame where the external hadrons have no transverse momentum

$$P^\mu = \left[P^-, \frac{M^2}{2P^-}, \mathbf{0}\right], \qquad P_h^\mu = \left[\frac{M_h^2}{2P_h^+}, P_h^+, \mathbf{0}\right], \tag{2.59}$$

$$q^\mu = \left[-xP^-, \frac{P_h^+}{z}, \boldsymbol{q}_T\right]. \tag{2.60}$$

As mentioned before, there is a connection between the transverse momentum of the virtual photon in this frame, and that of the detected hadron in the frame used in the rest of this thesis. More precisely: $\boldsymbol{q}_T = -\boldsymbol{P}_{h\perp}/z$. The transverse momentum part of the δ function is then $\delta^{(2)}\left(\boldsymbol{p}_T + \boldsymbol{q}_T - \boldsymbol{k}_T\right)$, or equivalently $\delta^{(2)}\left(\boldsymbol{p}_T - \boldsymbol{P}_{h\perp}/z - \boldsymbol{k}_T\right)$. We have seen that in the case of inclusive DIS, Sec. 2.1.2, when considering the leading terms in $1/Q$, the expression for the hadron tensor simplifies, and only the correlator integrated over transverse momentum is involved. However, this is not the case for Semi-Inclusive DIS, when the transverse momentum $\boldsymbol{q}_T$ is not integrated over, and it is of the order of hadronic scales. The correct expression is indeed

$$\begin{aligned} W^{\mu\nu}\left(q, P, S, P_h\right) = \sum_q e_q^2 4z \int \mathrm{d}^2\boldsymbol{p}_T \mathrm{d}^2\boldsymbol{k}_T \boldsymbol{\delta}^{(2)}\left(\boldsymbol{p}_T - \frac{\boldsymbol{P}_{h\perp}}{z} - \boldsymbol{k}_T\right) \\ \mathrm{Tr}\left[\Phi\left(x, \boldsymbol{p}_T, S\right)\gamma^\mu \Delta\left(z, \boldsymbol{k}_T\right)\gamma^\nu\right], \end{aligned} \tag{2.61}$$

where

$$\begin{aligned} \Phi\left(x, \boldsymbol{p}_T, S\right) &\equiv \int \mathrm{d}p^+ \Phi(p, P, S)\bigg|_{p^-=xP^-}, \\ \Delta\left(z, \boldsymbol{k}_T\right) &\equiv \frac{1}{4z}\int \mathrm{d}k^- \Delta\left(k, P_h\right)\bigg|_{k^+=P_h^+/z} \end{aligned} \tag{2.62}$$

are the transverse-momentum dependent correlators (Fig. 2.3). At leading twist, the correlator Φ in (2.62) gives rise to 8 Transverse Momentum Dependent Parton Distribution Functions (TMD PDFs or TMDs), corresponding to different polarization states of the nucleon and the quark, which are collected in Table 2.1. Being leading twist functions, they have a probabilistic interpretation (Fig. 2.4). For instance, a function which has attracted lot of attention in recent years is the Sivers function $f_{1T}^\perp\left(x, p_T\right)$, first proposed in [34, 56], which can be interpreted as the probability difference for a quark transverse momentum to point right or left, when the nucleon is moving towards us, and its spin is pointing upward. In this thesis

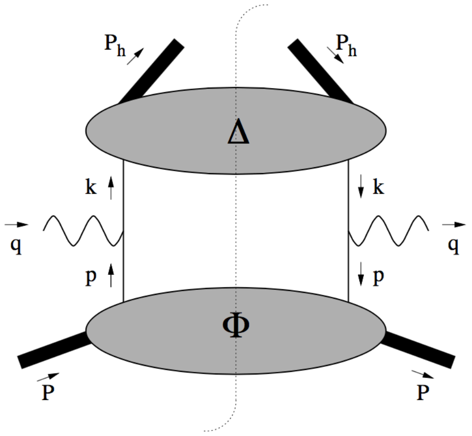


Figure 2.3: The cut diagram for SIDIS.

quark pol.

nucleon pol.

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

Table 2.1: The 8 leading twist TMD PDFs, for different nucleon and quark polarization. Functions in red are T-odd, while functions in black survive integration over the transverse momentum, and hence have a collinear PDF counterpart.

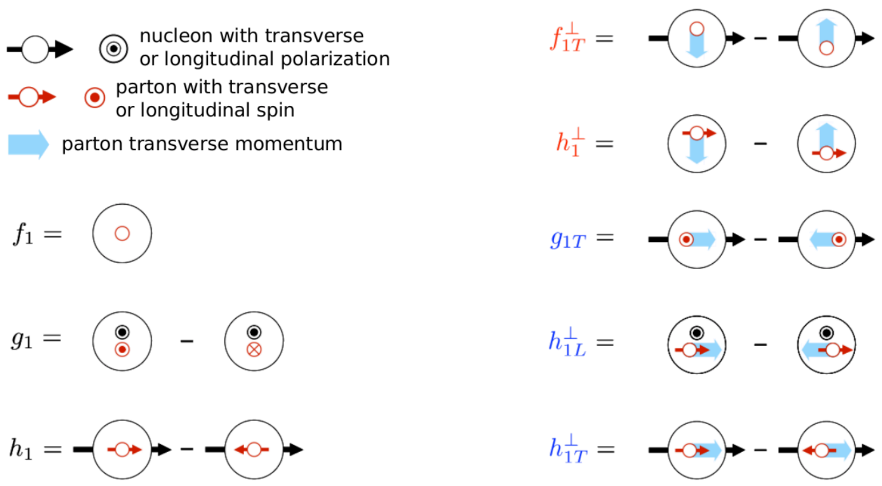


Figure 2.4: Probabilistic interpretation of leading twist TMDs. The nucleon is moving towards the reader.

we will concentrate on the most basic TMD, i.e., the function $f_1(x, p_T)$ describing the transverse momentum probability distribution of unpolarized quarks in an unpolarized nucleon. Besides giving precious information by itself, its knowledge is also needed for the extraction of other functions, since different TMDs have some common evolution properties, and since unpolarized TMDs appear in the denominator of asymmetries. Concerning the fragmentation correlator Δ, if we restrict ourselves to final hadrons with spin 0, or unreconstructed spin, we remain with 2 leading twist Transverse Momentum Dependent Fragmentation Functions (TMD FFs):

$$D_1(z, k_T), \qquad H_1^{\perp}(z, k_T), \tag{2.63}$$

describing respectively the fragmentation of an unpolarized and transversely polarized quark (the function $H_1^{\perp}$ is known as Collins function from [57]).

TMD factorization theorems establish that cross sections at $q_T \ll Q$ can be written as convolutions of TMD PDFs and FFs. For instance, the SIDIS unpolarized structure function in (2.56) can be expressed in an equivalent way as

$$\begin{aligned} F_{UU,T} \propto \sum_q e_q^2 \int d^2\boldsymbol{p}_{T1} d^2\boldsymbol{p}_{T2} \delta^{(2)}\left(\boldsymbol{q}_T + \boldsymbol{p}_T - \boldsymbol{k}_T\right) \\ \times f_1^{q/N}(x, |\boldsymbol{p}_T|) D_1^{h/q}(z, z\,|\boldsymbol{k}_T|) \end{aligned} \tag{2.64}$$

where $f_1^{q/N}$ and $D_1^{h/q}$ are respectively the TMD PDF for nucleon N and TMD FF for hadron h (we neglected for the moment the dependence on factorization scales, which will be discussed in Ch. 3).

It is worth noticing the relevance of gauge links in the case of Transverse Momentum Dependent Parton Distribution Functions. The quark fields in (2.62) are separated in the transverse plane, at variance with the integrated correlator in (2.21):

$$\Phi_{IJ}\left(x, \boldsymbol{p}_T, S\right) = \int \frac{d\xi^+ d^2\boldsymbol{\xi}_T}{(2\pi)^3} e^{ip\cdot\xi} \left\langle P, S \left| \overline{\psi}_J(0)\psi(\xi)_I \right| P, S \right\rangle \Big|_{\xi^-=0} \tag{2.65}$$

meaning that the gauge link to be inserted in the above equation will be different from 1, regardless of the gauge choice. Gauge links are responsible for properties peculiar to TMDs. One of these is the appearance of a new class of divergences, so called *rapidity divergences* [58], which set some of the evolution properties of these functions. Another important effect is the T-oddness of certain TMDs. This last property is due to the fact that gauge links are process dependent: the path in (2.24) is determined by the process considered, and in particular depends on whether the gluon radiation is in the final or initial state [59]. The leading twist T-odd TMDs are the Sivers [34, 56] and Boer-Mulders [60] functions, respectively $f_{1T}^{\perp}$ and $h_1^{\perp}$ in Table 2.1, and the Collins fragmentation function $H_1^{\perp}$ [57].

2.3 Drell-Yan

Drell-Yan is the process where a lepton-antilepton pair from the decay of a virtual photon or Z-boson is produced in the high energy scattering of two hadrons (we will not consider polarized beams in this thesis):

$$N_1(P_1) + N_2(P_2) \to \gamma^*/Z(q) + X \to \ell^+(l) + \ell^-(l') + X. \tag{2.66}$$

The cross section can be expressed in terms of lepton and hadron tensor as

$$\frac{d\sigma}{d^4q d\Omega} = \frac{\alpha_{em}^2}{4Q^4 s} L_{\mu\nu} W^{\mu\nu}, \tag{2.67}$$

where $Q^2 \equiv q^2$, and we have chosen as independent variables the 4-momentum q^μ of the vector boson and the solid angle Ω of the lepton l^+ in a given rest frame of the boson. One example of such a frame is the Collins-Soper frame[8], depicted in Fig. 2.5. Like for SIDIS, quantum interference between between different scattering amplitudes originates modulations in the lepton angles θ and ϕ. Although in this

---

[8] Another commonly chosen frame is the Gottfried-Jackson frame, where the $\hat{z}$ axis points along the direction of hadron N_1

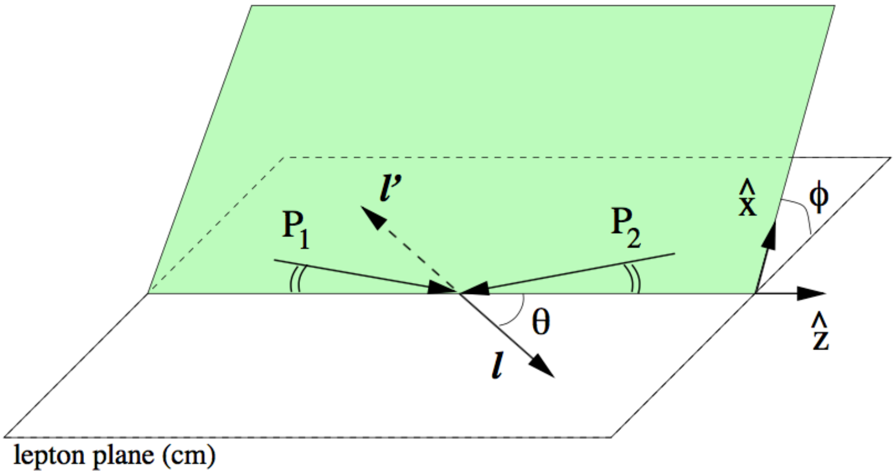


Figure 2.5: The Collins-Soper frame. The two hadron momenta define the $\hat{x} - \hat{z}$ plane, while the $\hat{z}$ axis is defined as the direction bisecting the angle between $\boldsymbol{P}_1$ and $-\boldsymbol{P}_2$

thesis we will only deal with cross sections integrated over azimuthal angles, it is worth briefly discussing the angular structure, since this will be of some relevance when calculating the effect of experimental cuts on the lepton phase space, in Ch. 4. The cross section in terms of structure functions reads [61–63][9]

$$\begin{aligned} \frac{d\sigma}{d^4q d\Omega} = \frac{\alpha^2}{2Q^2 s} \Big[& W_T \left(1 + \cos^2\theta\right) + W_L \left(1 - \cos^2\theta\right) \\ & + W_{+0} \sin 2\theta \cos\phi + W_{+-} \sin^2\theta \cos 2\phi \Big] , \end{aligned} \tag{2.68}$$

where the subscripts denote contraction of the hadron tensor with different polarization vectors for the virtual boson. More precisely, the decomposition of the hadron tensor is

$$\begin{aligned} W^{\mu\nu} = & - (g^{\mu\nu} - T^\mu T^\nu) (W_T + W_{+-}) - 2X^\mu X^\nu W_{+-} \\ & + Z^\mu Z^\nu (W_L - W_T - W_{+-}) - (X^\mu Z^\nu + Z^\mu X^\nu) W_{+0}, \end{aligned} \tag{2.69}$$

where $T^\mu = q^\mu / Q$ is the time-like polarization direction, while the three orthogonal space-like vectors X^μ, Y^μ, Z^μ are defined according to the direction of $\hat{x}$ and $\hat{z}$ in

---

[9] The hadron tensor W in the convention of [61] should be divided by $(2\pi)^4$ to obtain the one in our convention

Fig. 2.5:

$$\begin{aligned} Z^\mu &= \frac{2}{s\sqrt{Q^2+q_T^2}}\left((P_2\cdot q)\,P_1^\mu-(P_1\cdot q)\,P_2^\mu\right), \\ X^\mu &= -\frac{2Q}{sQ_T\sqrt{Q^2+Q_T^2}}\left((P_2\cdot q)\,\tilde{P}_1^\mu+(P_1\cdot q)\,\tilde{P}_2^\mu\right), \end{aligned} \tag{2.70}$$

with $\tilde{P}_i^\mu \equiv P_i^\mu - q^\mu\,(P_1\cdot q)\,/Q^2$. In the same way, the lepton tensor can be decomposed in terms of a helicity basis. After integrating over the angles, the differential cross section reads

$$\frac{d\sigma}{d^4q}=\frac{4\pi\alpha_{em}^2}{3Q^2s}\left[2W_T+W_L\right]. \tag{2.71}$$

In the parton model, applicable when Q^2 is high, the intermediate boson is produced by the annihilation of a quark from one hadron and an antiquark from the other one. The tensor $W^{\mu\nu}$ is then

$$\begin{aligned} W^{\mu\nu} &= \sum_X \left\langle P_1P_2\left|J^{\mu\dagger}(0)\right|X\right\rangle \langle X\left|J^\nu(0)\right|P_1P_2\rangle\,(2\pi)^4\delta^{(4)}\left(P_1+P_2-P_X-q\right) \\ &= \int d^4x e^{-iqx}\left\langle P_1P_2\left|J^{\mu\dagger}(x)J^\nu(0)\right|P_1P_2\right\rangle, \end{aligned} \tag{2.72}$$

with $J^\mu(x)$ the electromagnetic quark current

$$J^\mu(x)=\sum_q e_q\overline{\psi}_q(x)\gamma^\mu\psi_q(x). \tag{2.73}$$

As in the case of SIDIS, one needs to distinguish two different kinematic regimes depending on q_T, i.e., the transverse momentum of the boson in any frame where the colliding hadrons are collinear. When q_T is of the order of the hard scale Q, the correct prediction is given by fixed-order collinear factorization, where the hard cross section is convoluted with ordinary PDFs. We will discuss collinear factorization for Drell-Yan in Ch. 5. In chapter 4 instead we will present a phenomenological analysis of the regime $q_T \ll Q$, where CSS resummation, or equivalently, TMD factorization should be applied. In this regime, the structure function W_L can be neglected, and the term W_T can be written as a convolution of two TMD PDFs, one for each of the incoming hadrons

$$\begin{aligned} \frac{\mathrm{d}\sigma}{\mathrm{d}Q^2\mathrm{d}q_T^2\mathrm{d}y} &\propto \sum_q e_q^2\int d^2\boldsymbol{p}_{T1}d^2\boldsymbol{p}_{T2}\delta^{(2)}\left(\boldsymbol{q}_T-\boldsymbol{p}_{T1}-\boldsymbol{p}_{T2}\right) \\ &\times\left[f_1^{q/N_1}(x_1,|\boldsymbol{p}_{T1}|)f_1^{\bar{q}/N_2}(x_2,|\boldsymbol{p}_{T2}|)+\quad q\leftrightarrows\bar{q}\quad\right]. \end{aligned} \tag{2.74}$$

In (2.74), we wrote the differential cross section in terms of common experimental variables: Q^2, q_T and the rapidity y of the vector boson in the c.m. of the two hadrons

$$y = \frac{1}{2}\log\left(\frac{q^0 + q_z}{q^0 - q_z}\right). \tag{2.75}$$

The longitudinal fractional momenta in the TMD formula above are

$$x_1 = \frac{Q}{\sqrt{s}}e^{y}, \quad x_2 = \frac{Q}{\sqrt{s}}e^{-y}. \tag{2.76}$$

2A Appendix: Notation and conventions

Lorentz vectors and metric

In this thesis, we use round brackets to express a vector in terms of its Minkowski components:

$$v^\mu = \left(v^0, v^x, v^y, v^z\right)$$

, while we use square brackets for light-cone components:

$$\begin{aligned} v^\mu &= \left[v^-, v^+, \mathbf{v_T}\right] \\ &\equiv \left[\frac{v^0 - v^z}{\sqrt{2}}, \frac{v^0 + v^z}{\sqrt{2}}, v^x, v^y\right]. \end{aligned}$$

Vectors in the 2-dimensional transverse space are denoted in bold-face.
We use the metric tensor defined as

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

Plus-distributions

Plus-distributions are defined as

$$(f(x))_+ \equiv \lim_{\varepsilon \to 0}\left\{f(x)\theta(1 - x - \varepsilon) - \delta(1 - x - \varepsilon)\int_0^{1-\varepsilon} f(y)\mathrm{d}y\right\}.$$

For instance, one has, for a regular test function $T(x)$:

$$\int_x^1 \mathrm{d}y\frac{T(y)}{(1-y)_+} = \int_x^1 \mathrm{d}y\frac{T(y) - T(1)}{1 - y} - T(1)\ln\frac{1}{1-x}.$$

Chapter 3

Low transverse momentum: formalism

In this Chapter we will discuss more in detail the theoretical description of transverse momentum dependent cross sections in SIDIS,

$$\frac{\mathrm{d}\sigma^{\ell\, N\to\ell'\, h\, X}}{\mathrm{d}Q^2\mathrm{d}x\mathrm{d}z\mathrm{d}P_{hT}^2}, \tag{3.1}$$

and Drell-Yan

$$\frac{\mathrm{d}\sigma^{N_1\, N_2\to\ell^+\,\ell^-\, X}}{\mathrm{d}Q^2\mathrm{d}y\mathrm{d}q_T^2}, \tag{3.2}$$

concentrating on the TMD formalism, valid at low-q_T. As seen in Ch. 2, cross sections in (3.1) and (3.2) can be expressed in terms of transverse and longitudinal structure functions. The precise relations, considering electromagnetic scattering, are

$$\frac{\mathrm{d}\sigma^{\ell\, N\to\ell'\, h\, X}}{\mathrm{d}Q^2\mathrm{d}x\mathrm{d}z\mathrm{d}P_{hT}^2} = \frac{(2\pi)^2\alpha_{em}^2}{xQ^4}\left\{\left(1-y+\frac{y^2}{2}\right)F_{UU,T}^{SIDIS} + (1-y)\,F_{UU,L}^{SIDIS}\right\}, \tag{3.3}$$

and

$$\frac{\mathrm{d}\sigma^{N_1\, N_2\to\ell^+\,\ell^-\, X}}{\mathrm{d}Q^2\mathrm{d}y\mathrm{d}q_T^2} = \frac{2\pi^2\alpha_{em}^2}{3Q^2 s}\left[2F_{UU,T}^{DY} + F_{UU,L}^{DY}\right], \tag{3.4}$$

where for uniformity of notation we have redefined W_T and W_L in (2.71) as $F_{UU,T}^{DY}$ and $F_{UU,L}^{DY}$. The hadron tensors in (3.3) and (3.4) in the region $q_T \ll Q$ satisfy TMD factorization, i.e., factorization in terms of Transverse Momentum Distributions. In the following we will not consider the contribution of longitudinal structure functions, being negligible in this kinematic region. Up to now, we only presented parton model formulas (Eqs. (2.64) and (2.74)). The full QCD result, including scale dependencies generated by factorization and renormalization, reads [26, 64]

$$
\begin{aligned}
F_{UU,T}^{SIDIS} &= \sum_q \left|\mathcal{H}_{SIDIS}(Q;\mu)^2\right| \int d^2\mathbf{p}_T d^2\mathbf{k}_T \delta^{(2)}\left(\mathbf{p}_T + \mathbf{q}_T - \mathbf{k}_T\right) \\
&\times f_1^{q/N}\left(x, |\mathbf{p}_T|;\mu;\zeta_f\right) D_1^{h/q}\left(z, z\,|\mathbf{k}_T|;\mu;\zeta_D\right) \\
&\equiv \sum_q \left|\mathcal{H}_{SIDIS}(Q;\mu)^2\right| \int \frac{d^2\mathbf{b}}{(2\pi)^2} e^{-i\mathbf{q}_T\cdot\mathbf{b}} \\
&\times \tilde{f}_1^{q/N}\left(x, b;\mu;\zeta_f\right) \tilde{D}_1^{h/q}\left(z, b;\mu;\zeta_D\right),
\end{aligned}
\tag{3.5}
$$

$$
\begin{aligned}
F_T^{DY} &= \sum_q \left|\mathcal{H}_{DY}(Q;\mu)^2\right| \int d^2\mathbf{p}_{T1} d^2\mathbf{p}_{T2} \delta^{(2)}\left(\mathbf{p}_{T1} + \mathbf{p}_{T2} - \mathbf{q}_T\right) \\
&\times f_1^{q/N_1}\left(x, |\mathbf{p}_{T1}|;\mu;\zeta_{f1}\right) f_1^{\bar{q}/N_2}\left(x, |\mathbf{p}_{T2}|;\mu;\zeta_{f2}\right) \\
&\equiv \sum_q \left|\mathcal{H}_{DY}(Q;\mu)^2\right| \int \frac{d^2\mathbf{b}}{(2\pi)^2} e^{i\mathbf{q}_T\cdot\mathbf{b}} \\
&\times \tilde{f}_1^{q/N_1}\left(x_1, b;\mu;\zeta_{f1}\right) \tilde{f}_1^{\bar{q}/N_2}\left(x_2, b;\mu;\zeta_{f2}\right).
\end{aligned}
\tag{3.6}
$$

We defined $\tilde{f}_1$ and $\tilde{D}_1$ as the Fourier transformed of the respective TMDs, since TMD evolution is more easily written in the conjugate transverse space **b**, which ensures transverse momentum conservation:

$$
\begin{aligned}
f_1^{q/N}\left(x, |\mathbf{p}_T|;\mu;\zeta_f\right) &\equiv \int \frac{\mathrm{d}^2\boldsymbol{b}}{(2\pi)^2} e^{i\boldsymbol{p}_\mathrm{T}\cdot\boldsymbol{b}} \tilde{f}_1^{q/N}\left(x, b;\mu;\zeta_f\right), \\
D_1^{h/q}\left(z, z\,|\mathbf{k}_T|;\mu;\zeta_D\right) &\equiv \int \frac{\mathrm{d}^2\mathbf{b}}{(2\pi)^2} e^{-i\boldsymbol{k}_\mathrm{T}\cdot\mathbf{b}} \tilde{D}_1^{h/q}\left(z, b;\mu;\zeta_D\right).
\end{aligned}
\tag{3.7}
$$

The hard factor $\mathcal{H}_X(Q;\mu)$ is process dependent and is calculated in a fixed-order expansion in the strong coupling α_s. TMD PDFs and FFs obey evolution equations in two independent scales: besides the renormalization scale μ, the rapidity scale ζ appears, as a consequence of the regularization of *rapidity divergences*. This kind of divergences cannot be classified as infrared nor ultraviolet. They are rather artificial divergences, introduced when separating the cross section in terms of two TMDs: if a unregulated definition was adopted, the gauge link inside a TMD would contain gluons moving in the light-cone direction opposite to that of the parent nucleon. Also, that would not take into account the role of soft gluon radiation, which is embodied in the *soft factor*.[1] The regularization is done by including one part of the soft factor in each of the TMDs, and setting a rapidity cutoff y_c to separate left-moving ($y < y_c$) and right-moving ($y > y_c$) particles. For instance, in the case of SIDIS, the two scales ζ_f and ζ_D are defined by

$$\zeta_f = x^2 M_N^2 e^{2(y_N - y_c)}, \qquad \zeta_D = \frac{M_h^2}{z^2} e^{2(y_c - y_h)}, \tag{3.8}$$

where $M_{N(h)}$ and $y_{N(h)}$ are the mass and rapidity of the initial nucleon (final hadron). By construction one has $\zeta_f \zeta_D = Q^4$. A common choice is to set $\zeta_f = \zeta_D = Q^2$. A convenient choice for the scale μ is instead $\mu \simeq Q$, to avoid large logarithms $\ln(Q^2/\mu^2)$ in the hard factor.

3.1 TMD evolution

3.1.1 Perturbative terms

The Renormalization Group Equation, originating from the renormalization of ultraviolet divergences, dictates the evolution in μ:

$$\frac{\partial \ln \tilde{f}_1 (x, b; \mu, \zeta)}{\partial \ln \mu} = \gamma_f \left(\alpha_s(\mu); \zeta/\mu^2\right), \tag{3.9}$$

where γ_f is called the *TMD anomalous dimension*[2]. The evolution with respect to the rapidity scale ζ is instead related to the derivative of the soft factor with respect to the rapidity cutoff y_c, and has the form

$$\frac{\partial \ln \tilde{f}_1 (x, b; \mu, \zeta)}{\partial \ln \sqrt{\zeta}} = \tilde{K} (b; \mu) . \tag{3.10}$$

The *rapidity anomalous dimension* $\tilde{K}$ also depends on μ, and evolves through its own RGE:

$$\frac{d\tilde{K} (b; \mu)}{d \ln \mu} = -\gamma_K(\alpha_s(\mu)), \tag{3.11}$$

where γ_K is called the *cusp anomalous dimension*. An important relation between the anomalous dimensions follows from the continuity of second derivatives of the TMD function:

$$\frac{\partial^2 \ln \tilde{f}_1}{\partial \ln \sqrt{\zeta}\, \partial \ln \mu} = \frac{\partial^2 \ln \tilde{f}_1}{\partial \ln \mu\, \partial \ln \sqrt{\zeta}}. \tag{3.12}$$

---

[1] The soft factor is a vacuum expectation value of a loop of Wilson lines, describing emission of gluons in the central rapidity and near-on-shell region. For all the details on TMD definition and evolution, including the origin of Eq. (3.10), see [26, 64].

[2] Evolution equations and anomalous dimension for TMD PDFs and FFs are identical.

Inserting Eqs. (3.9) to (3.11), this translates into

$$\frac{\partial \gamma_f\left(\alpha_s(\mu);\zeta/\mu^2\right)}{\partial \ln\sqrt{\zeta}} = -\gamma_K\left(\alpha_s(\mu)\right). \tag{3.13}$$

Hence, evolving the TMD anomalous dimension γ_f in rapidity, from the starting point $\zeta_0 = \mu^2$, yields the relation

$$\gamma_f\left(\alpha_s(\mu);\zeta/\mu^2\right) = \gamma_f\left(\alpha_s(\mu);1\right) - \gamma_K(\alpha_s(\mu)) \ln\frac{\sqrt{\zeta}}{\mu}. \tag{3.14}$$

The quantities $\gamma_f\left(\alpha_s(\mu);1\right)$, $\tilde{K}\left(b;\mu\right)$,$\gamma_K(\alpha_s(\mu))$ all obey a perturbative expansion in powers of α_s:

$$\begin{aligned}\gamma_f\left(\alpha_s(\mu);1\right) &= \sum_{n=1}^{\infty}\left(\frac{\alpha_s(\mu)}{\pi}\right)^n \gamma_f^{(n)},\\ \gamma_K\left(\alpha_s(\mu)\right) &= \sum_{n=1}^{\infty}\left(\frac{\alpha_s(\mu)}{\pi}\right)^n \gamma_K^{(n)},\\ \tilde{K}\left(b;\mu\right) &= \sum_{n=1}^{\infty}\left(\frac{\alpha_s(\mu)}{\pi}\right)^n \tilde{K}^{(n)}(b;\mu).\end{aligned} \tag{3.15}$$

Perturbative coefficients up to order α_s^3 can be found in [65, 66]. These expansions are valid as long as the scale μ is within a perturbative range, and the last expression is intended for small values of b ($b \ll 1/\Lambda$). Here we report just the first order coefficients:

$$\gamma_f^{(1)} = C_F\frac{3}{2}, \quad \gamma_K^{(1)} = 2C_F, \quad \tilde{K}^{(1)}(b;\mu) = -C_F\ln\left(\frac{\mu^2 b^2}{b_0^2}\right), \tag{3.16}$$

with $b_0 \equiv 2e^{-\gamma_E}$.

TMD evolution is thus an evolution in the bi-dimensional μ, ζ space, which can be applied in b-space through a multiplicative evolution kernel $\mathrm{E}(\mu,\zeta;\mu_0,\zeta_0)$:

$$\tilde{f}_1\left(x,b;\mu,\zeta\right) = \mathrm{E}(\mu,\zeta;\mu_0,\zeta_0)\tilde{f}_1\left(x,b;\mu_0,\zeta_0\right). \tag{3.17}$$

In this work, we will make use of a solution for the evolution kernel adopted in previous phenomenological studies (see, e.g., [25, 64, 67]), which is also the solution most close to the standard CSS formalism.[3] Since the perturbative expressions

---

[3] Neglecting higher order corrections, our choice coincides with the standard CSS; the explicit proof of this correspondence can be found for instance in Appendix B of [32], while a comparison of different solutions for TMD evolution can be found in [65, 68].

for the initial scale TMD $\tilde{f}_1(x, b; \mu_0, \zeta_0)$ and in the evolution kernel in Eq. (3.17) contain logarithms of the form $\ln\left(\frac{\mu_0^2 b^2}{b_0^2}\right)$ and $\ln\left(\frac{\zeta_0}{\mu_0^2}\right)$, we minimize these terms by adopting the natural choice for central values of the initial scales:

$$\mu_0 = \mu_b, \qquad \zeta_0 = \mu_b^2, \qquad \mu_b \equiv \frac{b_0}{b}. \tag{3.18}$$

Moreover, we evolve first in ζ and then in μ:

$$\begin{aligned}\tilde{f}_1(x, b; \mu, \zeta) =& \tilde{f}_1\left(x, b; \mu_b, \mu_b^2\right)\\ \times& \left(\frac{\sqrt{\zeta}}{\mu_b}\right)^{\tilde{K}(b;\mu_b)} \exp\left\{\int_{\mu_b}^{\mu} \frac{d\mu'}{\mu'}\left[\gamma_f\left(\alpha_s\left(\mu'\right);1\right) - \ln\frac{\sqrt{\zeta}}{\mu'}\gamma_K\left(\alpha_s\left(\mu'\right)\right)\right]\right\}.\end{aligned} \tag{3.19}$$

The second line of the above equation is the evolution kernel $\mathrm{E}(\mu, \zeta; \mu_0 = \mu_b, zeta_0 = \mu_b^2)$ within our prescription. As can be checked from (3.16), this choice of scales implies $\tilde{K} = 0$ up to order α_s. Scale variations can be studied by setting $(\mu_0, \zeta_0) = (c_{\mu_0}\mu_b, c_{\zeta_0}\mu_b^2)$, and varying the factors c_i.

In principle, the regularity of $\tilde{f}_1$ as a function of μ and ζ should guarantee that the evolution kernel $\mathrm{E}(\mu, \zeta; \mu_0, \zeta_0)$ is independent of the path chosen to integrate the evolution equations from (μ_0, ζ_0) to (μ, ζ). However, the perturbative truncation of $\tilde{K}$ to a given order in α_s breaks the continuity condition (3.12), introducing some path dependence. A detailed study on the issue can be found in [69].

At small values of b, i.e., $b \ll 1/\Lambda$, TMDs can be calculated perturbatively, through an Operator Product Expansion (OPE) [70] around $b = 0$, and their expression reads[4]

$$\begin{aligned}\tilde{f}_{1\,OPE}^{a/N}(x, b; \mu, \zeta) &= \sum_i \int_x^1 \frac{\mathrm{d}\hat{x}}{\hat{x}} \tilde{C}_{a/i}^{in}\left(x/\hat{x}, b; \mu, \zeta, \alpha_s(\mu)\right) f_1^{i/N}(\hat{x}, \mu),\\ \tilde{D}_{1\,OPE}^{h/a}(z, b; \mu, \zeta) &= \frac{1}{z^2}\sum_i \int_x^1 \frac{\mathrm{d}\hat{z}}{\hat{z}} \tilde{C}_{i/a}^{out}\left(z/\hat{z}, b; \mu, \zeta, \alpha_s(\mu)\right) D_1^{h/i}(\hat{z}, \mu),\end{aligned} \tag{3.20}$$

where $f_1^{i/N}(\hat{x}, \mu)$ $(D_1^{h/i}(\hat{z}, \mu))$ is the collinear PDF(FF) for parton species i. The explicit calculation of the Wilson coefficients $\tilde{C}^{in}$ and $\tilde{C}^{out}$ to order α_s can be found in [26, 64], while the expressions to order α_s^2 can be found for instance in [66, 68]. Potentially large logarithmic terms in these expressions are controlled by the choice of scales in (3.18).

---

[4] Our $\tilde{C}^{out}$ coefficients are defined in a different way from [26, 64]:

$$\tilde{C}^{out}\Big|_{\text{ours}}(z, b; \ldots) = z^2 \tilde{C}^{out}\Big|_{\text{Collins}}(z, b; \ldots).$$

This is done to facilitate the comparison to collinear factorized formulas in Ch. 4.

3.1.2 Logarithmic ordering

In this subsection we specify and motivate the notation we will use to indicate different levels of perturbative accuracy in the TMD evolution formula. The TMD evolution kernel E (also known as Sudakov factor) can be written as an infinite series of logarithms

$$\mathrm{E}^2(Q, Q^2; \mu_b, \mu_b^2) = 1 + \sum_{n=1}^{\infty} \alpha_s^n \sum_{k=1}^{2n} \tilde{E}^{(n,k)} \ln^k \frac{Q^2}{\mu_b^2}. \tag{3.21}$$

In (3.21) we wrote E^2 because we are interested in the logarithmic expansion of the cross section, which always contains two TMDs. Terms with $k = 2n$ are the only terms taken into account in a *leading logarithmic* (LL) expression, and more generally $k = 2n-m$ $(m \geq 0)$ corresponds to N^mLL logarithmic accuracy. Different levels of logarithmic accuracy correspond to different perturbative accuracy of the anomalous dimensions in (3.15). The remaining perturbative ingredients for the cross section, i.e., the hard factor $\mathcal{H}$ and the Wilson coefficients $\tilde{C}^{in(out)}$ also have a perturbative expansion in α_s, but without logarithms of the type $\ln Q^2/\mu_b^2$. They should however be taken to the proper order in α_s, to attain the desired logarithmic accuracy. Let us consider, indeed, an expansion of (3.21) to N^mLL, and check what is the effect of multiplying by a term proportional to α_s^p:

$$\begin{aligned} \alpha_s^p \mathrm{E}^2_{\mathrm{N}^m\mathrm{LL}} &= \alpha_s^p \sum_{n=1}^{\infty} \alpha_s^n \sum_{k=1}^{2n-m} \tilde{E}^{(n,k)} \ln^k \frac{Q^2}{\mu_b^2}, \\ &= \sum_{n'=1+p}^{\infty} \alpha_s^{n'} \sum_{k=1}^{2n'-2p-m} \tilde{E}^{(n'-p,k)} \ln^k \frac{Q^2}{\mu_b^2} \sim \mathrm{N}^{m+2p}\mathrm{LL}, \end{aligned} \tag{3.22}$$

where in the last line we have defined $n' \equiv n + p$, and the symbol $\sim$ denotes that what we obtain are terms of logarithmic accuracy N^{m+2p}LL. This means that in a NNLL framework, one should take $\mathcal{H}$ and $\tilde{C}^{in(out)}$ at least up to order α_s, in order to take into account all contributions with the proper logarithmic counting. This is not necessary at LL and NLL, where one can safely take them at leading order, i.e., α_s^0, as higher order terms would introduce corrections of higher logarithmic accuracy. There is however a different convention, often adopted in studies within the CSS framework (see, e.g., [19, 22, 71, 72]), where at N^mLL the coefficients $\mathcal{H}$ and $\tilde{C}^{in(out)}$ are taken at order α_s^m. To distinguish, we will denote calculations done with this convention with a $'$ symbol, i.e., $\mathrm{N}^m\mathrm{LL}'$. In any case the cusp anomalous dimension is taken at one order higher, α_s^{m+1}, since it multiplies a logarithm $\ln(Q/\mu_b)$, while γ_f and $\tilde{K}$ are taken at order α_s^m. The two different prescriptions are summarized in Table 3.1.

	$\mathcal{H}, \tilde{C}^{in(out)}$	γ_K	γ_f	$\tilde{K}$
LL	α_s^0	α_s^1	α_s^0	α_s^0
$\mathrm{N}^{m\geq 1}\mathrm{LL}$	α_s^{m-1}	α_s^{m+1}	α_s^m	α_s^m
$\mathrm{N}^{m\geq 1}\mathrm{LL}'$	α_s^m	α_s^{m+1}	α_s^m	α_s^m

Table 3.1: Nomenclature for different logarithmic orders in the TMD formula. Notice the distinction between $\mathrm{N}^m\mathrm{LL}$ and $\mathrm{N}^m\mathrm{LL}'$ for $m \geq 1$.

3.1.3 Non-perturbative terms and b^* prescriptions

At large b, $b \gtrsim 1/\Lambda$, not only the OPE in (3.20) is not applicable, but also the strong coupling $\alpha_s(\mu_b)$ in the perturbative expansions in (3.15) and in the Wilson coefficients hits the Landau pole. Since the Fourier integral in (3.7) extends over the full **b** space, a common way to avoid this issue is to make the substitution $b \longrightarrow b^*(b)$ inside $\tilde{f}_1\left(x, b; \mu; \zeta_f\right)$ and $\tilde{D}_1\left(z, b; \mu; \zeta_D\right)$, where $b^*(b)$ is a cutoff function, defined in such a way that $b^* \simeq b$ when b is below a certain threshold $b < b_{max}$, and $b^* \simeq b_{max}$ beyond the threshold. This is commonly accompanied by the introduction of a function $\tilde{f}_{NP}(x, b, Q^2)$ ($\tilde{D}_{NP}(z, b, Q^2)$), to describe nonperturbative physics at large b. The final expression for the evolved TMD PDF becomes then

$$\begin{aligned}\tilde{f}_1\left(x, b; \mu, \zeta\right) =& \tilde{f}_{1\,OPE}\left(x, b^*; \mu_{b^*}, \mu_{b^*}^2\right)\, \tilde{f}_{NP}(x, b, \zeta) \\ \times& \left(\frac{\sqrt{\zeta}}{\mu_{b^*}}\right)^{\tilde{K}(b;\mu_{b^*})} \exp\left\{\int_{\mu_{b^*}}^{\mu} \frac{d\mu'}{\mu'}\left[\gamma_f\left(\alpha_s\left(\mu'\right); 1\right) - \ln\frac{\sqrt{\zeta}}{\mu'}\gamma_K\left(\alpha_s\left(\mu'\right)\right)\right]\right\}\end{aligned} \tag{3.23}$$

(analogously for TMD FFs). A classical choice for the function $b^*(b)$, introduced in [16] and used in many phenomenological works, is

$$b^*(b) = \frac{b}{\sqrt{1 + \frac{b^2}{b_{max}^2}}}, \tag{3.24}$$

but of course this is not the only possible choice. We will explore also an exponential form, adopted in [67], which yields a more sharp transition (see Fig. 3.1:

$$b^*_{exp}(b) = b_{max}\left[1 - \exp(-b^4/b_{max}^4)\right]^{\frac{1}{4}}. \tag{3.25}$$

Practically, the b^* function can be chosen in such a way to modify not only the large distance region $b \gtrsim 1/\Lambda$, but also the very short distance region, $b \lesssim 1/Q$, as

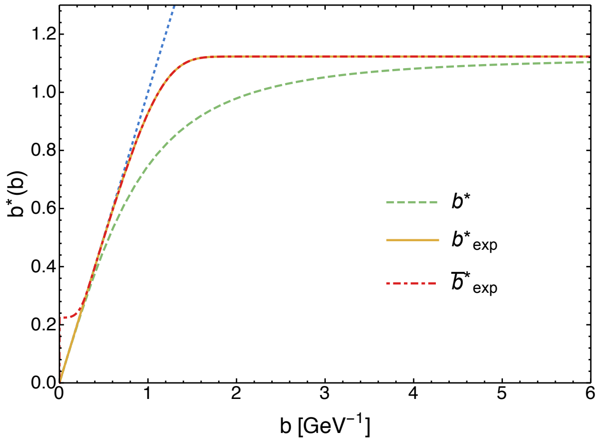


Figure 3.1: A comparison of the different forms for the function b^*. Dashed: Eq. (3.24); full line: Eq. (3.25); dot-dashed: Eq. (3.26). $b_{max} = 2e^{-\gamma_E} GeV^{-1}$ for all the curves, while b_{min} in Eq. (3.26) is set to $2e^{-\gamma_E}/5GeV^{-1}$ for illustrative purposes. The dotted line represents $b^* = b$, for comparison.

proposed in [73]. Such a modification is legitimate since it only affects the region where TMD factorization loses predictive power (in q_T-space the modification essentially affects the tail of TMDs, at $q_T \sim Q$), and is similar to the prescription introduced in [72], where logarithms $\ln(Q^2 b^2)$ in the resummed formula are replaced by $\ln(Q^2 b^2 + 1)$. In our work we make use of the following function

$$\bar{b}^*_{exp}(b) = b_{max}\left[\frac{1 - \exp(-b^4/b^4_{max})}{1 - \exp(-b^4/b^4_{min})}\right]^{\frac{1}{4}}, \tag{3.26}$$

already used in [25], with $b_{min} \sim 1/Q$. The advantages of prescriptions like (3.26) are:

- the scale μ_{b^*} in the Sudakov exponent in (3.23) never becomes greater than $\mu \sim Q$. This avoids the use of resummation in an unphysical region;
- consequently, resummed contributions are damped in the region $q_T \sim Q$, where they lose accuracy, thus improving the matching to fixed-order collinear factorization;
- the integral over $\int d^2\mathbf{q}_T$ of TMDs, or equivalently, the value of the Fourier transformed TMD $\tilde{f}_1(x, b; \mu, \zeta)$ in (3.23) at $b = 0$, is different from zero, at

variance with the standard CSS prescription. Its value has a precise physical meaning in relation to the q_T-integrated cross section, (2.29), as we will show.

The function $\tilde{f}_{NP}$ in (3.23) is the product of two terms:

$$\tilde{f}^a_{NP}(x, b, \zeta) = \tilde{f}^a_{1\,NP}(x, b) \left(\frac{\sqrt{\zeta}}{Q_0}\right)^{\tilde{K}_{NP}(b)} . \tag{3.27}$$

$\tilde{f}_{1\,NP}$ is the nonperturbative counterpart of $\tilde{f}_{1\,OPE}$ for large b, describing the intrinsic transverse momentum distribution from the parton-in-nucleon wave function, while $\tilde{K}(b)$ is the nonperturbative extension of the rapidity evolution kernel (we recall indeed that the kernel is b-dependent and is derived from the soft factor). In the above equation we have explicitly reinserted dependence on the flavor a, since the intrinsic distribution $\tilde{f}^a_{1\,NP}$ can have flavor-dependence, while the nonperturbative part of the evolution is independent of flavor and external hadron species, and also of the process. Dependence on the hard scale $\zeta \sim Q$ comes only from $\tilde{K}_{NP}$. Being nonperturbative functions, $\tilde{f}^a_{1\,NP}$ and $\tilde{K}_{NP}$ can be predicted through models, or extracted from fits to data with a suitable parametrization. The initial rapidity scale Q_0 in (3.27) should be though of as the scale of the model for $\tilde{f}^a_{1\,NP}$. It is usually set of the order of $1 \sim 2$ GeV. Different fits, like [19, 20, 29], have used an overall gaussian form for $\tilde{f}_{NP}$:

$$\tilde{f}^a_{NP}(x, b, \zeta) \propto \exp\left\{-b^2\left(g_1(x) + g_2 \ln\frac{\zeta}{Q_0^2}\right)\right\}, \tag{3.28}$$

but also a negative exponential form has been used, as well as more involved forms [25, 29, 74]. One should keep in mind that the actual TMD in the fit is a convolution of the nonperturbative form $\tilde{f}_{NP}$ and the perturbative parts, and hence details concerning the way the two parts are separated, like the value of b_{max} and the form of the function b^*, can affect the form of the nonperturbative function preferred by the fit.

3.2 Matching to fixed-order

As we have discussed in Sec. 2.2.2, the collinear factorization formula should be applied for transverse-momentum dependent cross sections at $q_T \sim O(Q)$, while TMD factorization is applicable at $q_T \ll Q$ and is subject to power corrections of the order $O(q_T/Q)$. One should thus switch from one formalism to the other when describing data from low to high q_T. However, an abrupt switch would not yield the most accurate predictions in the intermediate region where both the formalisms are applicable with some degree of accuracy[5]. Hence it is common to use the so-called $W+Y$ prescription [16, 71] to match the two regions[6]. In this prescription, W stands for the TMD factorized cross section in (3.5) and (3.6), while Y is a correction term to be applied in the intermediate region $M \ll q_T \lesssim Q$:

$$\begin{aligned}
&\frac{\mathrm{d}\sigma^X}{\mathrm{d}q_T\mathrm{d}\ldots}(\text{matched}) = W + Y, \\
&W \equiv \frac{\mathrm{d}\sigma^X}{\mathrm{d}q_T\mathrm{d}\ldots}(\text{TMD}), \\
&Y \equiv \frac{\mathrm{d}\sigma^X}{\mathrm{d}q_T\mathrm{d}\ldots}(\text{F.O.}) - \frac{\mathrm{d}\sigma^X}{\mathrm{d}q_T\mathrm{d}\ldots}(\text{asy}),
\end{aligned} \tag{3.29}$$

where $\mathrm{d}\sigma^X/(\mathrm{d}q_T\mathrm{d}\ldots)$ denotes either (3.1) or (3.2), and labels in parentheses denote, respectively, the TMD cross section, the fixed order collinear-factorized cross section (Eq. (2.50) in the case of SIDIS, the analogue for DY), and the asymptotic expansion discussed in Sec. 2.2.2. The idea behind (3.29) is the following: when $q_T \ll Q$, the collinear expression should be well approximated by its asymptotic expression, yielding a negligible contribution from the Y term. At higher q_T, the TMD formula needs $O(q_T/Q)$ corrections. These are well approximated by Y, since it contains precisely those terms from the collenar formula that are subleading. Eventually, when q_T is of the order of Q, the W term and the asymptotic term in (3.29) should cancel, leaving the fixed-order cross section. One should keep in mind that in practice cancellations are never perfect. Concerning the high-q_T limit of (3.29), it is true that the W term, when expanded to a given order in α_s, yields precisely the asymptotic expression at that order (see, e.g., Sec. 3.2 of [51] for the explicit formulae at leading order). However, in practice there is only a narrow range in q_T (if any) where this approximation works (see, e.g., Sec. 1.4 of [71]), after which one should switch to a pure collinear-factorized formula. In studies on q_T spectra for Tevatron and LHC (e.g., [71, 77, 78]), this switching usually happens at $q_T < Q$. The situation is particularly dramatic for SIDIS, since it has been shown that a true cancellation never occurs, due to the influence of nonperturbative components [32]. Also at very low q_T, errors in the cancellation of the asymptotic with the fixed-order can be sizeable, especially at low Q. A way to practically avoid the issue, as proposed in [73], is to introduce two damping functions, one to gradually set to 0 the Y term at low q_T, and one to set to 0 the W and asymptotic terms at relatively high q_T, before these terms become negative (which usually happens at $q_T < Q$).

---

[5] See, e.g., the discussion in Sec. 1.3 of [71].

[6] Alternative matching procedures were proposed also in, e.g., Refs. [75, 76].

3A Appendix: expansion of the evolution factor to $O(\alpha_s^2)$ and logarithmic ordering

To motivate the expansion of the evolution kernel in Eq. (3.21) and the logarithmic ordering in Table 3.1, we will discuss the fixed-order α_s^2 expansion of E, or, better, of E^2, since we always have two TMDs in the cross section formula. For simplicity, we will not consider scale variations at this stage, and set $\mu = Q$, $\zeta = Q^2$:

$$\mathrm{E}^2(Q, Q^2; \mu_b, \mu_b^2) = \\ \exp\left\{\tilde{K}(b;\mu_b)\ln\frac{Q^2}{\mu_b^2} + \int_{\mu_b^2}^{Q^2}\frac{d\mu'^2}{\mu'^2}\left[\gamma_f\left(\alpha_s\left(\mu'\right);1\right) - \ln\frac{\sqrt{\zeta}}{\mu'}\gamma_K\left(\alpha_s\left(\mu'\right)\right)\right]\right\}. \tag{3.30}$$

Since we are interested in a order α_s^2 expansion, we can retain just the first two terms in each of the perturbative series of the anomalous dimensions (3.15):

$$\begin{aligned}&\mathrm{E}^2(Q, Q^2; \mu_b, \mu_b^2) \simeq \\ &\exp\left\{\sum_{n=1,2} a_s^n(\mu_b)\tilde{K}^{(n)}(b;\mu_b)\ln\frac{Q^2}{\mu_b^2}\right. \\ &\left.+\int_{\mu_b^2}^{Q^2}\frac{d\mu'^2}{\mu'^2}\sum_{n=1,2} a_s^n(\mu')\left[\gamma_f^{(n)} - \ln\frac{\sqrt{\zeta}}{\mu'}\gamma_K^{(n)}\right]\right\},\end{aligned} \tag{3.31}$$

where $a_s \equiv \alpha_s/(4\pi)$, and we have consequently redefined the perturbative coefficients $\gamma_f^{(n)}$, $\gamma_K^{(n)}$ and $\tilde{K}^{(n)}$ to compensate for the factor of 4. In order to analytically solve the integral over μ', we need an expression for the running coupling. In the context of a second-order approximation, it is sufficient to consider the leading order RGE for a_s:

$$a_s(\mu) = \frac{a_s(Q)}{1 + a_s(Q)\beta_0\ln\left(\mu^2/Q^2\right)} \simeq a_s(Q)\left[1 + a_s(Q)\beta_0\ln\left(Q^2/\mu^2\right) + O\left(a_s^2\right)\right], \tag{3.32}$$

where we expanded around the hard scale Q. Inserting this expression in (3.31), one obtains

$$\begin{aligned}\mathrm{E}^2 \simeq \exp\Bigg\{&a_s(Q)\left[\tilde{K}^{(1)}\ln\frac{Q^2}{\mu_b^2} + \int_{\mu_b^2}^{Q^2}\frac{d\mu'^2}{\mu'^2}\left(\gamma_f^{(1)} - \gamma_K^{(1)}\ln\frac{Q^2}{\mu'^2}\right)\right] \\ &+ a_s^2(Q)\beta_0\left[\tilde{K}^{(1)}\ln^2\frac{Q^2}{\mu_b^2} + \int_{\mu_b^2}^{Q^2}\frac{d\mu'^2}{\mu'^2}\left(\gamma_f^{(1)}\ln\frac{Q^2}{\mu'^2} - \gamma_K^{(1)}\ln^2\frac{Q^2}{\mu'^2}\right)\right] \\ &+ a_s^2(Q)\left[\tilde{K}^{(2)}\ln\frac{Q^2}{\mu_b^2} + \int_{\mu_b^2}^{Q^2}\frac{d\mu'^2}{\mu'^2}\left(\gamma_f^{(2)} - \gamma_K^{(2)}\ln\frac{Q^2}{\mu'^2}\right)\right] + O\left(a_s^3\right)\Bigg\}.\end{aligned} \tag{3.33}$$

At this point we can carry out the integration through the formula:

$$\begin{aligned}\int_{\mu_b^2}^{Q^2} \frac{d\mu'^2}{\mu'^2} \ln^k \left(\frac{Q^2}{\mu'^2}\right) &= \int_{\ln(\mu_b^2)}^{\ln Q^2} d\ln\mu'^2 \ln^k \left(\frac{Q^2}{\mu'^2}\right) \\ = \int_0^{\ln(Q^2/\mu_b^2)} dx x^k &= \frac{1}{k+1} \ln^{k+1} \left(\frac{Q^2}{\mu_b^2}\right), \qquad (k \geq 0)\end{aligned} \tag{3.34}$$

obtaining

$$\begin{aligned}\mathrm{E}^2 \simeq \exp\Bigg\{& a_s(Q) \left[\left(K^{(1)} + \gamma_f^{(1)}\right) L - \frac{1}{2}\gamma_K^{(1)} L^2\right] \\ &+ a_s^2(Q) \left[\left(K^{(2)} + \gamma_f^{(2)}\right) L + \left(\beta_0 K^{(1)} + \frac{1}{2}\beta_0\gamma_f^{(1)} - \frac{1}{2}\gamma_K^{(2)}\right) L^2 \right. \\ &\left. - \frac{1}{3}\beta_0\gamma_K^{(1)} L^3\right]\Bigg\},\end{aligned} \tag{3.35}$$

where

$$L \equiv \ln \frac{Q^2}{\mu_b^2}. \tag{3.36}$$

It is useful to introduce a shorthand notation:

$$\mathrm{E}^2 \simeq \exp\left\{\sum_{n=1}^{2} a_s^n(Q) \sum_{k=1}^{n+1} E^{(n,k)} L^k\right\} \tag{3.37}$$

where $E^{(n,k)}$ are the coefficients of $a_s^n(Q)L^k$ in (3.35):

$$\begin{aligned}&E^{(1,1)} = \tilde{K}^{(1)} + \gamma_f^{(1)}, \quad E^{(1,2)} = -\frac{1}{2}\gamma_K^{(1)} \\ &E^{(2,1)} = K^{(2)} + \gamma_f^{(2)}, \quad E^{(2,2)} = \beta_0 K^{(1)} + \frac{1}{2}\beta_0\gamma_f^{(1)} - \frac{1}{2}\gamma_K^{(2)}, \quad E^{(2,3)} = -\frac{1}{3}\beta_0\gamma_K^{(1)}.\end{aligned} \tag{3.38}$$

We can now expand the exponential to order a_s^2:

$$\begin{aligned}\mathrm{E}^2 \simeq& 1 + a_s(Q) \sum_{k=1}^{2} E^{(1,k)} L^k + a_s^2(Q) \left[\sum_{k=1}^{3} E^{(2,k)} L^k + \frac{1}{2}\left(\sum_{k=1}^{2} E^{(1,k)} L^k\right)^2\right] \\ &+ O\left(a_s^3\right) \\ \simeq& 1 + a_s(Q) \sum_{k=1}^{2} E^{(1,k)} L^k + a_s^2(Q) \sum_{k=1}^{4} \widetilde{E}^{(2,k)} L^k + O\left(a_s^3\right),\end{aligned} \tag{3.39}$$

where

$$\begin{aligned} \widetilde{E}^{(2,1)} &= E^{(2,1)}, & \widetilde{E}^{(2,2)} &= E^{(2,2)} + \tfrac{1}{2}\left[E^{(1,1)}\right]^2 \\ \widetilde{E}^{(2,3)} &= E^{(2,3)} + E^{(1,1)}E^{(1,2)}, & \widetilde{E}^{(2,4)} &= \tfrac{1}{2}\left[E^{(1,2)}\right]^2 \end{aligned} \quad . \tag{3.40}$$

By induction, one can conclude that the Sudakov factor has the logarithmic expansion in (3.21), and that in order to achieve a LL, NLL, NNLL logarithmic resummation, the perturbative expansions should be taken as in Table 3.1.

3B Appendix: hard factor and Wilson coefficients at $O(\alpha_s)$

In this Appendix we collect the $O(\alpha_s)$ expressions for the hard factors $\mathcal{H}_X$ in Eqs. (3.5)-(3.6) and the Wilson coefficients $\tilde{C}^{in(out)}$ in Eq (3.20), since they will be used in Ch. 4, and there are different conventions for their definition. For their derivation, we refer the reader to [26, 64]. Here we follow the convention of those references, apart from the Wilson coefficient $\tilde{C}^{out}$, for which

$$\tilde{C}^{out}\Big|_{\text{ours}}(z, b; \ldots) = z^2 \tilde{C}^{out}\Big|_{\text{Collins}}(z, b; \ldots). \tag{3.41}$$

Hard factors

$$\begin{aligned} \left|\mathcal{H}_{SIDIS}(Q;\mu)^2\right| &= 1 + \frac{\alpha_s C_F}{\pi}\left[-\frac{3}{2}\ln\frac{\mu^2}{Q^2} - \frac{1}{2}\ln^2\frac{\mu^2}{Q^2} - 4\right] + O\left(\alpha_s^2\right) \\ \left|\mathcal{H}_{DY}(Q;\mu)^2\right| &= 1 + \frac{\alpha_s C_F}{\pi}\left[-\frac{3}{2}\ln\frac{\mu^2}{Q^2} - \frac{1}{2}\ln^2\frac{\mu^2}{Q^2} - 4 + \pi^2\right] + O\left(\alpha_s^2\right). \end{aligned} \tag{3.42}$$

TMD PDF Wilson coefficients

$$\begin{aligned} &\tilde{C}^{in}_{q/q'}(x, b; \mu, \zeta, \alpha_s(\mu)) = \delta_{qq'}\delta(1-x) + \\ &+ \delta_{qq'}\frac{\alpha_s C_{\text{F}}}{2\pi}\left\{ \ln\left(\frac{2e^{-\gamma_{\text{E}}}}{b\,\mu}\right)^2 \left[\left(\frac{2}{1-x}\right)_+ - 1 - x\right] + 1 - x + \right. \\ &\left. + \delta(1-x)\ln\left(\frac{2e^{-\gamma_{\text{E}}}}{b\,\mu}\right)^2\left[1 + \ln\left(\frac{\zeta}{\mu^2}\right)\right]\right\} + O\left(\alpha_s^2\right), \end{aligned} \tag{3.43}$$

$$\begin{aligned} &\tilde{C}^{in}_{q/g}(x, b; \mu, \zeta, \alpha_s(\mu)) = \frac{\alpha_s T_{\text{f}}}{2\pi}\left\{ \ln\left(\frac{2e^{-\gamma_{\text{E}}}}{b\,\mu}\right)^2 [1 - 2x(1-x)] + \right. \\ &\left. + 2x(1-x)\right\} + O\left(\alpha_s^2\right), \end{aligned} \tag{3.44}$$

TMD FF Wilson coefficients

$$\begin{aligned}
&\tilde{C}^{out}_{q'/q}(z, b; \mu, \zeta, \alpha_s(\mu)) = \delta_{qq'}\delta(1-z)+\\
&+\delta_{qq'}\frac{\alpha_s C_\mathrm{F}}{2\pi}\left\{\left[\ln\left(\frac{2e^{-\gamma_\mathrm{E}}}{b\,\mu}\right)^2+\ln z^2\right]\left[\left(\frac{2}{1-z}\right)_+ +1+z\right]+1-z+\right.\\
&\left.+\delta(1-z)\ln\left(\frac{2e^{-\gamma_\mathrm{E}}}{b\,\mu}\right)^2\left[1+\ln\left(\frac{\zeta}{\mu^2}\right)\right]\right\}+\mathcal{O}\left(\alpha_s^2\right),
\end{aligned} \tag{3.45}$$

$$\begin{aligned}
&\tilde{C}^{out}_{g/q}(z, b; \mu, \zeta, \alpha_s(\mu)) =\\
&\frac{\alpha_s C_\mathrm{F}}{2\pi}\frac{1}{z}\left\{\left[\ln\left(\frac{2e^{-\gamma_\mathrm{E}}}{b\,\mu}\right)^2+\ln z^2\right]\left[1+(1-z)^2\right]+z^2\right\}+\mathcal{O}\left(\alpha_s^2\right).
\end{aligned} \tag{3.46}$$

Chapter 4

Low transverse momentum: phenomenology

In this Chapter we present phenomenological studies on TMDs in SIDIS and Drell-Yan. The two processes are treated separately, since SIDIS data present issues when attempting a description beyond leading order. These are addressed in Sec. 4.1, where we provide theoretical and numerical evidences that help to identify the origin of the discrepancies. In Sec. 4.2 we show preliminary results for a fit of Drell-Yan data from both fixed-target and collider experiments, including the newest data from LHC.

4.1 TMDs in SIDIS

Early phenomenological studies on the CSS formalism in SIDIS can be found in [22], where data from the H1 [79] and ZEUS [80] collaborations at HERA were considered. The observables in that case are however different from the cross section in (3.1). In one case the transverse energy flow was measured, which is equivalent to

$$\frac{\mathrm{d}\Sigma_z}{\mathrm{d}x\mathrm{d}Q^2\mathrm{d}q_T} = \sum_h \int_{z_{min}}^1 \mathrm{d}z \left(z \cdot \frac{\mathrm{d}\sigma^{\ell\, N \to \ell'\, h\, X}}{\mathrm{d}x\mathrm{d}z\mathrm{d}Q^2\mathrm{d}q_T} \right), \tag{4.1}$$

and hence a sum over hadron species and an average over z of the cross section are involved. In the other case, data differential in P_{hT} and integrated over the other variables were presented:

$$\frac{\mathrm{d}\sigma^{\ell\, N \to \ell'\, h\, X}}{\mathrm{d}P_{hT}}. \tag{4.2}$$

The fact that the cross section in (4.2) is differential in P_{hT} and integrated over z poses complications in the theoretical analysis of data, since the low and high q_T ranges are mixed. Leaving to future studies a reanalysis of these data, in this work we will discuss the unpolarized P_{hT}-differential multiplicities in SIDIS, integrated over the azimuthal angle. To distinguish from collinear multiplicities, Eq. (2.27), we will denote them as $M^{\mathrm{h}}_{P_T}$:

$$M^{\mathrm{h}}_{P_T}\left(x, z, Q^2, P_{hT}\right) = \frac{\mathrm{d}\sigma^{\mathrm{h}}/\left(\mathrm{d}x\,\mathrm{d}Q^2\,\mathrm{d}z\,\mathrm{d}P_{hT}\right)}{\mathrm{d}\sigma^{\mathrm{DIS}}/\left(\mathrm{d}x\mathrm{d}Q^2\right)}. \tag{4.3}$$

Different phenomenological TMD studies have been performed on SIDIS multiplicites of the form (4.3), following the release of large statystics data from HERMES [81] and COMPASS [82] collaborations. Some of these studies are in a parton model framework (like, e.g., [23, 24]), while others (e.g., [25]) are at most at NLL, in the nomenclature of Table 3.1. Valuable information was provided by these pioneering studies. However, higher order analyses were not performed, and experimental issues were present: the COMPASS data set, which was the set with the higher statistics and kinematic coverage, was affected by normalization errors[1]. This allowed to reproduce the shape of the data but not their normalization. A new set of data by the same collaboration has been released only recently [83].

In this thesis, we will show that going at the first non-trivial order ($\mathcal{O}(\alpha_s)$) in the TMD formula apparently introduces a serious discrepancy with data, lowering theoretical predictions by a factor ~ 0.5. We will then argue that this discrepancy is probably due to the fact that most data lie in a q_T region where the formalism is not applicable, at least without modifications. At fixed-target energies, it is hard if not impossible to identify a region of validity of the TMD formalism. Considerations based on perturbative calculations indicate that P_{hT} should be extremely small if we want q_T/Q logarithms to dominate. However, nonperturbative contributions should be large in that region and hamper a perturbative analysis.

4.1.1 The range for transverse-momentum resummation

The issue of what is practically the q_T range for TMDs is quite controversial. In all Drell-Yan studies (e.g., [19, 20, 25, 29, 30]), the chosen range is $q_T/Q \lesssim 0.1-0.25$ (both for fixed target and collider regimes), while restrictions adopted in SIDIS studies are definitely loose: the focus is more on the value of P_{hT}, so that data with q_T close to Q are often accepted as TMD data (see, e.g., [23–25]). This choice was in part dictated by the gaussian-like shape of the data, which seems to suggest that intrinsic transverse momentum is playing a role in SIDIS even sometimes

---

[1] See the *erratum* to [82]

at $q_T \sim Q$, especially when Q is few GeVs. We cannot discard this hypothesis. However, it must be noticed that the TMD formalism is beyond its domain of applicability in that region, and hence it cannot be expected to give meaningful results: as we have seen, TMD factorization is equivalent to transverse-momentum resummation, and hence is applicable in the q_T range where resummation is meaningful. A commonly adopted rough criterion is $q_T \ll Q$, since this means large logarithms of the form $\ln(Q/q_T)$. A more refined criterion can be to look at the asymptotic cross section. This contains the parts of the fixed-order cross section diverging at least as $1/q_T^2$ (Eq. (2.54)), which are precisely those terms that, order by order, need to be resummed. Resummation should be applied where the asymptotic term is a good approximation of the fixed-order calculation. Moreover, going to larger q_T, the asymptotic curve rapidly goes to negative values. At this point, not only a resummation formula, but also a W+Y construction would be affected by large errors (see again Sec. 1.4 of [71]). To give an example of what can be the effect of such a criterion on data selection, in Fig. 4.1 we plot the fixed order collinear-factorized prediction and its asymptotic expansion (at LO), for a Drell-Yan and a SIDIS bin, with comparable values of $\sqrt{s}$, Q ($\simeq$ 4.5 GeV) and x. The region where the asymptotic term is positive, and hence meaningful, is $q_T \lesssim 1.4$ GeV in both cases. Beyond this region it is in principle hard to justify the use of resummed formulas. However, the range covered by data is markedly different: at variance with Drell-Yan, almost all data in the considered SIDIS bin lie beyond this 'safe zone'[2]. It follows, then, that SIDIS studies for kinematics similar to Fig. 4.1 necessarily have to violate the above criterion. It is interesting to notice that, instead, selection cuts adopted in Drell-Yan studies generally agree with the criterion[3].

4.1.2 The integral of TMDs

From the above considerations, it is important to understand what one should expect when adopting a TMD formalism to attempt a description of SIDIS data. To this aim, we will consider the integral over $\mathbf{q}_T$ of the TMD-factorized cross section

$$\int \mathrm{W} \equiv \int \mathrm{d}^2\mathbf{q}_T \frac{\mathrm{d}\sigma^{\ell N \to \ell' h X}}{\mathrm{d}Q^2 \mathrm{d}x \mathrm{d}z \mathrm{d}^2\mathbf{q}_T}(\mathrm{TMD}), \tag{4.4}$$

---

[2] One might be induced to think that this is specific of the smaller z bins, since $q_T = P_{hT}/z$, and hence this bins cover higher values in q_T. However, as we shall see in the following, the observations presented here are quite general for low-energy SIDIS, and our conclusions do not have a marked dependence on the value of z.

[3] The cut $q_T/Q \lesssim 0.1 - 0.25$ in general selects data within our 'safe zone'. For instance, the analyses in [19, 20, 25] considered the first 7 data points from the E288 bin in Fig. 4.1, while [29] considered the first 5 points. A complete collection of plots like Fig. 4.1 for fixed-target Drell-Yan can be found in Ch. 5.

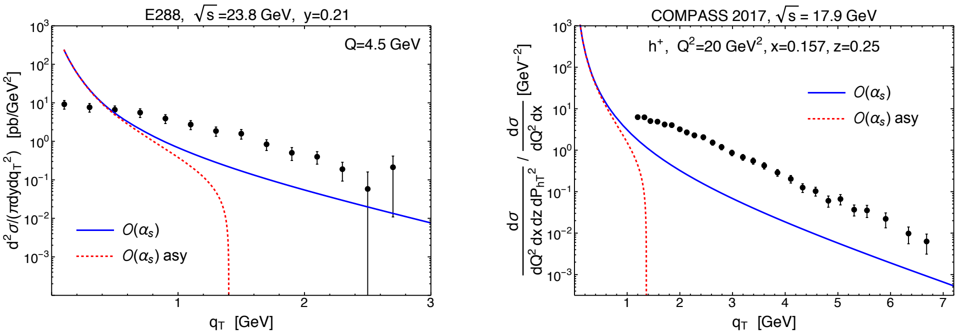


Figure 4.1: The fixed order (full curve) and asymptotic (dashed curve) $O(\alpha_s)$ calculations for a Drell-Yan (left) and a SIDIS (right) bin. The range where the two theoretical curves overlap should be considered as the range where transverse-momentum resummation is needed. To show the available experimental ranges, data from E288 [13] and COMPASS [83] are superimposed.

where (TMD) denotes that formula (3.5) is adopted for the q_T-differential cross section. By the properties of Fourier transforms, the above intergal can be written in terms of $b-$space distributions, evaluated at $b \equiv 0$:

$$\begin{aligned}\int \mathrm{W} = \sigma_0 \sum_q \left|\mathcal{H}_{SIDIS}(Q;Q)^2\right| \\ \times \tilde{f}_1^{q/N}\left(x, b \equiv 0; Q; Q^2\right) \tilde{D}_1^{h/q}\left(z, b \equiv 0; Q; Q^2\right),\end{aligned} \tag{4.5}$$

where σ_0 is the cross section prefactor

$$\sigma_0 \equiv \frac{z^2 4\pi\alpha_{em}^2}{xQ^4}\left(1 - y + \frac{y^2}{2}\right). \tag{4.6}$$

Comparing the value of this integral to the integral of experimental data, or to the theoretical integrated cross section in collinear factorization, Eq. (2.29), will explain why a $O(\alpha_s^0)$ formula (as the one in [25]) can approach data, and higher order calculations cannot.

Analytic expressions for the integral of TMDs

In the standard CSS prescription, the TMD integral over the full q_T space is zero.[4] However, as we mentioned in Sec. 3.1.3, this is due to the scale μ_b^* in (3.23) going unphysically to infinity for $b \to 0$. In q_T-space, this is reflected on the TMD going large and negative at large q_T, compensating the integral of

---

[4] See Appendix A of [73] for an explicit proof.

the small-q_T region. This is an incidental feature of the prescription used for the b^* prescription, as discussed in Sec. 3.1.3. Prescriptions of the $\bar{b}^*$ type, like (3.26), while not affecting the TMD in the physical region, modify the large-q_T tail, yielding a non-zero value for $\int$ W. Indeed, these prescriptions give $\bar{b}^* \longrightarrow b_{min}$ when $b \longrightarrow 0$, with $b_{min} \sim 1/Q$. Inserting our solution (3.23) for TMD evolution, the integral in Eq. (4.5) reads:

$$\begin{aligned}\int \mathrm{W} = \sigma_0 \sum_q \left|\mathcal{H}_{SIDIS}(Q;Q)^2\right| \left(\frac{Q}{\mu_{b_{min}}}\right)^{2\tilde{K}\left(b\equiv 0;\mu_{b_{min}}\right)} \\ \times \exp\left\{2\int_{\mu_{b_{min}}}^{Q} \frac{d\mu'}{\mu'}\left[\gamma_f\left(\alpha_s\left(\mu'\right);1\right) - \ln\frac{Q}{\mu'}\gamma_K\left(\alpha_s\left(\mu'\right)\right)\right]\right\} \\ \times \tilde{f}_{1\,OPE}\left(x, b_{min}; \mu_{b_{min}}, \mu^2_{b_{min}}\right) \tilde{f}_{NP}(x, b\equiv 0, Q^2) \\ \times \tilde{D}_{1\,OPE}\left(z, b_{min}; \mu_{b_{min}}, \mu^2_{b_{min}}\right) \tilde{D}_{NP}(z, b\equiv 0, Q^2),\end{aligned} \tag{4.7}$$

where we chose the hard scales as $\mu = Q$, $\zeta_{f(D)} = Q^2$, and the initial scales as $\mu_0 = \mu_{b^*}$, $\zeta_{0f(D)} = \mu^2_{b^*}$. In the above equation, we have substituted b^* with b_{min}, i.e., its limit for $b \longrightarrow 0$. A reasonable choice for b_{min} is the one that gives

$$\mu_{b_{min}} = Q, \tag{4.8}$$

i.e., $b_{min} = b_0/Q$, since this prevents unphysical Sudakov resummation (i.e., resummation with $\mu_b > Q$). With this choice, perturbative evolution terms are set to 1 in (4.7). By definition, also the nonperturbative terms $\tilde{f}_{NP}$ and $\tilde{D}_{NP}$ must be equal to 1 at $b = 0$ (they should not affect formulas in the perturbative regime $b \longrightarrow 0$). In conclusion, we are left with

$$\begin{aligned}\int \mathrm{W} = \sigma_0 \sum_q \left|\mathcal{H}_{SIDIS}(Q;Q)^2\right| \\ \times \tilde{f}_{1\,OPE}\left(x, b_0/Q; Q, Q^2\right) \tilde{D}_{1\,OPE}\left(z, b_0/Q; Q, Q^2\right),\end{aligned} \tag{4.9}$$

where $\tilde{f}_{1\,OPE}$ and $\tilde{D}_{1\,OPE}$ are convolutions of Wilson coefficients $\tilde{C}^{in}$ and $\tilde{C}^{out}$ with collinear PDFs and FFs, respectively (Eq. (3.20)). When the hard factor and the Wilson coefficients are taken at $O(\alpha_s^0)$, as in a LL or NLL framework, the explicit expression becomes

$$\int \mathrm{W}\Big|_{O(\alpha_s^0)} = \sigma_0 \sum_q \frac{e_q^2}{z^2} f_1^{q/N}(x,Q) D_1^{h/q}(z,Q). \tag{4.10}$$

This corresponds to the leading-order expression for the integrated cross section in collinear factorization, (2.29):

$$\int \mathrm{W}\bigg|_{O(\alpha_s^0)} = \frac{\mathrm{d}\sigma^h}{\mathrm{d}x\mathrm{d}Q^2\mathrm{d}z}\bigg|_{O(\alpha_s^0)} . \tag{4.11}$$

Expressions beyond zeroth order for $\mathcal{H}$ and $\tilde{C}^{in(out)}$ contribute in a NLL' calculation, or at higher logarithmic accuracy. Isolating only the terms proportional to $O(\alpha_s)$ in the TMD integral, with the standard scale choices of (4.9), one obtains

$$\begin{aligned}\int \mathrm{W}\bigg|_{O(\alpha_s^1)} = \sigma_0 \sum_q \frac{e_q^2}{z^2}\frac{\alpha_s}{\pi}\bigg\{ & -4C_F f_1^{q/N}(x,Q) D_1^{h/q}(z,Q) \\ & + \left[\tilde{C}_{(1)}^{in}(b=0)\otimes f_1\right]^{q/N}(x,Q)\, D_1^{h/q}(z,Q) \\ & + f_1^{q/N}(x,Q)\, \left[\tilde{C}_{(1)}^{out}(b=0)\otimes D_1\right]^{h/q}(z,Q)\bigg\},\end{aligned} \tag{4.12}$$

where we have explicitly inserted the expression (3.42) for the hard factor. The convolutions in the second and third line are

$$\begin{aligned} & \left[\tilde{C}_{(1)}^{in}(b=0)\otimes f_1\right]^{q/N}(x,Q) = \\ & \sum_{i=q,g}\int_x^1 \frac{\mathrm{d}\hat{x}}{\hat{x}}\tilde{C}_{q/i}^{in(1)}\left(x/\hat{x}, b_{min}; Q, Q^2, \alpha_s(Q)\right) f_1^{i/N}(\hat{x},Q), \\ & \left[\tilde{C}_{(1)}^{out}(b=0)\otimes D_1\right]^{h/q}(z,Q) = \\ & \sum_{i=q,g}\int_x^1 \frac{\mathrm{d}\hat{z}}{\hat{z}}\tilde{C}_{i/q}^{out(1)}\left(z/\hat{z}, b_{min}; Q, Q^2, \alpha_s(Q)\right) D_1^{h/i}(\hat{z},Q), \end{aligned} \tag{4.13}$$

and the expressions for the order $O(\alpha_s)$ part of Wilson coefficients in (4.13) can be derived from Eqs. (3.43) to (3.46):

$$\tilde{C}_{q/q'}^{in(1)}\left(x, b_{min}; Q, Q^2, \alpha_s(Q)\right) = \delta_{qq'}\frac{\alpha_s C_\mathrm{F}}{2\pi}(1-x), \tag{4.14}$$

$$\tilde{C}_{q/g}^{in}\left(x, b_{min}; Q, Q^2, \alpha_s(Q)\right) = \frac{\alpha_s T_\mathrm{f}}{2\pi}\left[2x(1-x)\right], \tag{4.15}$$

$$\begin{aligned}\tilde{C}_{q'/q}^{out(1)}\left(z, b_{min}; Q, Q^2, \alpha_s(Q)\right) =\delta_{qq'}\frac{\alpha_s C_\mathrm{F}}{2\pi}\bigg\{ & \ln z^2\left[\left(\frac{2}{1-z}\right)_+ + 1 + z\right] \\ & + 1 - z\bigg\},\end{aligned} \tag{4.16}$$

$$\tilde{C}^{out}_{g/q}\left(z, b_{min}; Q, Q^2, \alpha_s(Q)\right) = \frac{\alpha_s C_{\rm F}}{2\pi}\frac{1}{z}\left\{ \ln z^2 \left[1 + (1-z)^2\right] + z^2 \right\}. \tag{4.17}$$

If we compare the $O(\alpha_s)$ TMD integral, Eq. (4.12), to the $O(\alpha_s)$ expression for the integrated SIDIS cross section in collinear factorization (Appendix 4A), we see that the two expressions are different, and $\int$ W reproduces only one part of the coefficients of the full cross section:

$$\int \mathrm{W}\bigg|_{O(\alpha_s^1)} \neq \frac{\mathrm{d}\sigma^h}{\mathrm{d}x\mathrm{d}Q^2\mathrm{d}z}\bigg|_{O(\alpha_s^1)}. \tag{4.18}$$

This is consistent: integrating only the low-q_T contributions, contained in the TMD formula, should reproduce only one part of the full integrated cross section. The only exception is the order $O(\alpha_s^0)$ case, as we have seen above. The reason is that at order 0 in collinear factorization the only contribution is from $q_T = 0$, since there is no gluon radiation to produce a recoil transverse momentum.

Relation between the TMD integral and the asymptotic term

To better understand what is the relation between the TMD integral and collinear factorization, it is useful to consider the integral of the asymptotic cross section

$$\int \mathrm{d}^2\mathbf{q}_T \frac{\mathrm{d}\sigma^{\ell N \to \ell' h X}}{\mathrm{d}Q^2\mathrm{d}x\mathrm{d}z\mathrm{d}^2\mathbf{q}_T}(\text{asy}), \tag{4.19}$$

where the leading order expression for the integrand is given in (2.53).
To be more precise, Eq. (4.19) as it stands is ill-defined, since the asymptotic function is not integrable over q_T, due to collinear and soft singularities which arise as $q_T \longrightarrow 0$. To have a well-defined integral, one should consider the partonic version of (4.19)

$$\int \widehat{\text{asy}} \equiv \int \mathrm{d}^2\mathbf{q}_T \frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}Q^2\mathrm{d}\hat{x}\mathrm{d}\hat{z}\mathrm{d}^2\mathbf{q}_T}(\text{asy}), \tag{4.20}$$

and then convolute it with hadronic PDFs and FFs. After inclusion of proper virtual contributions at $q_T = 0$ and factorization of collinear singularities, the resulting cross section integrated over q_T is finite. This is what we will refer to as "integral of the asymptotic term":

$$\begin{aligned}\int \text{asy} \equiv &\sum_{i,j} \int_x^1 \frac{\mathrm{d}\hat{x}}{\hat{x}} f_1^{i/N}\left(\frac{x}{\hat{x}}, \mu\right) \int_z^1 \frac{\mathrm{d}\hat{z}}{\hat{z}} D_1^{h/j}\left(\frac{z}{\hat{z}}, \mu\right) \\ &\times \int \mathrm{d}^2\mathbf{q}_T \frac{\mathrm{d}\hat{\sigma}^{ji}}{\mathrm{d}Q^2\mathrm{d}\hat{x}\mathrm{d}\hat{z}\mathrm{d}^2\mathbf{q}_T}(\text{asy}).\end{aligned} \tag{4.21}$$

By sending the space-time dimensions to $d = 4 - 2\epsilon$, and hence the transverse dimensions to $2 - 2\epsilon$, divergences in the integral (4.20) will appear as poles in ϵ. Soft divergences are then canceled by virtual contributions, while collinear divergences are factorized into partonic PDFs and FFs. At this point an analytic solution for (4.21) can be found, provided one introduces an upper integration cutoff q_T^{cut}. This is needed since the asymptotic term is only meaningful at low q_T, where it approximates the fixed-order. The result of such a calculation can be found in [22]:

$$\begin{aligned}\int \text{asy}\Big|_{O(\alpha_s^1)} = \int \text{W}\Big|_{O(\alpha_s^1)} &+ \\ &\sum_q \frac{\sigma_0\alpha_s}{z^2\pi} e_q^2 \Bigg\{ -\frac{C_F}{2}\left(\ln^2\frac{Q^2}{\left(q_T^{cut}\right)^2} - \ln\frac{Q^2}{\left(q_T^{cut}\right)^2}\right)\times \\ &\times f_1^{q/N}(x,Q)D_1^{h/q}(z,Q)+ \\ &+\ln\frac{Q^2}{\left(q_T^{cut}\right)^2}\sum_q\Bigg[D_1^{h/q}(z,Q)\sum_{i=q,g}\int_x^1 \frac{\mathrm{d}\hat{x}}{\hat{x}}P_{qi}(\hat{x})f_1^{i/N}\left(\frac{x}{\hat{x}},Q\right)+ \\ &+ f_1^{q/N}(x,Q)\sum_{j=q,g}\int_z^1 \frac{\mathrm{d}\hat{z}}{\hat{z}}P_{jq}(\hat{z})D_1^{h/j}\left(\frac{z}{\hat{z}},\mu\right)\Bigg]\Bigg\},\end{aligned} \tag{4.22}$$

where P_{ij} are the splitting functions, and we have set the factorization and renormalization scales equal to Q. Thus, for $q_T^{cut} \simeq Q$, the TMD integral at order α_s corresponds to the integral of the asymptotic term at the same order. This relation is expected to hold also at higher orders.

Concluding, we have evidenced that:

1. the TMD cross section is supposed to yield a reliable prediction in the region where the fixed-order collinear cross section is approximated by the asymptotic expansion;

2. the integral of the TMD cross section can be calculated analytically, provided a $\bar{b}^*$ prescription is used, to avoid unphysical use of resummation; as we shall see in Sec. 4.1.3, the analytic integral thus calculated is a good approximation of the numerical integral of the W term, restricted to the region W> 0; its value is independent of non-perturbative terms, and thus predictable from pQCD;

3. at order $O(\alpha_s)$ in the hard factor and Wilson coefficients, the analytic integral above corresponds to the integral of the asymptotic term (after cancellation

of singularities); this is consistent with point 1: the W term should reproduce the fraction of data in the region where the asymptotic approximation works;

4. at order $O(\alpha_s^0)$, the integral of the W term corresponds to the full LO integrated cross section; this is due to the fact that, at this order, the q_T generated by perturbative QCD is 0, and hence there is no high-q_T contribution to the theoretical cross-section.

4.1.3 Comparison to data

For the reasons explained in the previous Sections, it is possible for TMD theory–when taken at the 0th order in the hard coefficients–to approach fixed-target SIDIS data, even tough most of them lie outside of what can be safely retained a TMD region. This was partially shown by previous studies like [23–25], which however could not make definitive conclusions concerning the COMPASS data [82], affected by an experimental issue in the determination of the normalization. The new and unbiased data released in [83] confirm the above statement: in Figs. 4.2 and 4.3 we show them in comparison with predictions obtained with the parameters extracted in [25]. Since no hadron identification is provided, we consider the sum of pion and kaon multiplicities. The data shown satisfy the selection cuts $P_{hT} < 0.9$ GeV, $0.2 < z < 0.74$, and $Q^2 > 1.4$ GeV^2. All the perturbative and nonperturbative ingredients are the same as in [25]. The overall χ^2/N_{data} from the comparison, including both positive and negative hadrons in the final state, is ~ 2.4, and goes down to ~ 1.7 when excluding the two lowest-Q^2 bins (i.e., selecting data with $Q^2 > 2.1$ GeV^2).

However, in agreement with the discussions in Sec. 4.1.2, when going from a NLL prescription–as the one used in [25]–to a NLL′ one (or higher order), where the hard factor and the Wilson coefficients are at order $O(\alpha_s)$, the comparison to data becomes extremely difficult. Indeed, one finds that the ratio data/theory is around ~ 2 for most bins. This is independent of nonperturbative choices, as the discrepancy does not improve when letting the parameters being chosen by a fit to data. In view of what we saw above, the explanation is rather simple: in a 0th-order prescription, the TMD cross section is such to reproduce the *full* LO cross section when integrated over q_T (Eq. (4.11)). Thus, NLL TMDs (as well as LL and parton model TMDs) are normalized to a value that approximates the *whole* q_T-integrated data, while NLL′ TMDs are normalized to the q_T-integral of just the TMD region (hence, an extremely narrow region, as shown in Fig. 4.1). To quantitatively see this effect, we took again the TMD parameterization from [25][5], but this time adding the $O(\alpha_s)$ terms for the hard factor and Wilson coefficients,

---

[5] With respect to [25], we used more recent sets for the collinear PDFs and FFs that enter the TMD formula: we used MMHT PDFs [5], and the DSSH17 set for kaon FFs.

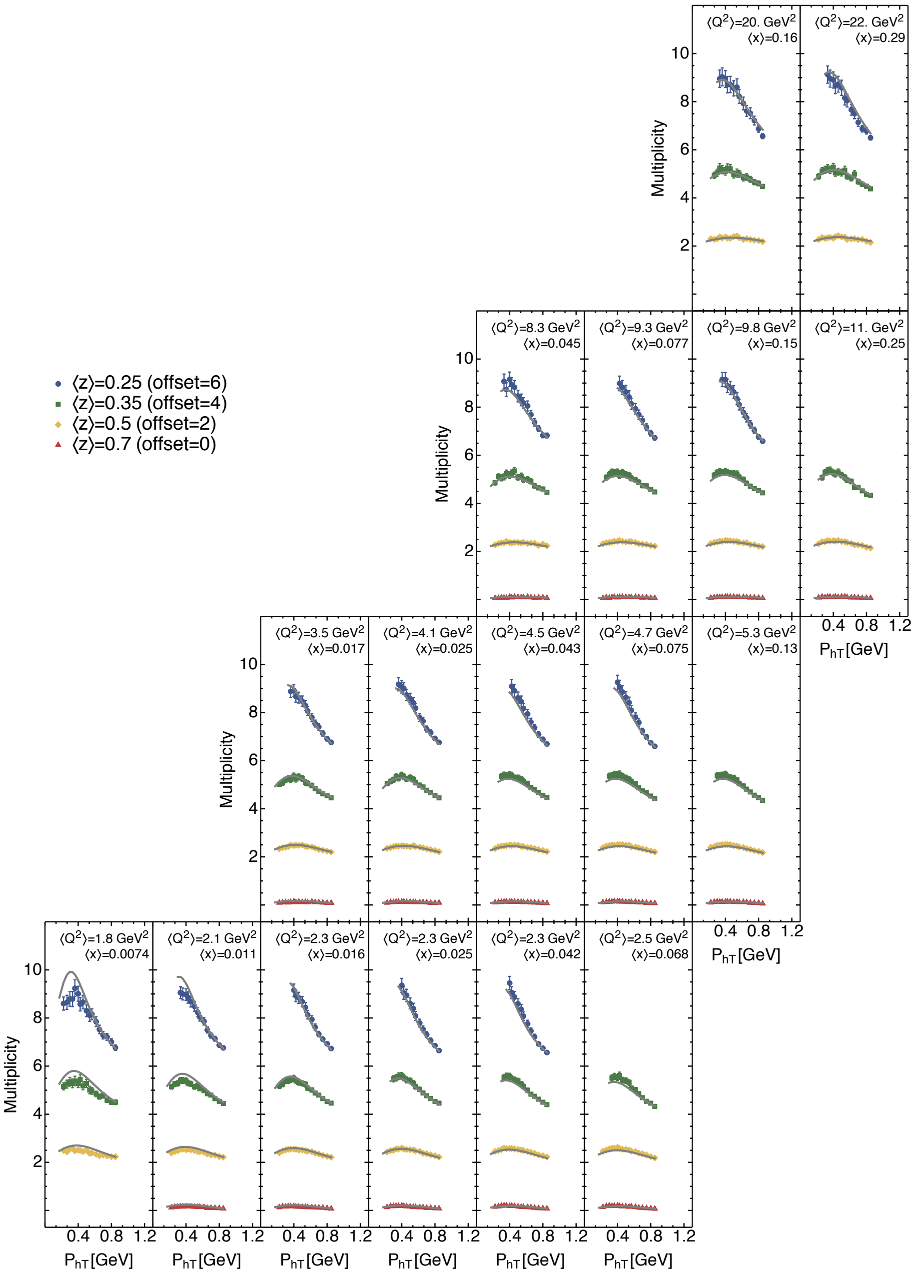


Figure 4.2: Comparison of the new COMPASS data [83] for SIDIS multiplicities, (4.3), to predictions obtained with the TMD functions extracted from other data in [25]. In this figure: negative hadron production off a deuteron target.

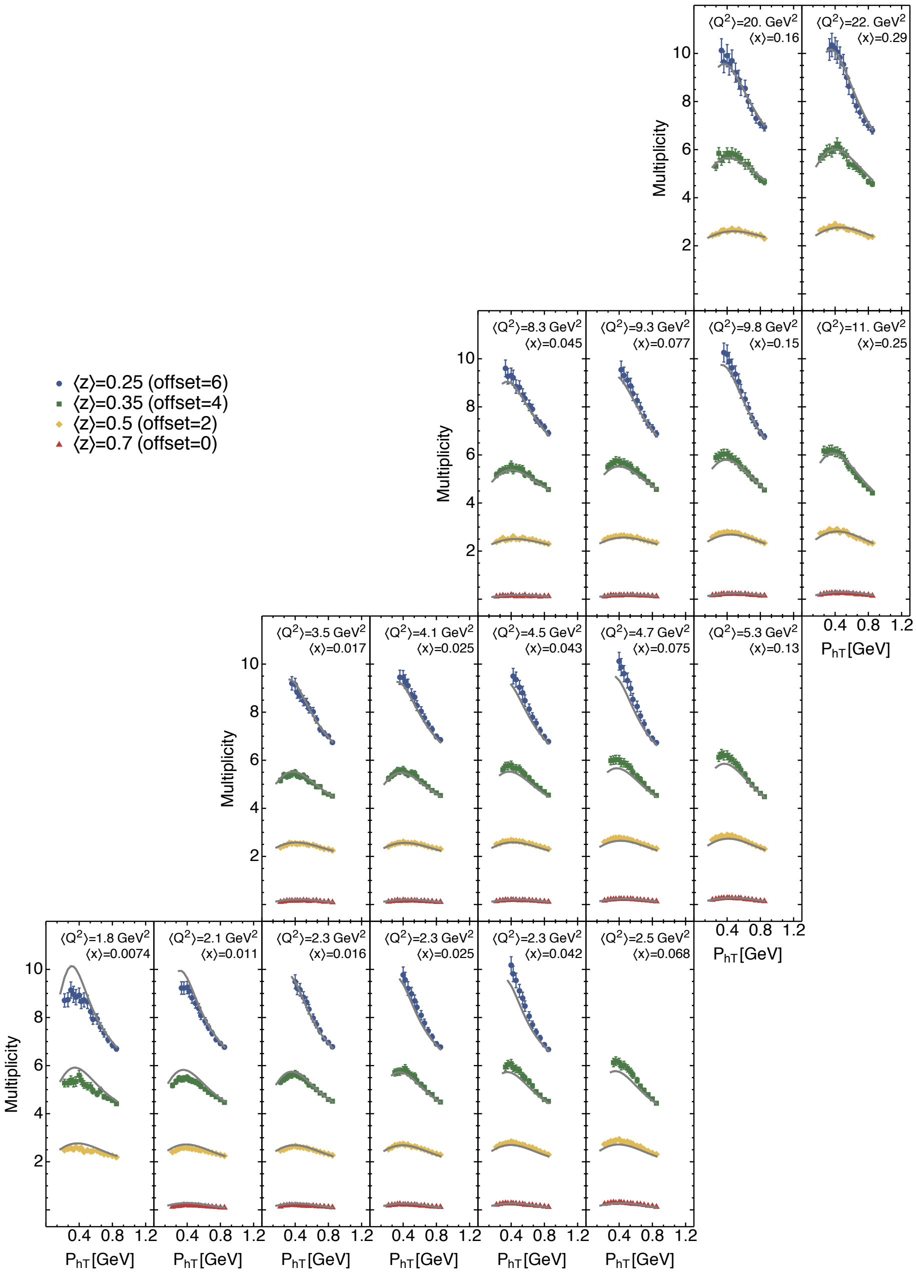


Figure 4.3: The same as Fig. 4.2, but for positive hadron production.

and we multiplied the theoretical predictions by independent normalization factors $N(x, z, Q^2)$, one for each bin in x, z and Q^2. These factors were then considered as free parameters, while all the nonperturbative parameters were kept fixed, and the χ^2 w.r.t. the data was minimized as a function of them[6]

$$\chi^2 \equiv \sum_{x,z,Q^2} \sum_{P_{hT}} \frac{\left[M^{\mathrm{h}}_{P_T}(\mathrm{exp}) - N(x,z,Q^2) M^{\mathrm{h}}_{P_T}(\mathrm{th})\right]^2}{\left(\delta M^{\mathrm{h}}_{P_T}\right)^2}. \tag{4.23}$$

The global $\chi^2/d.o.f.$, considering the number of parameters introduced, is ~ 1.34. The normalizations $N(x, z, Q^2)$ thus obtained should give an accurate estimate of the ratio $data/theory$ for the different bins.

According to the previous Section, these factors should be close to the ratio

$$\frac{\frac{\mathrm{d}\sigma^h}{\mathrm{d}x\mathrm{d}Q^2\mathrm{d}z}}{\int \mathrm{W}\big|_{O(\alpha_s^0)+O(\alpha_s^1)}}, \tag{4.24}$$

where the numerator is the the (q_T-integrated) SIDIS cross section in collinear factorization, while the denominator is the integral of the W term, as defined above. The order α_s^0 expression for $\int$ W was given in (4.10), while the α_s expression is reported in Eqs. (4.12)-(4.17). The order α_s expression for the integrated cross section can be found, for instance, in Appendix C of [84]. We do not expect Eq. (4.24) to reproduce *exactly* the phenomenological normalization factors, simply because the numerator of (4.24) is a good estimate for the integral of data over the *whole* q_T range, while we necessarily have to reduce the range when comparing to TMDs. The selection cuts in our analysis also exclude the high-q_T tail of TMDs, which is instead taken into account in the denominator of (4.24). Nonetheless, Fig. 4.4 shows an overall good agreement between the normalization factors fitted by the above procedure (black dots) to the values predicted by the collinear factorization formula[7] (4.24) (red crosses).

To check that the conclusions are quite general, we also performed an actual fit of, simultaneously, the nonperturbative parameters and the normalization factors, with the same exact parameterization and perturbative framework as above (i.e., at NLL′). To have a good estimate of the uncertainty on the parameters, we adopted a bootstrap procedure, repeating the fit for 200 sets of pseudo-data randomly generated around the measured values. Due to the over-parameterization introduced by

---

[6] $M^{\mathrm{h}}_{P_T}(\mathrm{exp})$ and $M^{\mathrm{h}}_{P_T}(\mathrm{th})$ are the experimental and theoretical multiplicities, respectively.

[7] Both the numerator and denominator of (4.24) have been calculated with APFEL++[85]. To produce the figure, we have used a LO expression for the numerator. A NLO formula does not improve the description.

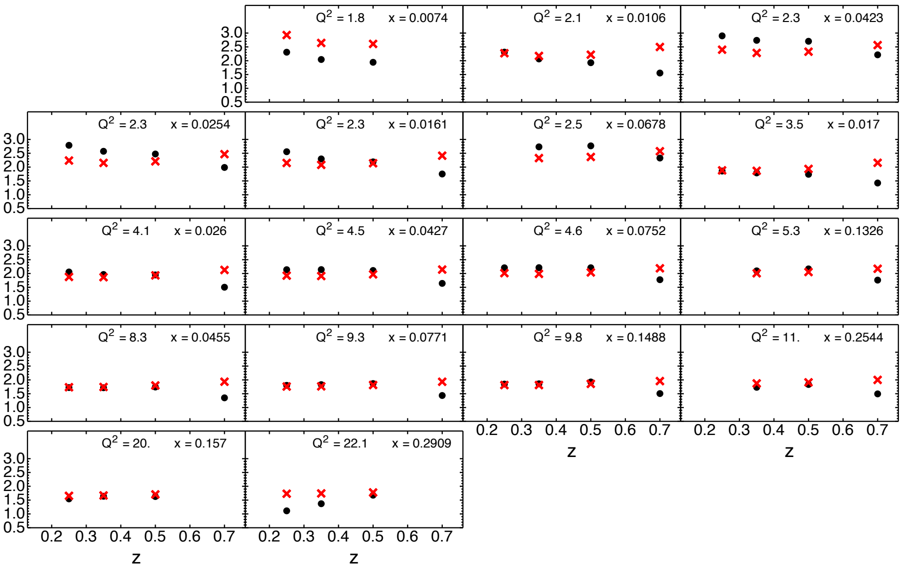


Figure 4.4: The normalization factors $N(x, z, Q^2)$ (black dots), obtained from a comparison to data as described in the text, compared to the theoretical ratio in (4.24) (red crosses), for the different bins of the new COMPASS data [83], in the case of positive hadron production. Here the collinear cross section, numerator of (4.24), is taken at LO.

g_1 $\left[\mathrm{GeV}^2\right]$	α	σ	λ $\left[\mathrm{GeV}^{-2}\right]$	g_2 $\left[\mathrm{GeV}^2\right]$
0.21 ± 0.04	1.56 ± 0.73	-0.12 ± 0.08	0.21 ± 0.21	0.10 ± 0.03

Table 4.1: Non-perturbative parameters for the TMD PDFs, in the notation of Eqs. (34)-(39) of [25], extracted from our toy fit with normalization factors. The reported values correspond to the mean and 68% confidence interval from the results of the 200 bootstrap replicas.

g_3 $\left[\mathrm{GeV}^2\right]$	β	δ
0.40 ± 0.09	4.03 ± 2.06	2.04 ± 0.76
γ	λ_F $\left[\mathrm{GeV}^{-2}\right]$	g_4 $\left[\mathrm{GeV}^2\right]$
0.24 ± 0.15	18.6 ± 1.4	0.18 ± 0.04

Table 4.2: Analogue of Tab. 4.1, but for TMD FFs.

the artificial normalization factors, the final nonperturbative parameters (reported in Tabs. 4.1-4.2) should not be considered as an accurate description of intrinsic transverse momentum. For the same reason, the resulting average $\chi^2/d.o.f.$ is particularly low: 0.51 . This should rather be considered as a toy fit, whose aim is to show that the normalization factors $N(x, z, Q^2)$ are quite general and independent of the nonperturbative parameters. Indeed, from Fig. 4.5, we can see that the agreement with the theoretically predicted normalizations is preserved also in this case.

To conclude, we showed that beyond-lowest-order TMD theory cannot reproduce SIDIS fixed-target data, at least when the data selection cuts in q_T are taken, as often done, beyond the safe region[8] $q_T \lesssim 0.2Q$. However, this discrepancy can be quantitatively predicted by theoretical arguments. For this reason, we are not facing a failure of the theory, but instead a confirmation of the fact that the considered data lie mostly outside of what–on theoretical grounds–can be considered a TMD region. Together with the fact that parton model and lowest-order approaches seem instead to work, with gaussian-like parameterizations, this seems to suggest that intrinsic transverse momentum is playing a major role in the observed data, but at the same time that the TMD factorization formalism may not be the appropriate tool to study this role.

---

[8] During the writing of the thesis, it was reported in [86] a fit of both Drell-Yan and SIDIS, confirming that acceptable results for SIDIS can be achieved only when using restrictive cuts $q_T/Q < 0.2 - 0.25$, thus considering a small percentage of the data.

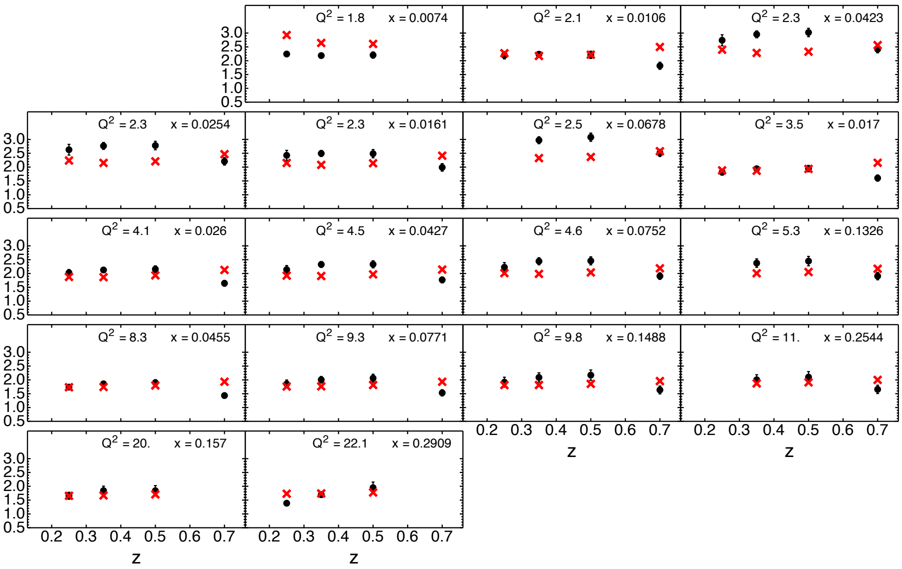


Figure 4.5: Same as Fig. 4.4, but with the normalization factors $N(x, z, Q^2)$ obtained from a simultaneous fit together with the nonperturbative parameters, as described in the text. The black dots and the error bars are obtained respectively as the mean and the standard deviation of the values from the 200 bootstrap replicas.

4.2 TMDs in Drell-Yan

In this Section we report our phenomenological study of TMDs in Drell-Yan[9]. Different studies are present in the literature: after the first analyses of transverse-momentum resummation for Tevatron data in the 1980-90s and early 2000s [17–20], a second wave of updated fits and phenomenological works was published in the last decade, following the release of newer data from Tevatron (e.g., see [74, 87]) and, later, of LHC data. The first TMD fit of simultaneously Drell-Yan and SIDIS data can be found in [25], and was conducted within the "standard" theoretical framework used in the previously-cited works (i.e., the one given in Eq. (3.23)). The first studies including LHC data are [29, 30], which used however a different choice for the TMD renormalization and rapidity scales, called the ζ *prescription* (see [29] for the description), which should allegedly improve perturbative stability. We retain it important to have an independent study of the most recent data in the standard framework, at the highest available perturbative accuracy. Moreover, we will show a TMD study of LHC data without making use of any normalization factor, contrarily to what was done in previous works, where predictions were normalized to the value of the total experimental cross section. This is crucial, since it allows TMD theory to be cleanly tested against the most precise data available, by reproducing not only the shape of transverse momentum cross sections, but also their absolute normalization.

4.2.1 Perturbative and nonperturbative choices

The theoretical framework for our calculations is described in Chapt 3, and condensed in Eq. (3.23). The b^* function is taken as in Eq. (3.26), i.e., with an exponential form and a b_{min} threshold. We set

$$b_{max} = 2e^{-\gamma_E}\ \text{GeV}^{-1} \simeq 1.123\ \text{GeV}^{-1}, \qquad b_{min} = \frac{2e^{-\gamma_E}}{Q}, \tag{4.25}$$

while the TMD factorization and rapidity scales are set to $\mu = \sqrt{\zeta} = Q$. Except for Sec. 4.2.4, where we compare different logarithmic orderings, all our result are at N^{3}LL. For the collinear PDFs, we use the MMHT NNLO set [5]. For what concerns the nonperturbative TMD, according to recent studies [29, 30], a function with a gaussian behavior at $b \to 0$ and a milder slope at large b should allow for an easier description of, simultaneously, low energy data and LHC data. We will thus use a q-Gaussian with q=2, since it has precisely the above properties. Our

---

[9] The term Drell-Yan is extended here to include lepton pair production in hadron-hadron collision via an intermediate Z-boson.

parameterization is:

$$f_{\mathrm{NP}}(x,b,\zeta) = \frac{\exp\left[-\frac{1}{4}g_2 \log\left(\frac{\zeta}{Q_0^2}\right) b^2\right]}{1+\frac{1}{4}g_1^2(x)b^2}, \tag{4.26}$$

where the x-dependence is governed by the function g_1 in the denominator:

$$g_1(x) = N_1 \frac{x^{\sigma}(1-x)^{\alpha}}{\hat{x}^{\sigma}(1-\hat{x})^{\alpha}}. \tag{4.27}$$

We choose $Q_0^2 = 1$ GeV^2 and $\hat{x} = 0.1$. The free parameters are thus four: α, σ, N_1, g_2.

4.2.2 Selected data sets

A description of the data sets selected in our analysis is given in Tab. 4.3. For all the sets, we adopt the selection cut $q_T < 0.2 Q_{min}$, where Q_{min} is the lower bound of the invariant mass bin.

Fixed-target Drell-Yan

We included in our analysis data from Drell-Yan on nuclear targets from Fermilab experiments E288 [13] and E605 [88]. The observable in this case is

$$\frac{E d^3\sigma}{d^3 q} \equiv \frac{2E}{\pi\sqrt{s}} \frac{d\sigma}{dx_F dq_T^2} = \frac{d\sigma}{\pi dy dq_T^2}. \tag{4.28}$$

The kinematic variables of the experiments are reported in Tab. 4.3. The center-of-mass energies range from ~ 20 to ~ 40 GeVs. Nuclear targets (copper and tungsten) are treated as an incoherent sum of protons and nucleons. These data sets are affected by large normalization errors, notably the E288 set, with a declared overall uncertainty of 25%, which we inserted in the correlation matrix as a completely correlated multiplicative uncertainty. We integrated our calculations over the respective bin widths in Q and q_T, and excluded bins with $9 < Q < 11$, to avoid the Υ resonance.

Tevatron

Tevatron data from CDF runs I-II [90, 91], and D0 runs I-II [92, 93] for Z-boson production were also included. The observable in this case is $\mathrm{d}\sigma/\mathrm{d}q_T$, except for D0 run II, where the data are divided by the total measured cross-section: $1/\sigma \cdot \mathrm{d}\sigma/\mathrm{d}q_T$. Our calculations are integrated over the physically allowed rapidity

Experiment	$\sqrt{s}$ (GeV)	y , x_F	Q (GeV)	$\mathrm{N_{pt}}$
E288 200[13]	19.4	$y = 0.4$	4-9	30
E288 300 [13]	23.8	$y = 0.21$	4-9	30
E288 400 [13]	27.4	$y = 0.03$	5-9, 11-14	65
E605 [88]	38.8	$x_F = 0.1$	7-9, 10.5-18	25
PHENIX [89]	200	$y = 1.2 - 2.2$	4.8-8.2	2
CDF run I [90]	$1.8 \cdot 10^3$	-	66-116	25
CDF run II [91]	$1.96 \cdot 10^3$	-	66-116	26
D0 run I [92]	$1.8 \cdot 10^3$	-	75-105	12
D0 run II [93]	$1.96 \cdot 10^3$	-	70-110	5
ATLAS 7 [94]	$7 \cdot 10^3$	$\|y\| < 2.4$ (in 3 bins)	66-116	18
ATLAS 8 [95]	$8 \cdot 10^3$	$\|y\| < 2.4$ (in 6 bins)	66-116	36
CMS 7 [96]	$7 \cdot 10^3$	$\|y\| < 2.1$	60 -120	3
CMS 8 [97]	$8 \cdot 10^3$	$\|y\| < 2.1$	60 -120	4
LHCb 7 [98]	$7 \cdot 10^3$	$2 < y < 4.5$	60-120	7
LHCb 8 [99]	$8 \cdot 10^3$	$2 < y < 4.5$	60-120	7
LHCb 13 [100]	$13 \cdot 10^3$	$2 < y < 4.5$	60-120	7
Total				302

Table 4.3: Data sets considered in our Drell-Yan analysis. The last column reports the number of included data points after the selection cut $q_T < 0.2Q$.

range, as well as over the bin width in Q and q_T. We did not make use of the narrow-width approximation for Z production, and included Z/γ^* interference. In the case of D0 run II, we divided our differential cross section by the total cross section calculated with DYNNLO [101, 102].

RHIC

We also considered the data from 200 GeV pp collisions at the Relativistic Heavy Ion Collider (RHIC) recently released by the PHENIX collaboration [89]. Although having few data points and large uncertainties, this is at the moment the only unpolarized set available at RHIC kinematics. We integrated our calculations over the bin width in Q, q_T and y.

LHC

From the ATLAS [94, 95], CMS [96, 97] and LHCb [98–100] experiments at LHC, we included transverse momentum distributions of Z-bosons. Data are provided without correction for experimental cuts on the transverse momentum and rapidity of the final leptons. Hence, in our calculations we had to numerically solve the integral of the lepton tensor over the lepton phase space, including the final-state cuts reported in the respective papers. For ATLAS and CMS, the published values correspond to the normalized fiducial cross section, $1/\sigma_{fid} \cdot \mathrm{d}\sigma_{fid}/\mathrm{d}q_T$, while LHCb reports the absolute fiducial cross section, $\mathrm{d}\sigma_{fid}/\mathrm{d}q_T$. We considered data corrected for final-state QED radiation, and, for ATLAS, we considered the combined electron and muon channels, in the rapidity-binned version. As we did for Tevatron, we integrated our formulas over the bin width in Q, q_T and y, including Z/γ^* interference, and calculating the total cross section with DYNNLO.

4.2.3 Statistical treatment

To calculate the χ^2 w.r.t. data, we reconstructed the covariance matrix from the published values for the (correlated and uncorrelated) experimental uncertainties:

$$(cov)_{ij} = \sigma^2_{i,\mathrm{unc}}\delta_{ij} + \sigma_{i,\mathrm{corr}}\sigma_{j,\mathrm{corr}}, \tag{4.29}$$

where $i, j = 1, \ldots, \mathrm{N_{pt}}$ are indexes for the data points. In general, there can be multiple sources for each kind of uncertainty:

$$\sigma^2_{i,\mathrm{unc}} = \sum_{m=1}^{\mathrm{N_{unc}}} \left(\sigma^{(m)}_{i,\,\mathrm{unc}}\right)^2, \quad \sigma_{i,\mathrm{corr}}\sigma_{j,\mathrm{corr}} = \sum_{n=1}^{\mathrm{N_{corr}}} \sigma^{(n)}_{i,\,\mathrm{corr}}\sigma^{(n)}_{j,\,\mathrm{corr}}. \tag{4.30}$$

Denoting as t_i and m_i respectively the theoretical prediction and measured value for data point i, one has

$$\chi^2 = \sum_{i,j}^{\mathrm{N_{pt}}} (t_i - m_i)\,(cov^{-1})_{ij}\,\left(t_j - m_j\right) . \tag{4.31}$$

When plotting results in presence of correlated errors, to have a more faithful visual representation of the agreement with data, one can rewrite Eq. (4.31) in terms of the so-called "systematic shifts" [103]:

$$\chi^2 = \sum_{i}^{\mathrm{N_{pt}}} \frac{1}{\sigma^2_{i,\mathrm{unc}}} \left(t_i + s_i - m_i\right)^2 + \sum_{j}^{\mathrm{N_{corr}}} \lambda_j^2 . \tag{4.32}$$

We notice that the first term in (4.32) has the same form as an uncorrelated χ^2, but with the theoretical predictions shifted by amounts s_i: $t'_i = t_i + s_i$, where

$$s_i = \sum_{n=1}^{\mathrm{N_{corr}}} \lambda_n \sigma^{(n)}_{i,\,\mathrm{corr}} . \tag{4.33}$$

The second term in (4.32) is called a "penalty term", as it gives the penalty to pay for rewriting the χ^2 in an uncorrelated form. The parameters λ_j can be calculated from (4.31). In our plots, we show only uncorrelated errors for data points, and the shifted theoretical curves. This represents more accurately the statistical agreement, provided the penalty terms are not extremely large. In any case, we show on each plot the total χ^2, taking into account correlations.

Particular attention is required for uncertainties of multiplicative nature, like normalization uncertainties. If these are calculated as

$$\sigma^{(n)}_{i,\,\mathrm{corr}} = \delta^{(n)}_{i,\,\mathrm{corr}} m_i , \tag{4.34}$$

i.e., as the relative multiplicative error times the measured value, then the χ^2 is no more quadratic in m_i and a bias in the minimization is introduced, known as the D'Agostini bias [103, 104]. We have explicitly checked that this bias is particularly severe for E288 data, which are affected by a 25% normalization error: when fitting this data set alone, even a theoretical prediction equal to 0 for all points would yield χ^2 values not far from 1. One way to avoid the bias is by using, instead of m_i in (4.34), independent theoretical predictions, like, for instance, predictions from a previous fit:

$$\sigma^{(n)}_{i,\,\mathrm{corr}} = \delta^{(n)}_{i,\,\mathrm{corr}} t_{0,i} . \tag{4.35}$$

This is the so-called "t_0 method" (see, e.g., [103]). In our case, for $t_{0,i}$ we use predictions from a preliminary fit.

In fits to data, theoretical uncertainties are often neglected in the expression for the χ^2. However, this turns out to be practically impossible in our case, since we have data of unprecedented precision: ATLAS measurements have relative errors smaller than 1%, in some cases of the order of few ‰. This is smaller than the effect of the collinear PDF uncertainty on TMD calculations. We thus decided to include in the covariance matrix the 68% PDF uncertainty propagated by the Hessian method.[10] To this purpose, we used the MMHT Hessian sets [5], and calculated the point-by-point uncertainty on the theoretical observable. The error thus calculated was then added in quadrature to the uncorrelated experimental uncertainties in (4.29). In this procedure, some choice for the nonperturbative TMD parameters in (4.26) is needed. We decided to fix these parameters by doing a preliminary fit without PDF errors in the χ^2. The Hessian uncertainties thus calculated were then used in the final fit. This assumes that the size of uncertainties is weakly dependent on TMD parameters.

To have a measure of the uncertainties on the extracted parameters, we adopted the bootstrap technique, by generating 100 sets of pseudo-data from the mean values and covariance matrix of the original data, and performing 100 separate fits. Confidence intervals for any quantity which is a function of the nonperturbative parameters are then calculated from the mean and standard deviation over replicas (see, e.g., [23, 25]).

4.2.4 Comparison to data

Perturbative QCD description of LHC data

As it has recently been shown [31], transverse-momentum resummation without the introduction of nonperturbative functions is in good agreement with LHC measurements of Z-boson transverse momentum distribution. To have cross-validated results, we performed a comparison of TMD predictions (in our framework) to all the LHC data reported in Tab. 4.3, setting the nonperturbative parameters in Eqs. (4.26)-(4.27) to 0. As can be seen in Fig. 4.6 for a selected bin, it is essential to make use of the highest logarithmic accuracy to achieve a good agreement with data, the description deteriorating extremely rapidly at lower orders. The total $\chi^2/\mathrm{N}_{\mathrm{data}}$, for all the LHC data considered, goes from 1.83 at N^{3}LL, to 6.81 at NNLL, and 95.58 at NLL. We explicitly checked that performing a fit of LHC data alone with a nonperturbative function does not improve the agreement, as the best fit parameters are equal to 0. In any case, it is important to check whether there exists a nonperturbative parameterization such to allow for a good description of,

---

[10] Perturbative uncertainties are difficult to give a statistical interpretation, and are thus not included at the moment.

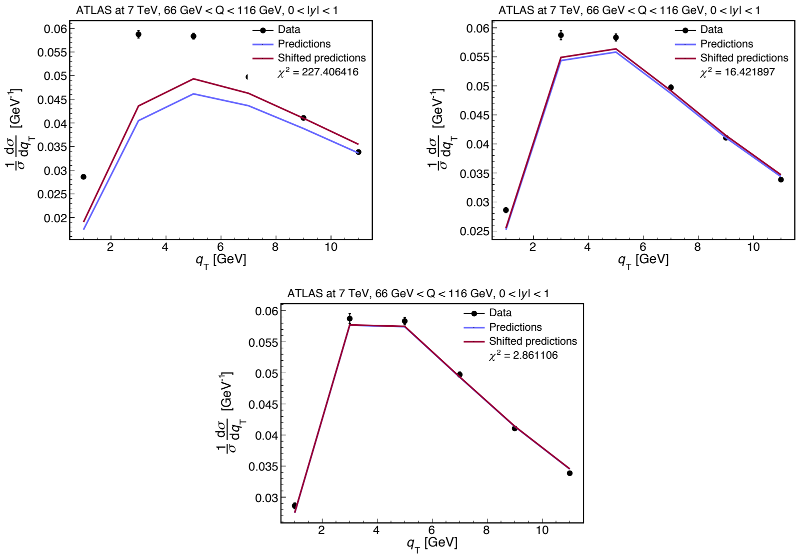


Figure 4.6: TMD predictions with the nonperturbative parameters set to 0, at increasing logarithmic accuracy. Top left: NLL; top right: NNLL; bottom: NNNLL. Only the central rapidity bin from [94] is shown. The χ^2 divided by the number of data points in the bin is calculated as described in Sec. 4.2.3. The shifted predictions are obtained by summing the shift parameters to the actual predictions (see Sec. 4.2.3, in particular Eqs. (4.32)-(4.33), for the definition of systematic shifts).

at the same time, lower energy data and LHC data. In this sense, LHC data will have the role of setting limits on nonperturbative effects, especially at low x.

Simultaneous description of data at different energies

In Tab. 4.4 we show the parameters for the function (4.26) obtained from two different fits: one including all the data in Tab. 4.3, and one excluding LHC data. The nonperturbative evolution of TMDs, encoded in the g_2 parameter, is slightly reduced when including CERN data, while the parameter N_1, describing the size of intrinsic transverse momentum, is compatible with the one extracted from the reduced set of data, and in general parameter uncertainties are greatly reduced. Thanks to the larger x span, and to the fact that in some cases a rapidity binning is provided, the impact of LHC data on the determination of the x dependence is remarkable. It allows, indeed, to extract the parameters α and σ, which would be otherwise completely unconstrained. In Fig. 4.7 is shown the function f_{NP} (Eq. (4.26)) at the initial scale $\zeta = 1$ GeV^2 and different values of x. It is clear how, to accommodate for LHC data, the intrinsic transverse momentum has to

parameter:	N_1	α	σ	g_2
LHC excluded	0.72 ± 0.24	5.9 ± 8.7	1.4 ± 3.1	0.098 ± 0.016
all data	0.92 ± 0.08	3.3 ± 0.5	0.74 ± 0.10	0.054 ± 0.006

Table 4.4: The nonperturbative parameters (Eqs. (4.26)-(4.27)), extracted from a fit of all the data (lower row), and a fit excluding LHC data (upper row).

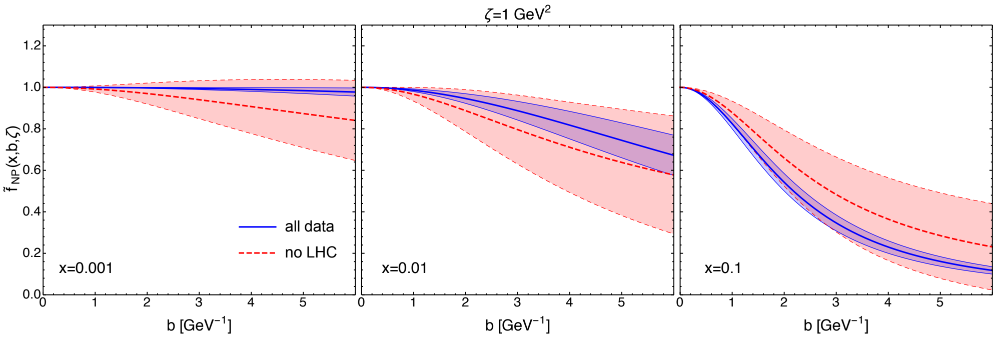


Figure 4.7: The nonperturbative function (4.26) at the initial scale $\zeta = 1$ GeV^2, as extracted from the fit to all data (full blue line), and the fit without LHC data (dashed red line). From left to right, $x = 10^{-3}, 10^{-2}, 10^{-1}$. For each value of b, the average and 68% confidence interval from the replicas is shown.

steeply decrease at low x.[11] The explicit nonperturbative transverse-momentum dependence at the initial scale $\zeta = 1$ GeV^2 is shown in Fig. 4.8, where we plot the function

$$\begin{aligned} f_{NP}(x, p_T, \zeta) &\equiv \int \frac{\mathrm{d}^2\boldsymbol{b}}{(2\pi)^2} e^{i\boldsymbol{p}_{\mathrm{T}}\cdot\boldsymbol{b}} \tilde{f}_{NP}(x, b, \zeta) \\ &\equiv \int \frac{\mathrm{d}b}{2\pi} b J_0(b p_T) \tilde{f}_{NP}(x, b, \zeta), \end{aligned} \tag{4.36}$$

multiplied by $2\pi p_T$ in order to have curves normalized to 1. In p_T space, we do not show curves from the fit without LHC data, since their uncertainty at the initial scale spans orders of magnitude.

The total $\chi^2/(d.o.f.)$ of the fit is 1.24 ± 0.01 when including all data, and 0.65 ± 0.02 when excluding LHC. The χ^2 over number of points for each experimental set is reported in Tab. 4.5. As anticipated in the previous section, ATLAS data are the most difficult to fit, due to the strikingly small errors, and reach values of χ^2

---

[11] Due to the properties of Fourier transforms, a large distribution in b space corresponds to a narrow distribution in momentum space.

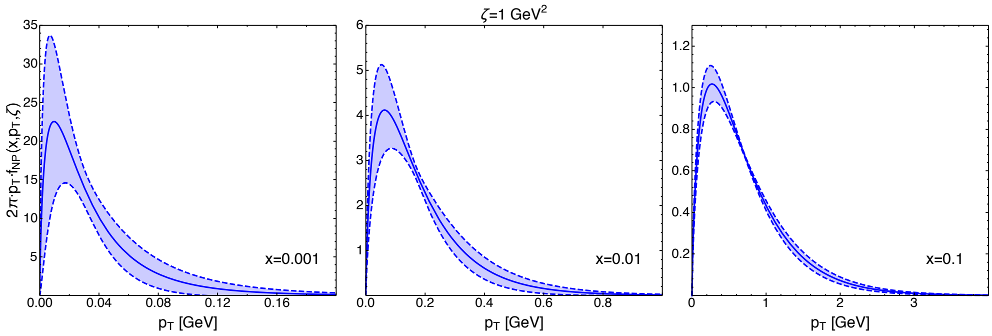


Figure 4.8: The intrinsic transverse momentum distribution as a function of p_T at the initial scale $\zeta = 1$ GeV^2. From left to right, $x = 10^{-3}, 10^{-2}, 10^{-1}$. For each value of p_T, we show the average and 68% confidence interval from the fit including all data. The curves correspond to $2\pi\, p_T\, f_{NP}\,(x, p_T, \zeta)$, hence their integral over p_T is 1.

of the order of some units. However the agreement with theory of the other LHC data is generally good, as well as the overall fit quality. In Figs. 4.9-4.12 we show curves obtained with the mean values of the nonperturbative parameters (second row of Tab. 4.4), compared to data. We stress that these are the first attempts at a simultaneous description of these data in the CSS formalism, and more detailed explorations are needed to draw sharp conclusions.

Experiment	$\chi^2/\mathrm{N}_{\mathrm{pt}}$ (all data)	$\chi^2/\mathrm{N}_{\mathrm{pt}}$ (no LHC)	N_{pt}
E288 200[13]	0.65 ± 0.05	0.56 ± 0.07	30
E288 300 [13]	0.68 ± 0.06	0.52 ± 0.05	30
E288 400 [13]	0.45 ± 0.03	0.49 ± 0.04	65
E605 [88]	0.87 ± 0.06	0.81 ± 0.09	25
PHENIX [89]	4.0 ± 0.4	2.3 ± 0.7	2
CDF run I [90]	0.64 ± 0.02	0.53 ± 0.02	25
CDF run II [91]	1.19 ± 0.09	0.76 ± 0.04	26
D0 run I [92]	0.75 ± 0.02	0.71 ± 0.03	12
D0 run II [93]	2.32 ± 0.07	2.24 ± 0.07	5
Total ($\sqrt{s}$ < 7TeV)	$(\chi^2/d.o.f.)$ 0.86 ± 0.03	$(\chi^2/d.o.f.)$ 0.65 ± 0.02	220
ATLAS 7 [94]	4.30 ± 0.08	-	18
ATLAS 8 [95]	2.65 ± 0.05	-	36
CMS 7 [96]	2.16 ± 0.01	-	3
CMS 8 [97]	1.58 ± 0.03	-	4
LHCb 7 [98]	1.31 ± 0.01	-	7
LHCb 8 [99]	0.51 ± 0.01	-	7
LHCb 13 [100]	0.81 ± 0.03	-	7
Total (all experiments)	$(\chi^2/d.o.f.)$ 1.24 ± 0.01	-	302

Table 4.5: The values of χ^2 over number of points for the different experimental sets considered, in the case of a fit including all data (first column), and one excluding LHC data (second column).

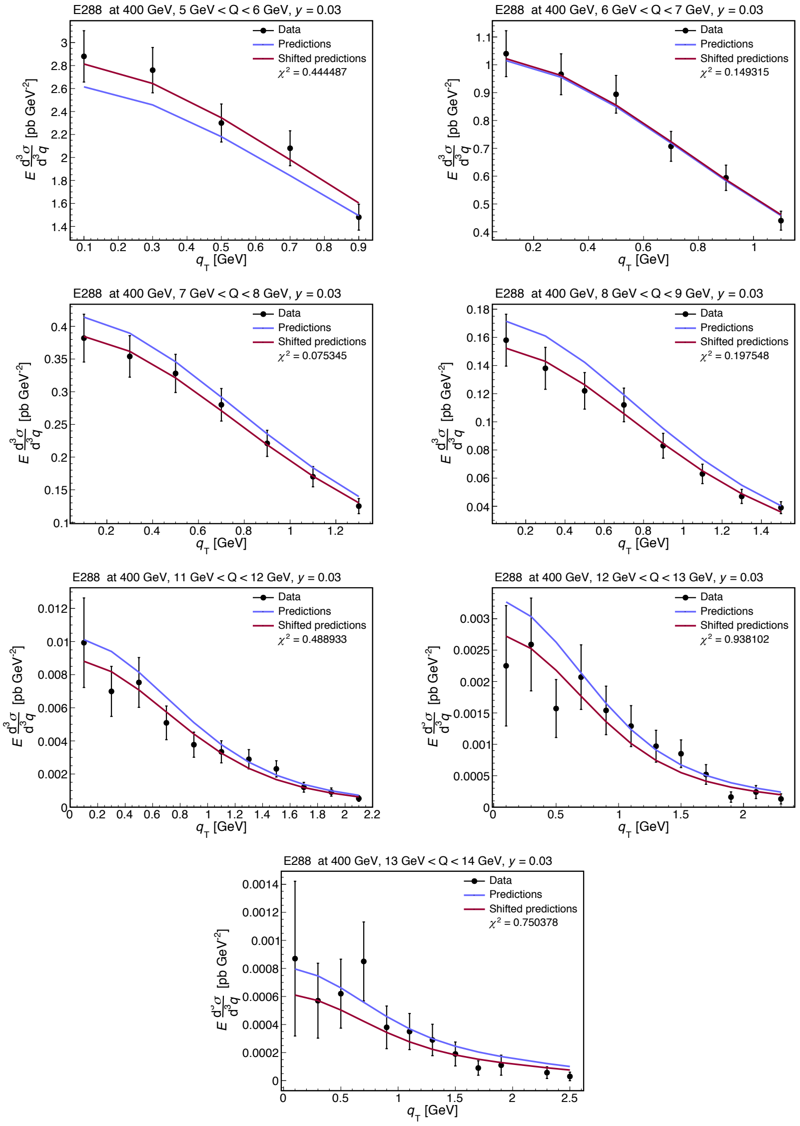


Figure 4.9: TMD curves obtained with the mean nonperturbative parameters from the fit to all data (see Tab. 4.4). Data are from the fixed-target Drell-Yan experiment E288, for $\sqrt{s} = 400$ GeV. The "shifted" curves are obtained by adding to theoretical predictions the systematic shifts, as described in Sec. 4.2.3. This gives a more faithful representation of the statistical agreement with data. The χ^2 over number of data is reported for each bin.

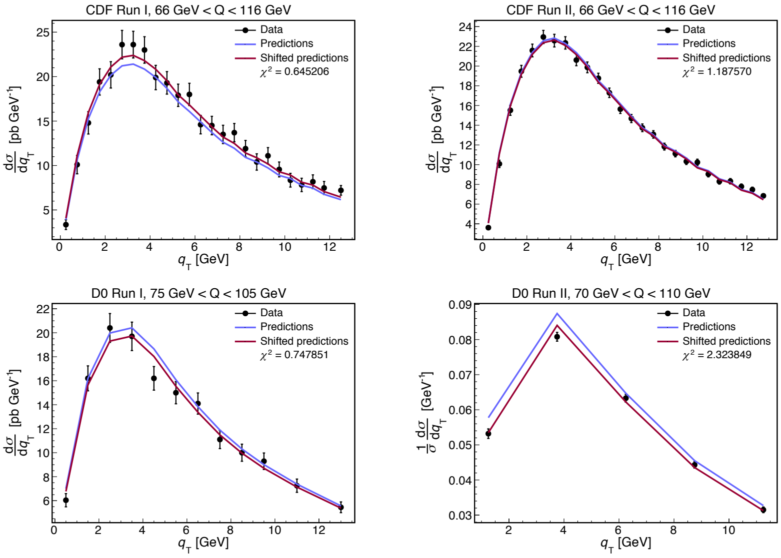


Figure 4.10: Same as Fig. 4.9, for Tevatron Z-production data. Upper panels: CDF run I and II. Lower panels: D0 run I and II.

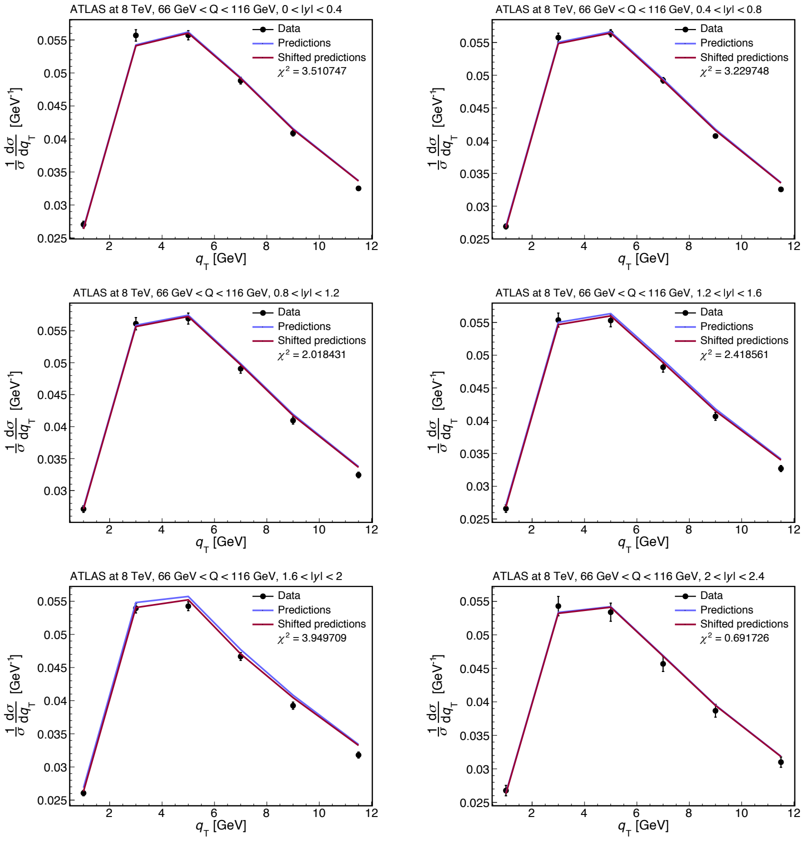
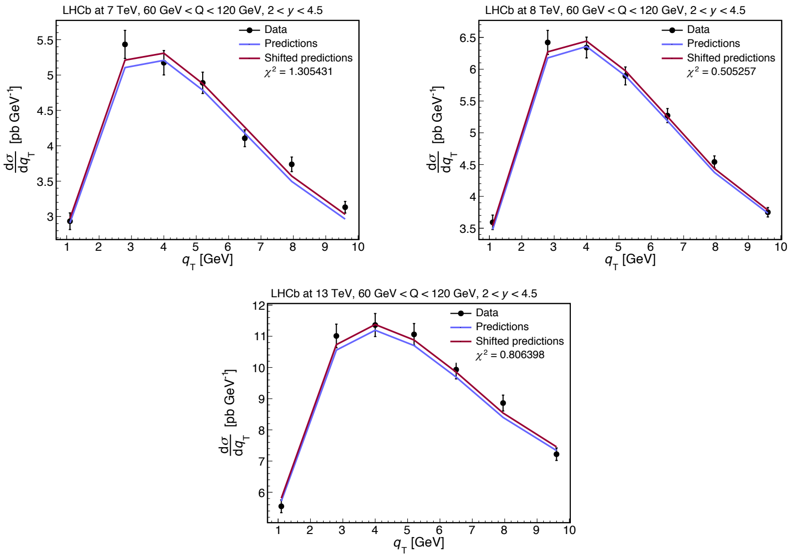


Figure 4.11: Same as Fig. 4.9, for Z-production data at ATLAS ($\sqrt{s}$ = 8 TeV).



LHCb at 7 TeV, 60 GeV < Q < 120 GeV, 2 < y < 4.5

LHCb at 8 TeV, 60 GeV < Q < 120 GeV, 2 < y < 4.5

LHCb at 13 TeV, 60 GeV < Q < 120 GeV, 2 < y < 4.5

Figure 4.12: Same as Fig. 4.9, for Z-production data at LHCb at different center-of-mass energies ($\sqrt{s}$ = 7, 8, 13 TeV).

4A Appendix: difference between the TMD integral and the integrated SIDIS cross section at $\mathcal{O}(\alpha_s)$

In this Appendix we report the theoretical formula for the SIDIS cross section integrated over transverse momentum at $\mathcal{O}(\alpha_s)$, for the reader who wants to compare it with the integral of the TMD cross section (Sec. 4.1.2). This expression can be found, for instance, in [84]:

$$\begin{aligned}\left.\frac{d\sigma^h}{dx\,dQ^2\,dz}\right|_{\mathcal{O}(\alpha_s^1)} = \sigma_0 \sum_{ff'} \frac{e_f^2}{z^2}\left(\delta_{f'f}+\delta_{f'g}\right)\frac{\alpha_s}{\pi}\Bigg\{ &\left[D_1^{h/f'} \otimes C_1^{f'f} \otimes f_1^{f/N}\right](x,z,Q) \\ &+\frac{1-y}{1+(1-y)^2}\left[D_1^{h/f'} \otimes C_L^{f'f} \otimes f_1^{f/N}\right](x,z,Q)\Bigg\},\end{aligned} \tag{4.37}$$

where the first term of the sum gives the structure function F_1, and the second term the longitudinal structure function F_L. Square brackets denote a double convolution over both x and z. The QCD coefficients $C_1^{f'f}$ and $C_L^{f'f}$ are calculated from first order perturbation theory, and read

$$\begin{aligned} C_1^{qq} = \frac{C_F}{2}\Bigg\{ &-8\delta(1-x)\delta(1-z) \\ &+\delta(1-x)\left[P_{qq}(z)\ln\frac{Q^2}{\mu_F^2}+L_1(z)+L_2(z)+(1-z)\right] \\ &+\delta(1-z)\left[P_{qq}(x)\ln\frac{Q^2}{\mu^2}+L_1(x)-L_2(x)+(1-x)\right] \\ &+2\frac{1}{(1-x)_+}\frac{1}{(1-z)_+}-\frac{1+z}{(1-x)_+}-\frac{1+x}{(1-z)_+}+2(1+xz)\Bigg\},\end{aligned} \tag{4.38}$$

$$\begin{aligned} C_1^{gq} = \frac{C_F}{2}\Bigg\{ &P_{gq}(z)\left[\delta(1-x)\ln\left(\frac{Q^2}{\mu_F^2}z(1-z)\right)+\frac{1}{(1-x)_+}\right] \\ &+z\delta(1-x)+2(1+x-xz)-\frac{1+x}{z}\Bigg\}\end{aligned} \tag{4.39}$$

$$\begin{aligned} C_1^{qg} = \frac{1}{4}\Bigg\{ &\delta(1-z)\left[P_{qg}(x)\ln\left(\frac{Q^2}{\mu^2}\frac{1-x}{x}\right)+2x(1-x)\right] \\ &+P_{qg}(x)\left[\frac{1}{(1-z)_+}+\frac{1}{z}-2\right]\Bigg\},\end{aligned} \tag{4.40}$$

$$C_L^{qq} = 2C_F xz, \tag{4.41}$$

$$C_L^{gq} = 2C_F x(1-z), \tag{4.42}$$

$$C_L^{qg} = 2x(1-x). \tag{4.43}$$

In the above expressions μ (μ_F) denotes the factorization scale for PDFs (FFs), while $L_{1,2}$ are abbreviations for the following logarithmic functions:

$$L_1(\xi) \equiv \left(1+\xi^2\right)\left(\frac{\ln(1-\xi)}{1-\xi}\right)_+, \qquad L_2(\xi) \equiv \frac{1+\xi^2}{1-\xi}\ln\xi. \tag{4.44}$$

To facilitate the comparison, one can rewrite the integral of the TMD cross section, Eq. (4.12), in a similar fashion to Eq. (4.37):

$$\int W\bigg|_{O(\alpha_s^1)} = \sigma_0 \sum_{ff'} \frac{e_f^2}{z^2}\left(\delta_{f'f}+\delta_{f'g}\right)\frac{\alpha_s}{\pi}\left[D_1^{h/f'} \otimes C_{TMD}^{f'f} \otimes f_1^{f/N}\right](x,z,Q), \tag{4.45}$$

where

$$\begin{aligned} C_{TMD}^{qq} = \frac{C_F}{2}\Bigg\{ & -8\delta(1-x)\delta(1-z) \\ & + \delta(1-x)\big[2L_2(z)+(1-z)\big] \\ & + \delta(1-z)\big[(1-x)\big]\Bigg\}, \end{aligned} \tag{4.46}$$

$$C_{TMD}^{gq} = \frac{C_F}{2}\Bigg\{P_{gq}(z)\big[\delta(1-x)\ln\left(z(1-z)\right)\big] + z\delta(1-x)\Bigg\}, \tag{4.47}$$

$$C_{TMD}^{qg} = \frac{1}{4}\Bigg\{\delta(1-z)2x(1-x)\Bigg\}. \tag{4.48}$$

We notice that the $O(\alpha_s^1)$ contribution to the TMD integral (4.45) can be obtained from the full QCD calculation (4.37), after subtraction of the following terms:

- terms contributing to the longitudinal structure function F_L (indeed, this structure function is subleading in the TMD framework);
- terms that “correlate” the x and z variable, i.e., terms that do not contain $\delta(1-x)$ or $\delta(1-z)$ in the perturbative coefficients;
- a term $C_F/2\,\{\delta(1-x)[L_1(z)-L_2(z)]+\delta(1-z)[L_1(x)-L_2(x)]\}$ must be subtracted from the C_1^{qq} coefficient (4.38).

Chapter 5

High transverse momentum: phenomenology

In Ch. 4 we discussed Drell-Yan data at different energies, in the low-q_T framework of Transverse Momentum Distributions. A complete description of the available data requires the matching of the TMD formalism with the collinear formalism at higher transverse momentum (see Secs. 2.2 and 3.2 for an introductive discussion about the two regimes and their matching). While this matching seems to be under control at the kinematics of Tevatron and LHC[1], the issue has never been addressed—to our knowledge—in the context of low-energy Drell-Yan experiments. What is more, there are not even studies showing the agreement between these data and collinear-factorized formulas at high q_T, where *high* here means beyond the TMD region. In the context of SIDIS, a recent study [33] has shown remarkable difficulties in the description of data in the collinear-factorization region. In this Chapter we will show that a similar disagreement is present also in low-energy ($\sqrt{s} \lesssim 60$ GeV) Drell-Yan, and we will explore different possible sources of the discrepancy. The work reported here has been published in [105].

5.1 Issues of the matching formalism

In low-energy Drell-Yan, admittedly, the q_T-span of data is more reduced than in SIDIS, and rarely exceeds $q_T \sim Q$. However, as we discussed in Secs. 3.2 and 4.1, the TMD formalism, and even its extension through a W+Y formalism, is well understood only in a narrow region $q_T \ll Q$, beyond which we expect the collinear formalism to give at least an approximate description of data. In this section,

---

[1] See for instance [31, 77], and references therein.

we briefly show what are the limits of the matching formalism in the regime we consider.

The W+Y formalism (Sec. 3.2) is known to present difficulties in SIDIS, most evident when the center-of-mass energy and the Q value are not too high (but persisting also at a HERA-like kinematics)[32]. In particular, it was shown that the W+Y sum is not able to match the fixed order calculation. Indeed, in this regime, contrary to the ideal case $Q \longrightarrow \infty$, the nonperturbative components affect the high-q_T tail of the W term, preventing a proper cancellation against the asymptotic term in Y. To overcome this and other problems, in [73] a modified W+Y prescription was proposed. In this scenario, the W and asymptotic terms are gradually damped with the growing of q_T, in such a way to leave only the fixed-order calculation. In the notation of Sec. 3.2:

$$\frac{\mathrm{d}\sigma^X}{\mathrm{d}q_T\mathrm{d}\ldots}(\text{matched}) \equiv \Xi\left(\frac{q_T}{Q},\eta\right)\left[\mathrm{W} - \frac{\mathrm{d}\sigma^X}{\mathrm{d}q_T\mathrm{d}\ldots}(\text{asy})\right] + \frac{\mathrm{d}\sigma^X}{\mathrm{d}q_T\mathrm{d}\ldots}(\text{F.O.}), \tag{5.1}$$

where Ξ is a damping function, such to be ~ 1 in the TMD region, and approximately ~ 0 when TMD approximations are not reliable. In the cited work, the following example was used:

$$\Xi\left(\frac{q_T}{Q},\eta\right) \equiv \exp\left[-\left(\frac{q_T}{\eta Q}\right)^{\lambda}\right], \tag{5.2}$$

using the values $\lambda = 8$ and $\eta = 0.34$ for phenomenological tests. By making this choices, the authors agreed that, to have reliable predictions, the transition to fixed-order theory should happen for q_T between $\sim Q/4$ and $\sim Q/2$, at the considered kinematics.

Since the causes of the issues mentioned for SIDIS are quite general, it is natural to expect a similar scenario in the Drell-Yan case. This is indeed what is shown by Fig. 5.1. In the left panel, we show the effect of taking the TMD function extracted in [25], and extrapolating it beyond the fit region. First of all, it is remarkable that the TMD curve is always distant from the asymptotic curve, preventing any cancellation. Moreover, we show the difference occurring when using the two different choices for the b^* function in Eqs. (3.24) and (3.25). While the details of this function should only affect the interplay between perturbative and nonperturbative contributions in the low-q_T spectrum, they are shown here to affect the behavior of the curves in the matching region and beyond. In the right panel, we show the application of the modified matching formalism in Eq. (5.2). When using the values quoted above for the damping parameters λ and η, the transition to pure fixed-order theory occurs around $q_T \sim 0.4Q$ (full curve). In this case, the matching looks smooth, but leaves out a number of data points at higher q_T, which should be described by collinear factorization, and are not. For

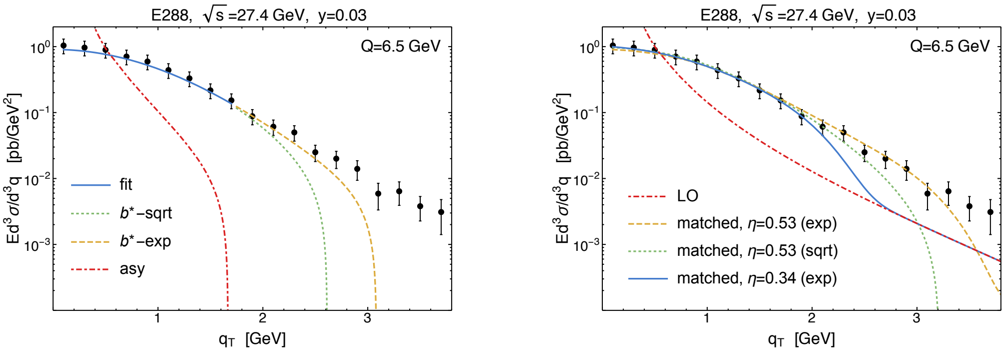


Figure 5.1: **Left**: the TMD cross section (full line) from the fit in [25], when extended beyond the fit region, shows markedly different behavior depending on the functional form chosen for b^* in Eq. (3.23) the dotted line is obtained with the square-root form, (3.24), while the dashed line with the exponential form (3.25). In both cases, $b_{\mathrm{max}} = 1.123\ \mathrm{GeV}^{-1}$. The asymptotic curve is also plotted (at LO, to be consistent with the fit). **Right**: matched curve obtained from the same TMD, with the procedure described in [73], Sec. IX: the full line is obtained with the damping function Ξ (see Eq. (5.2)), using the parameter $\eta = 0.34$ advocated in that paper. This choice enforces a transition to the pure LO prediction between $q_T \simeq Q/4$ and $q_T \simeq Q/2$, and the result is insensitive to the TMD tail and to the choice of the b^* function. The dashed and dotted curves instead show the effect of a slight variation of the parameter η controlling the transition point, for our two types of b^* prescriptions. For a larger value of η the matching to fixed order is not working due to the incomplete cancellation between the TMD tails and the Y term. Data are taken from [13].

a different choice of the η parameter–such to yield a larger non-zero region for the damping function–the matched curve never approaches the fixed-order one, and is completely dependent on the details of the b^* function (dashed and dotted curves). Predictions in this case are thus of dubious interpretation. Above all, we recall that the use of resummation, and of the W+Y prescription, is affected by uncontrolled errors beyond the region where the asymptotic term is positive (see also discussions in Sec. 4.1)

In conclusion, there is presently not a good understanding of the TMD-to-collinear matching in the low-energy Drell-Yan regime. In the following, we will approach the problem starting from the high-q_T side, i.e., from the point of view of collinear factorization.

5.2 Comparison of collinear factorization to data

In this section we provide comparisons of fixed-order collinear factorization predictions for the available q_T-dependent Drell-Yan data in the fixed target (or low-energy collider) regime, from Fermilab, CERN and RHIC experiments, mainly for proton-proton collisions. The center-of-mass energies of the experiments taken into account lie in the range 20 GeV $\leq \sqrt{s} \leq$ 60 GeV (except for RHIC, where $\sqrt{s}$=200 GeV), while the invariant mass of the Drell–Yan lepton pair lies in the range $4.5 \leq Q \leq 13.5$ GeV. For all our theoretical predictions, we use the DYqT [77, 106] and CuTe [107] codes, obtaining completely equivalent results for the fixed-order differential cross sections, at both LO QCD ($O(\alpha_s)$) and NLO QCD $\left(O\left(\alpha_s^2\right)\right)$. These codes also provide an all-order resummation of logarithms in q_T/Q in the cross section, which become relevant toward low q_T. This enables us to study the asymptotic expansion of the resummed result, which we will make use of below. We note that we have also performed cross-checks using the numerical codes of Refs. [108] and [109]. Throughout this paper, the CT14 PDF set [110] will be our default choice.

E866

The E866/NuSea experiment [111] was a fixed-target Drell–Yan experiment designed to measure the internal structure of the nucleon, in particular the asymmetry of down and up antiquarks in the sea, using di-muon events originating from the collision of an 800-GeV proton beam with hydrogen and deuterium targets ($\sqrt{s}$ = 38.8 GeV). The measurement of the q_T-distribution of the muon pair is presented in [112], a Fermilab PhD thesis, and results are given in terms of the differential cross section:

$$\frac{E d^3\sigma}{d^3 q} \equiv \frac{2E}{\pi\sqrt{s}} \frac{d\sigma}{dx_F dq_T^2} = \frac{d\sigma}{\pi dy dq_T^2} \,. \tag{5.3}$$

Data are reported for different bins in $x_F = 2p_L/\sqrt{s}$, ranging from -0.05 to 0.8, and are integrated over different ranges in the invariant mass Q of the muon pair.

The comparison of our LO and NLO theoretical calculations with the experimental data is shown in Fig. 5.2 for the bin $0.15 \leq x_F \leq 0.35$ and for the invariant mass range 4.2 GeV $\leq Q \leq$ 5.2 GeV. The lower part of the plot shows the ratio (data-theory)/theory. The error margins of the data points correspond to the sum in quadrature of statistical and systematic uncertainties, including also an overall normalization uncertainty of 6.5%, as indicated in [112]. Our theoretical predictions are computed at the average Q value and x_F of each bin (Q = 4.7 GeV and x_F = 0.25 in the case of Fig. 5.2). The left plot of Fig. 5.2 shows the comparison of the experimental data with NLO QCD $\left(O\left(\alpha_s^2\right)\right)$ predictions for central values of the factorization and renormalization scales, $\mu_R = \mu_F = Q$. The 90% confidence interval of the CT14 PDF set [110] is included in the plot, but the corresponding variation is barely visible.

An immediate observation from Fig. 5.2 is that the NLO cross section is below

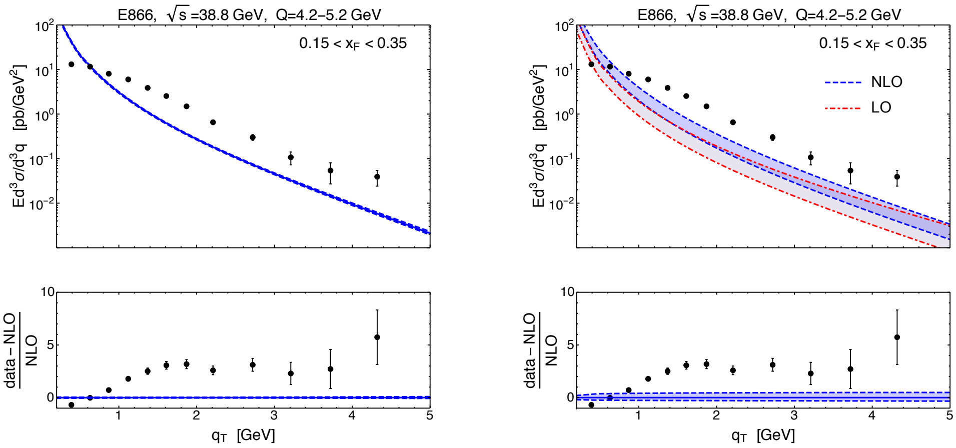


Figure 5.2: Transverse-momentum distribution of Drell–Yan di-muon pairs at $\sqrt{s}$ = 38.8 GeV in a selected invariant mass range and Feynman-x range: experimental data from Fermilab E866 (hydrogen target) [112] compared to LO QCD and NLO QCD results. **Left**: NLO QCD $\left(\mathcal{O}\left(\alpha_s^2\right)\right)$ calculation with central values of the scales $\mu_R = \mu_F = Q = 4.7$ GeV, including a 90% confidence interval from the CT14 PDF set [110]. **Right**: LO QCD and NLO QCD theoretical uncertainty bands obtained by varying the renormalization and factorization scales independently in the range $Q/2 < \mu_R, \mu_F < 2Q$.

the E866 data at high transverse momenta, $q_T \gtrsim 3$ GeV, even within the relatively large uncertainties that the data have here. The NLO cross section falls below the data even much more severely at lower q_T closer to the "matching regime" with TMD physics, where the experimental uncertainties are much smaller. This provides further evidence to our observation above that this regime is presently not well understood theoretically. At the same time we emphasize that data from [112], *integrated over* q_T, are in good agreement with theoretical predictions and are commonly used in global PDF fits [113, 114] (see, for instance, Section 5.1 of [112], where the only relevant discrepancy concerns the lowest mass point ($\langle Q \rangle \simeq 4.4$ GeV) for $0.05 < x_F < 0.25$ (Figs. 5.1-5.5)). This suggests that TMD physics may be the main player for the cross section up to relatively high q_T, since the tail at very large q_T makes only a small contribution to the cross section.

The right plot of Fig. 5.2 shows the effect of varying the renormalization and factorization scales independently in the range $Q/2 < \mu_R, \mu_F < 2Q$, both for the LO QCD ($\mathcal{O}(\alpha_s)$) and the NLO QCD $\left(\mathcal{O}\left(\alpha_s^2\right)\right)$ calculation. The fact that, for $q_T \gtrsim 2.5$ GeV, the NLO uncertainty band overlaps with (and is eventually included in) the LO uncertainty band provides some indication that perturbation theory is well-behaved for this process. On the other hand, we also observe that the NLO scale uncertainty band is only marginally more narrow than the LO one.

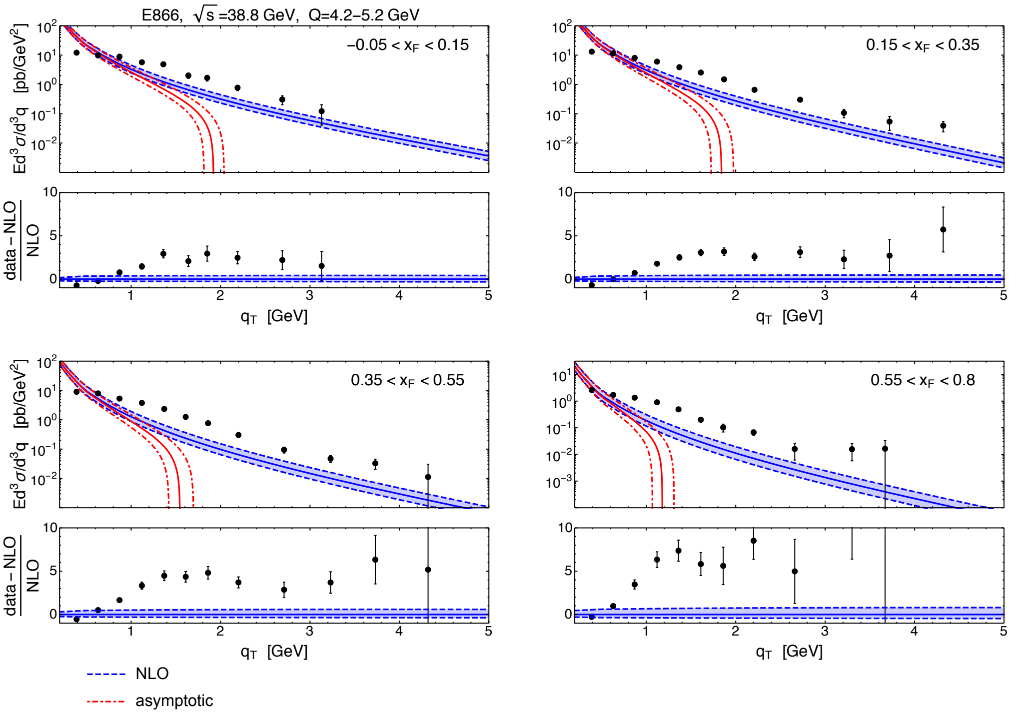


Figure 5.3: E866: comparison between experimental data and NLO QCD predictions for different x_F bins. We also show the low-q_T asymptotic part of the cross section. For details, see text.

We have also considered different PDF choices (CTEQ 10 [115], NNPDF 2.3 [116] and MSTW2008 [113]), obtaining very similar results: the different curves lie within the uncertainty bands shown in the right plot of Fig. 5.2. Such a mild PDF dependence was expected, since the PDFs are well constrained and have small uncertainties in the x-range probed in this process. We conclude that PDF uncertainties (unless they are grossly underestimated by the parameterizations) cannot explain the discrepancy between theory and data at high q_T.

The comparison between data and theory for other x_F bins (Fig. 5.3) and for a different invariant mass range (Fig. 5.4) gives the same qualitative results. The upper part of each plot contains the NLO QCD $\left(\mathcal{O}\left(\alpha_s^2\right)\right)$ prediction (blue) with its uncertainty band obtained through the customary scale variation ($Q/2 < \mu_R, \mu_F < 2Q$) around the central value Q of the invariant mass range. The lower part of each plot again shows the ratio (data-theory)/theory. We also plot the asymptotic expansion of the resummed calculation (red lines). The asymptotic result coincides with the fixed order prediction in the region of very low transverse

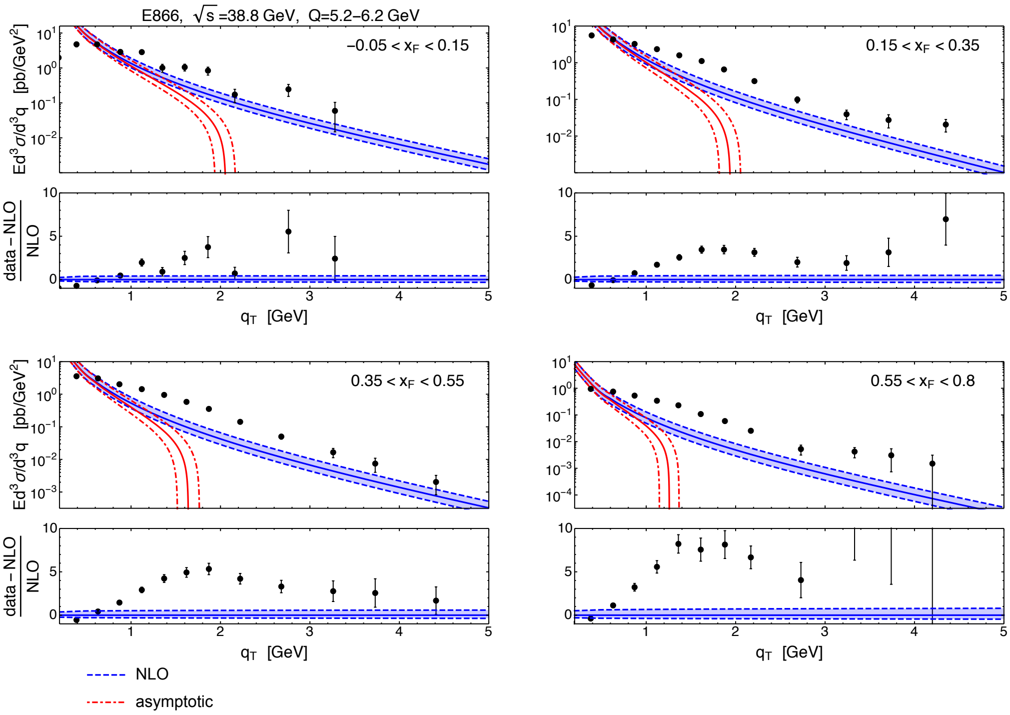


Figure 5.4: E866: comparison between experimental data and NLO QCD predictions for different invariant mass bins. We also show the low-q_T asymptotic part of the cross section. For details, see text.

momenta, but it becomes very small (and eventually negative) with increasing q_T. We show the asymptotic piece in order to obtain a rough guide concerning the region where the fixed-order calculation may start to become reliable [71]: ideally, when q_T is large enough that the difference between the fixed-order and asymptotic calculations (Y term) exceeds the full (W+Y) cross section, one should switch from W+Y to the fixed-order result to obtain more reliable predictions. This occurs for q_T values around 1-2 GeV in the present case. Figures 5.3 and 5.4 show the same qualitative features seen above: the overall agreement between theory and high-q_T data is poor. In general, the disagreement between data and theoretical predictions seems to become worse with increasing Feynman-x_F and to be only mildly dependent on the invariant mass Q of the lepton pair.

R209

The R209 experiment [14, 117] (two proton beams colliding at a center-of-mass energy of $\sqrt{s} = 62$ GeV) was carried out at the CERN ISR (Intersecting Storage Rings) to search for new particles and test scaling models. The differential cross section $d\sigma/dq_T^2$ for the production of a muon pair with transverse-momentum q_T is reported in [118] for the invariant mass range 5 GeV $< Q <$ 8 GeV. The low transverse momentum part of these data has been included in extractions of TMDs [19, 20]. Studies of the whole q_T spectrum can be found in [119, 120].

Comparisons of our NLO results to the R209 data are shown in Fig. 5.5. Again NLO is below the data at high q_T, although the discrepancy is not as statistically significant in this case as for the E866 data. We note that a similar gap between data and theory was reported in [120] in the context of a LO calculation. There, the so-called "k_T-factorization" formalism was claimed to account for the discrepancy. In contrast, in [119] the $W + Y$ formalism was reported to match the data over the whole q_T range.

E288

The E288 experiment [13] measured the invariant cross section $Ed^3\sigma/d^3q$, at fixed photon rapidity, for the production of $\mu^+\mu^-$ pairs in the collision of a proton beam with a fixed target composed of either Cu or Pt. The measurements were performed using proton incident energies of 200, 300 and 400 GeV, producing three different data sets. The respective center of mass energies are $\sqrt{s} = 19.4, 23.8, 27.4$ GeV. Our results are shown in Figs. 5.6, 5.7, 5.8, 5.9. The comparison to data shows the same features as before. We have tested the importance of nuclear effects by computing the cross sections also with the nCTEQ15 [121] and CT14 [110] nuclear PDFs. These turn out to lead to almost indistinguishable results. We

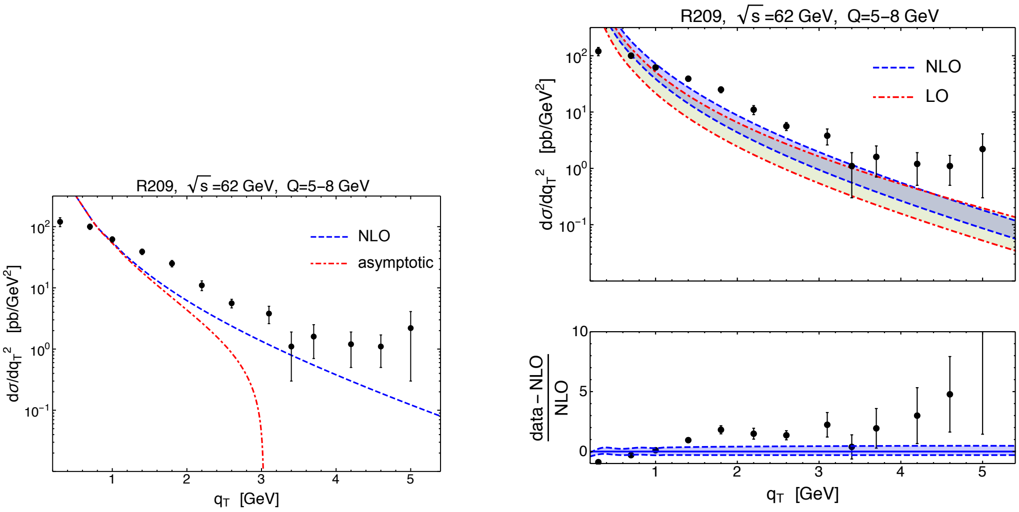


Figure 5.5: **Left**: R209 data [118] compared to NLO QCD $\left(\mathcal{O}\left(\alpha_s^2\right)\right)$. The dashed line shows the asymptotic part. Theoretical results are integrated over the Q range. We have chosen $\mu_R = \mu_F = Q$. **Right**: scale variations ($Q/2 < \mu_R, \mu_F < 2Q$) at LO and NLO.

note that the low transverse momentum part of the E288 data has been used for extractions of TMDs [19, 20, 25, 29, 74].

E605

We also consider the set of measurements of $Ed^3\sigma/d^3q$ in the E605 [88] experiment, extracted from an 800-GeV proton beam incident on a copper fixed target ($\sqrt{s} = 38.8$ GeV). Results at fixed $x_F = 0.1$ are shown in Fig. 5.10. The low transverse momentum part of these data has also been included in extractions of TMDs [19, 20, 25].

PHENIX

Finally, we also compare to the recent measurement [89] performed by the PHENIX collaboration at the Relativistic Heavy Ion Collider in pp collisions at $\sqrt{s}$=200 GeV. The experimental points are taken from Fig. 33 of [89] and compared to LO QCD and NLO QCD, including theoretical uncertainties, in Fig. 5.11. The asymptotic expansion of the W term to NLO is also shown. Evidently, the comparison between NLO and the data is overall satisfactory in this case. It thus appears that there is a qualitative difference between the fixed-target and collider regimes.

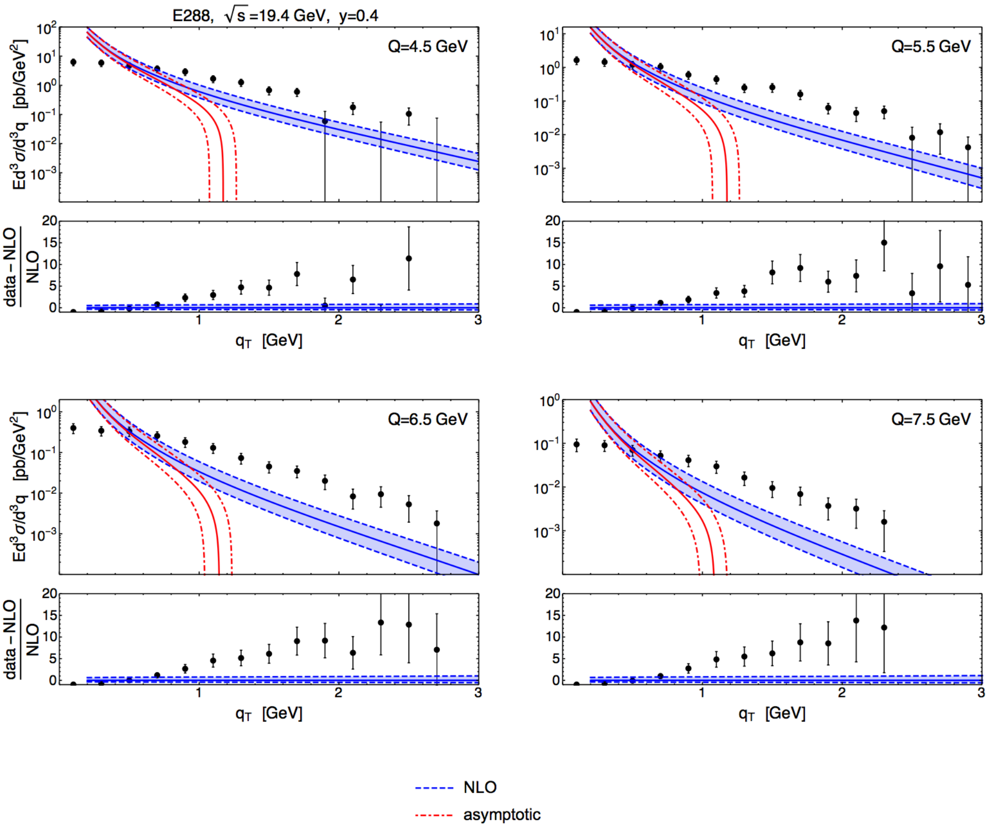


Figure 5.6: E288: experimental data vs. NLO QCD predictions for y=0.4 and different invariant mass bins.

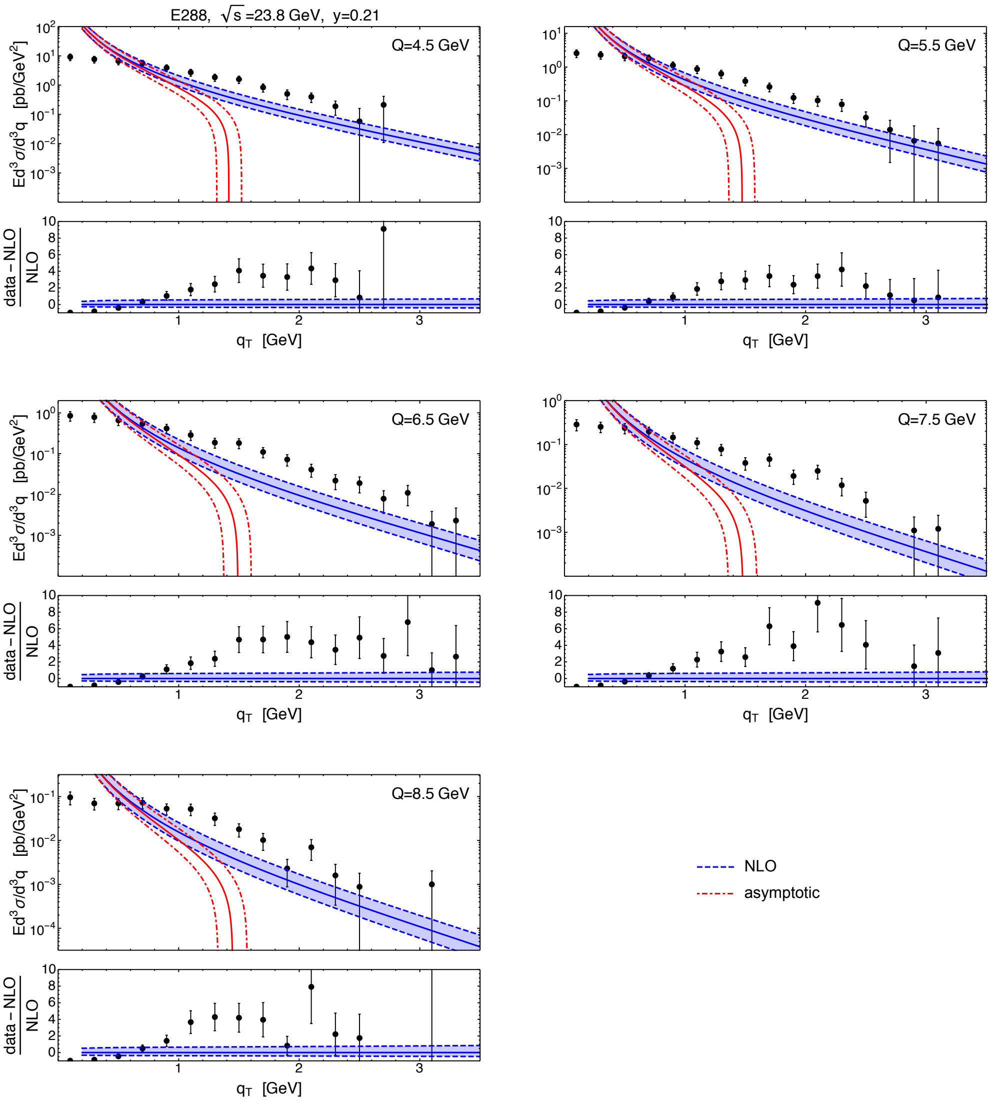


Figure 5.7: E288: experimental data vs. NLO QCD predictions for y=0.21 and different invariant mass bins.

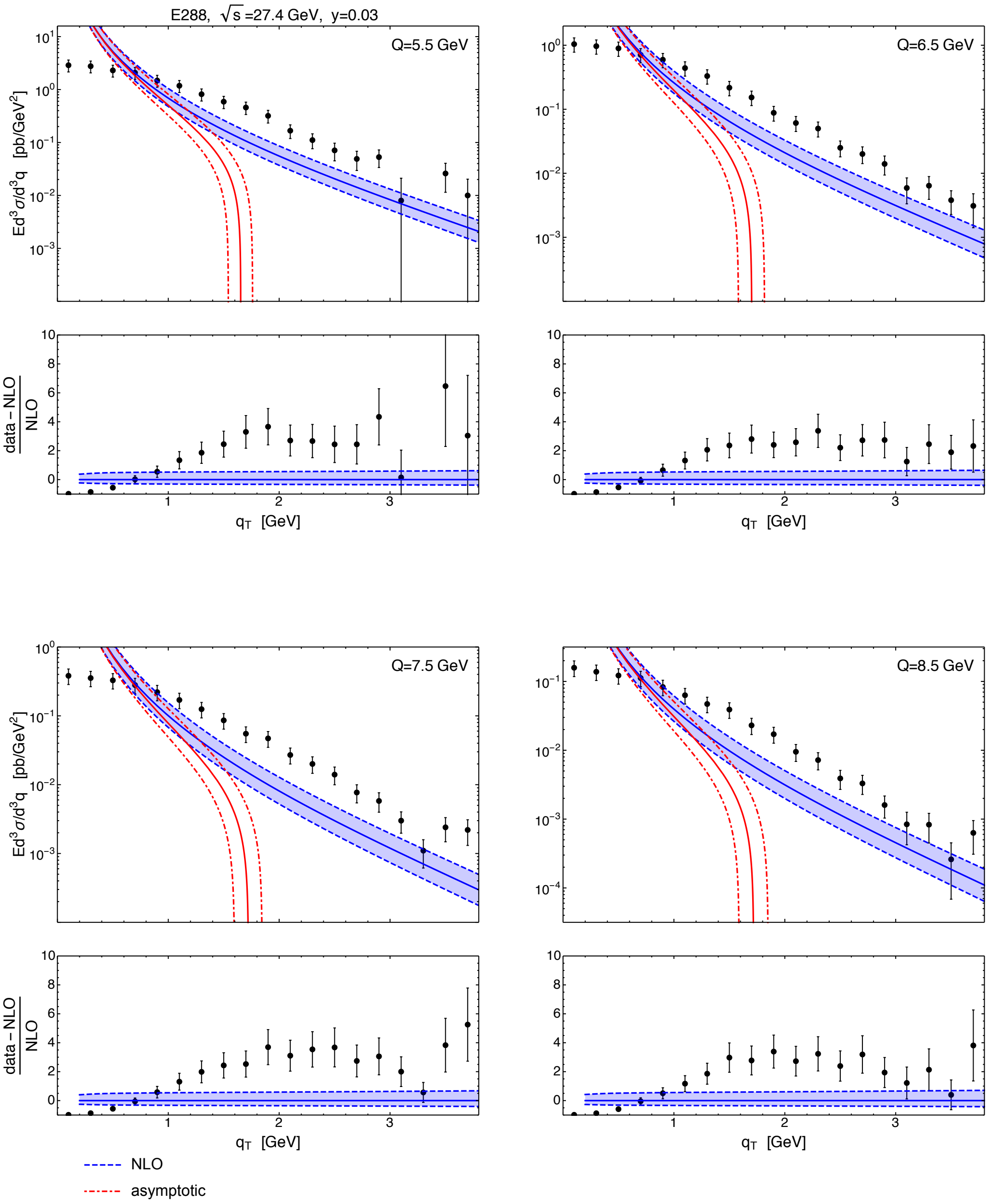


Figure 5.8: Additional plots for E288: experimental data vs. NLO QCD predictions for y=0.03 and different invariant mass bins.

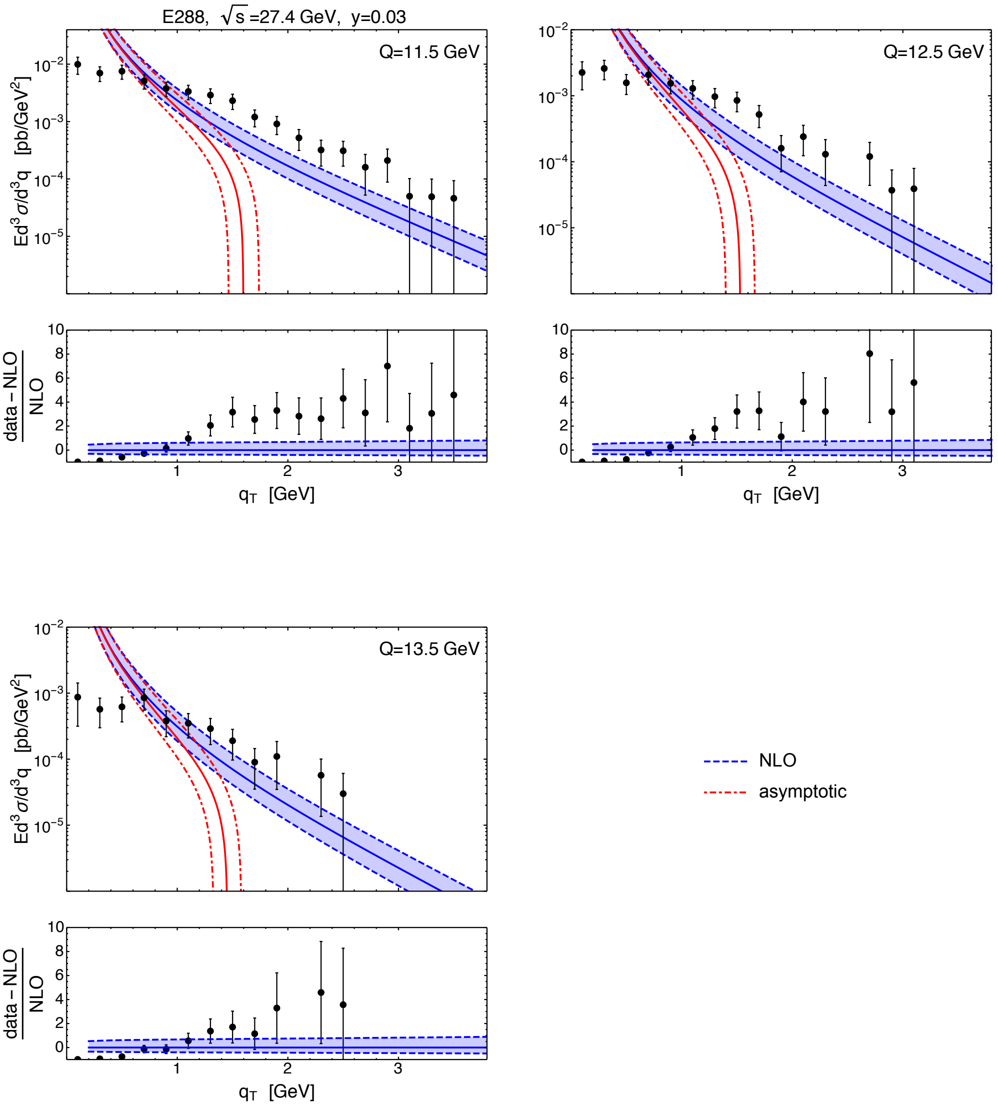


Figure 5.9: Additional plots for E288: experimental data vs. NLO QCD predictions for y=0.03 and different invariant mass bins.

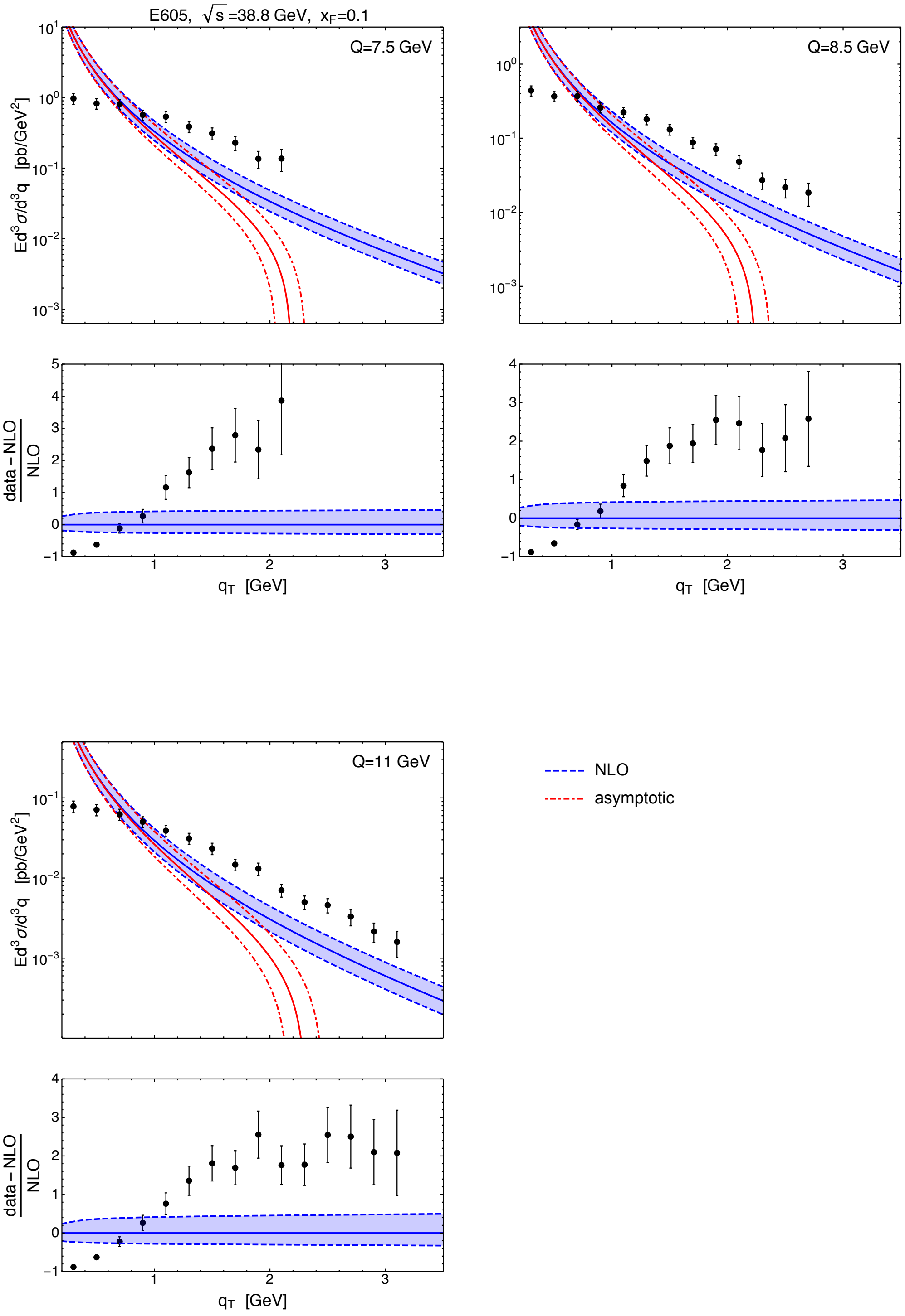


Figure 5.10: E605: experimental data vs. NLO QCD predictions for x_F=0.1 and different invariant mass bins.

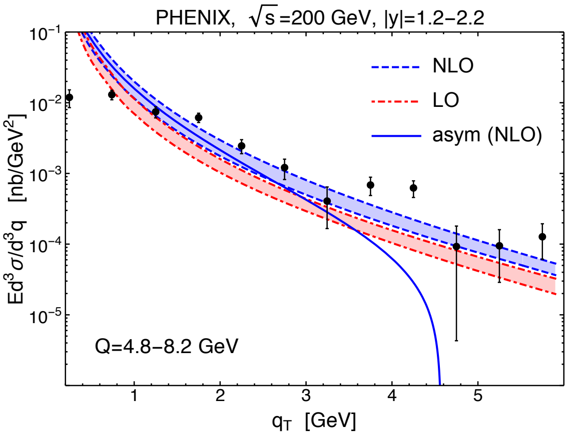


Figure 5.11: PHENIX: experimental data vs. NLO QCD predictions for 1.1< $|y|$ <2.2 and 4.8 GeV $< Q <$ 8.2 GeV.

5.3 Threshold resummation

As we have seen in Fig. 5.2, the NLO corrections to the q_T-differential cross sections are quite sizable. It is therefore important to investigate in how far beyond-NLO perturbative corrections might be relevant for obtaining a better agreement with the data. For the kinematics relevant for the Fermilab and CERN experiments, the invariant mass and transverse momentum of the Drell–Yan pair are such that the production is relatively close to partonic threshold, where a new class of logarithms (separate from that mentioned above at low q_T) arises. Indeed, in the collinear-factorized formula for observables like (5.3),

$$\frac{\mathrm{d}\sigma}{\mathrm{d}q_T\mathrm{d}y\mathrm{d}Q^2} \equiv \int \mathrm{d}x_a \int \mathrm{d}x_b f_a^1(x,\mu_F^2) f_b^1(x_b,\mu_F^2) \frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}q_T\mathrm{d}y\mathrm{d}Q^2}, \tag{5.4}$$

the partonic cross section is actually a function of the variables $\hat{y}_T$ and r:

$$\begin{aligned} \frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}q_T\mathrm{d}y\mathrm{d}Q^2} &\equiv H_{ab}\left(r, \hat{y}_T, \frac{\mu_F^2}{Q^2}, \frac{\mu_R^2}{Q^2}, \alpha_s(\mu_R^2)\right), \\ r &\equiv \frac{q_T}{m_T}, \\ \hat{y}_T &\equiv \frac{q_T + m_T}{\sqrt{\hat{s}}}, \end{aligned} \tag{5.5}$$

where $m_T = \sqrt{Q^2 + q_T^2}$ is the transverse mass, and $\hat{s} = x_a x_b s$ is the partonic center-of-mass squared. We notice that $\hat{y}_T \leq 1$, and for $\hat{y}_T \to 1$ the partonic center-of-mass energy is just sufficient to produce the lepton pair with invariant mass Q and transverse momentum q_T, i.e., it is near production *threshold*. Close to threshold, soft and/or collinear gluon radiation yields a characteristic logarithmic enhancement to the cross section, which, at nth order in perturbation theory, has the form $\alpha_s^n \ln^m \left(1 - \hat{y}_T^2\right)$, with $m \leq 2n$. The summation of these logarithms to all orders is known as *threshold resummation*. We note that $\hat{y}_T$ becomes especially large when the partonic momentum fractions x_a, x_b approach their lower integration limits. Since the PDFs rise steeply towards small argument, this enhances the relevance of the threshold regime, and the soft-gluon effects are relevant even when the hadronic center-of-mass energy is much larger than the produced transverse mass and transverse momentum of the final state.

Large corrections from threshold resummation have been found previously in purely hadronic single-inclusive processes such as $pp \to \pi X$ [122, 123]. For the high-q_T Drell-Yan cross section $pp \to \gamma^* X \to \ell^+ \ell^- X$, the corresponding study was carried out in Sec. IV of [105], adopting the formalism developed in Refs. [124–129]. Referring the reader to the original paper for formal details, we just report here the comparison of threshold-resummed formulas to data at a fixed value of Q and different values of $\sqrt{s}$ (Figs. 5.12-5.13). We notice that the NLO expansion of the resummed formula (black dashed curve) accurately reproduces the NLO result (blue solid curve, with uncertainty bands). This provides some confidence that threshold resummation correctly describes the dominant parts of the cross section to all orders, and that subleading contributions not addressed by resummation are reasonably small. In the left part of Fig. 5.12 we also show the scale uncertainty band for the NLL matched result (red dot-dashed curve), which is barely broad enough to be visible. Evidently, resummation leads to a strong reduction in scale dependence, as one would expect from a result that incorporates the dominant contributions to the cross section at all orders.

Overall, we notice a significant increase of the cross section due to NLL resummation, with respect to the NLO results shown in Sec. 5.2. The enhancement is more pronounced for the case of E288 than for E866 since, for a given Q, at E288 energy one is closer to threshold because of the lower c.m.s. energy. However, despite the increase, the NLL result unfortunately still remains well below the E288 and E866 experimental data at high q_T. We thus conclude that NLL high-q_T threshold resummation alone is not able to lead to a satisfactory agreement with the data.

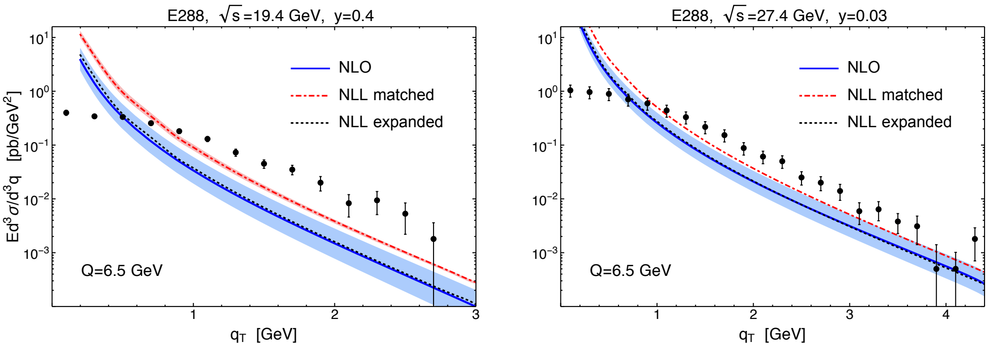


Figure 5.12: E288: experimental data vs. threshold-resummed predictions at NLL+NLO QCD for two different rapidity bins and two different center-of-mass energies.

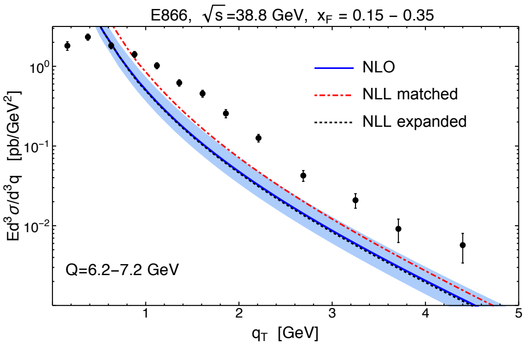


Figure 5.13: E866: experimental data vs. threshold-resummed predictions at NLL+NLO QCD for a selected (x_F, Q) bin.

5.4 Intrinsic-k_T smearing and power corrections

The factorized cross section, Eq. (5.4), receives corrections that are suppressed by inverse powers of $Q \sim q_T$. Little is known so far about the structure and size of such power corrections for the high-q_T Drell–Yan cross section. It is an interesting question whether the discrepancies between perturbative predictions and the high-q_T experimental data seen above might be explained by power corrections. We will try here to address this question from a phenomenological point of view.

As a simple way of modeling power corrections we estimate below the impact of a nonperturbative partonic "intrinsic" transverse momentum k_T on the Drell–Yan q_T spectrum. Such an "intrinsic-k_T smearing" is a phenomenological model that has been invoked in the early literature in cases where collinear factorization was found to underestimate transverse momentum spectra, like for inclusive prompt photon and pion production in hadronic collisions (see for instance [130–132]). For inclusive processes such as these and the high-q_T Drell–Yan process considered here, no general factorization theorem is known that would extend to arbitrary kinematics of the partonic process. For prompt photons, factorization has been established, however, for near-threshold kinematics and low k_T in the framework of the "joint resummation" formalism [133–135], and for high-energy (small-x) dynamics [136]. A technical challenge for all these approaches is the potential for an artificial singularity when the total transverse momentum of the initial state partons is comparable to the observed transverse momentum. A method for dealing with this issue was proposed in Ref. [137] and found to give rise to power corrections to the cross section. A full treatment of the Drell–Yan cross section may require implementation of perturbative joint resummation along with a study of corrections in inverse powers of Q or q_T. Rather than pursuing this elaborate framework, for the purpose of obtaining a simple estimate of the potential size of such higher-order perturbative and power-suppressed nonperturbative effects, we resort to an implementation of a simple model of intrinsic-k_T smearing that will be described now.

5.4.1 Overview of the formalism

The collinear factorization formula for the process $h_1 h_2 \longrightarrow \gamma^* X$ may be adapted from Eq. (5.4) and reads at LO ($O(\alpha_s)$):

$$\begin{aligned} E\frac{d^3\sigma}{d^3\mathbf{q}} \equiv \frac{d\sigma}{dy\, d^2\mathbf{q_T}} &= \sum_{a,b} \int dx_a\, dx_b\, f_{a/h_1}\left(x_a, Q^2\right) f_{b/h_2}\left(x_b, Q^2\right) \\ &\quad \times \frac{d\hat{\sigma}^{ab\to\gamma^* c}}{d\hat{t}} \frac{\hat{s}}{\pi}\, \delta\left(\hat{s} + \hat{t} + \hat{u} - Q^2\right), \end{aligned} \tag{5.6}$$

where as before the $f_{a/h}(x_a, Q^2)$ are the usual collinear PDFs for partons $a = q, \bar{q}, g$ in hadron h. If one allows the incoming partons to have a small transverse momentum $\mathbf{k}_T$, Eq. (5.6) becomes [131]:

$$E\frac{d^3\sigma}{d^3\mathbf{q}} = \sum_{a,b}\int dx_a\, d^2\mathbf{k}_{aT}\, dx_b\, d^2\mathbf{k}_{bT}\, F_{a/h_1}\left(x_a, \mathbf{k}_{aT}, Q^2\right) F_{b/h_2}\left(x_b, \mathbf{k}_{bT}, Q^2\right) \\ \times \frac{\hat{s}}{x_a x_b s}\frac{d\hat{\sigma}^{ab\to\gamma^* c}}{d\hat{t}}\frac{\hat{s}}{\pi}\delta\left(\hat{s}+\hat{t}+\hat{u}-Q^2\right), \tag{5.7}$$

where the functions $F_{a/h}$ are a generalization of the PDFs, including a dependence on transverse momentum. Notice that the partonic Mandelstam invariants must be modified with the inclusion of $\mathbf{k}_T$, and consequently a factor $\hat{s}/(x_a x_b s)$ must be inserted to account for the modification of the partonic flux (see Appendix A of [131]). The modification of the partonic four-momenta is most often done according to two criteria: (1) the partons remain on-shell: $p_{a\mu}p_a^\mu = 0$, and (2) the light-cone momentum fractions retain the usual meaning, e.g.: $x_a = p_a^+/P_a^+$. This leads to the following choice, in terms of Minkowski components [131, 138]:

$$p_a^\mu \doteqdot \left(x_a\frac{\sqrt{s}}{2} + \frac{k_{aT}^2}{2x_a\sqrt{s}},\ \mathbf{k}_{aT},\ x_a\frac{\sqrt{s}}{2} - \frac{k_{aT}^2}{2x_a\sqrt{s}}\right), \tag{5.8}$$

and likewise for the other parton's momentum. Note that we use LO cross sections in Eq. (5.7) since a higher-order formulation is not really warranted for our simple model.

As mentioned above, the framework must become unreliable when k_{aT} or k_{bT} become of the order of the observed transverse momentum, and arguably well before. Large values of k_{aT} can make the partonic Mandelstam in the denominators of the LO hard-scattering cross sections unphysically small. In [131], the following condition was chosen to limit the size of, for example, k_{aT}:

$$k_{aT} < \min\left[x_a\sqrt{s}, \sqrt{x_a\,(1-x_a)\,s}\right]. \tag{5.9}$$

This ensures that each parton moves predominantly along the direction of its parent hadron, and that its energy does not exceed the hadron's energy (a similar condition was adopted in [139], for a study of SIDIS in the TMD regime). However, for $\sqrt{s} \simeq 40$ GeV (E866 and E605 experiments), this condition implies that k_{aT} may still reach values as high as 20 GeV. In our numerical analysis we therefore prefer to introduce an additional cutoff $k_{T\max}$ on both k_{aT} and k_{bT} and will test the dependence of the results on this cutoff.

For the generalized PDFs in Eq. (5.7), the most common choice is

$$F_{a/h}\left(x_a, \mathbf{k}_{aT}, Q^2\right) = f_{a/h}\left(x_a, Q^2\right) \frac{1}{\pi \left\langle k_T^2 \right\rangle} \exp\left[-\frac{k_{aT}^2}{\left\langle k_T^2 \right\rangle}\right], \tag{5.10}$$

where $\left\langle k_T^2 \right\rangle$ is independent of flavor[2] and momentum fraction x_a, but does depend logarithmically on Q^2 because of soft gluon radiation. Instead of Eq. (5.10), one could also consider using the transverse momentum dependent PDFs extracted from the low-q_T spectra of Drell–Yan experiments (as given for instance in Refs. [19, 20, 25, 29]). However, these functions show a non-negligible tail at large k_T, where they lose physical meaning. Hence, if they are used inside a convolution such as Eq. (5.7), the result will strongly depend on the choice of the cutoff $k_{T\max}$, since the integrations (5.7) include contributions from this tail. This dependence will be mostly unphysical and is, in fact, precisely a manifestation of the artificial singularity arising in the partonic scattering functions at really large k_{aT} and k_{bT}. For this reason, we stick with Eq. (5.10); however, we tune $\left\langle k_T^2 \right\rangle$ to the width of the TMD PDFs taken from [25], evolved to the given Q^2. This is shown in Fig. 5.14 where the dashed lines show the evolved TMD of Ref. [25], evolved to Q= 4.7 GeV, normalized by dividing by its integral over $d^2\mathbf{k}_T$. We compare it to a pure Gaussian with a width tuned in such a way that the two distributions become very similar, except for the high-k_T tail. This "equivalent Gaussian" turns out to have a width of $\left\langle k_T^2 \right\rangle = (0.95\ \text{GeV})^2$. It is this Gaussian that we use for our numerical studies presented below.

Our choice of an x-*in*dependent Gaussian width in Eq. (5.10) is motivated by the fact that the x-dependence of $\left\langle k_T^2 \right\rangle$ is still not well constrained in the present TMD fits [25]. Different parametrizations have been proposed in the literature [19, 20], including also x-independent choices [29, 74, 140]. A dependence of $\left\langle k_T^2 \right\rangle$ on x is a natural feature in the joint resummation formalism [137]. In any case, for the mostly exploratory study presented here, an x-independent value of $\left\langle k_T^2 \right\rangle$ appears adequate. Since our goal is to give an upper limit for the k_T-smearing effects, we use the largest value of $\left\langle k_T^2 \right\rangle$ found in [25] (see Fig. 10 there), which occurs at $x = 0.06$.

5.4.2 Numerical results

In Fig. 5.15 we show the effect of k_T-smearing, Eq. (5.7), for E866 kinematics. The $\left\langle k_T^2 \right\rangle$ of the Gaussian is taken as in Fig. 5.14. The impact of smearing on the cross

---

[2] We remark that the initial parton "a" can also be a gluon. Every k_T-smearing model has to make an assumption for the average gluon transverse momentum, which is usually taken to be the same as that for the quarks. We note that perturbative resummations predict dependence of $\left\langle k_T^2 \right\rangle$ on parton flavor [137].

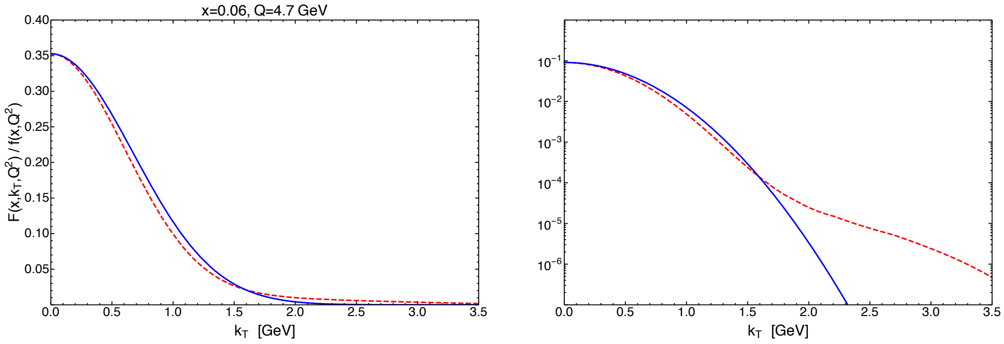


Figure 5.14: Comparison between the TMD of Ref. [25], evolved to the scale Q= 4.7 GeV and divided by its integral over $d^2\mathbf{k}_T$ (dashed line), with the Gaussian (5.10) with $\langle k_T^2 \rangle = (0.95\ \text{GeV})^2$ (full line). Left panel: linear scale; Right panel: logarithmic scale.

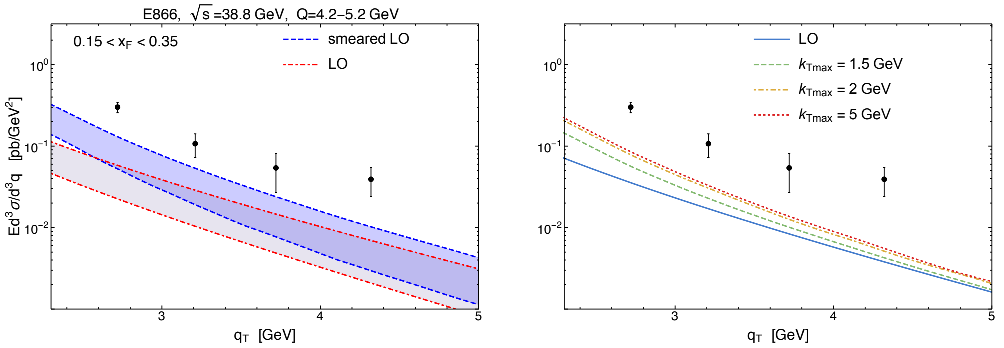


Figure 5.15: Left panel: the effect of k_T-smearing (dashed blue lines), with the cutoff $k_{T\max}$ in Eq. (5.7) set to 2 GeV. The bands correspond to variation of factorization and renormalization scales between $Q/2$ and $2Q$. For comparison, the calculation in ordinary collinear factorization at LO is also shown (red dotted lines). Right panel: the effect of varying the cutoff $k_{T\max}$ in Eq. (5.7). Here the curves correspond to the central values $\mu_R = \mu_F = Q$. For $k_{T\max} \geq 2$ GeV, which corresponds to the 99% percentile of the gaussian in Eq. (5.10), independence from the cutoff is reached.

section overall remains mild, as long as the cutoff $k_{T\max}$ is chosen below 2 GeV. Especially the regime $q_T \simeq Q$ is only little affected by k_T-smearing. We conclude that, although k_T-smearing does somewhat improve the comparison with the data, its effects do not appear to be sufficiently large to lead to a satisfactory agreement. We note that at lower c.m.s. energies as relevant for E288, one is forced to choose smaller cutoffs since the reach in q_T is more limited in these cases.

Chapter 6

High transverse momentum: T-odd effects in perturbative QCD and longitudinal single-spin asymmetries in SIDIS

As we have seen in Ch. 2, different structure functions are present in SIDIS, depending on the polarization states of beam and target. Each of these functions can be described in two different frameworks: at low q_T, they are expressed in terms of TMD PDFs and FFs, while at $q_T \sim O(Q)$ they are calculated in collinear factorization. A complete analysis of the structure functions in the TMD formalism can be found in [50], while a study on the matching of the low- and high-q_T formalisms was reported in [51]. However, for a couple of structure functions, the collinear pQCD calculation is not present in the literature. In the terminology of Sec. 2.2.1, these are $F_{UL}^{\sin\phi_h}$ and $F_{UL}^{\sin 2\phi_h}$, i.e., the $\sin\phi_h$ and $\sin 2\phi_h$ modulations for the case of unpolarized lepton and longitudinally polarized nucleon. Interestingly, these were the first single spin asymmetries observed in SIDIS [37–39]. Due to the low values of P_{hT} involved, these data were studied in a TMD framework, even if their interpretation is not straightforward, due to the competition of different mechanisms[40–44]. In any case, it is interesting to have collinear-factorized predictions for these kind of asymmetries, also because a future Electron-Ion Collider (EIC) will have an extended kinematic range, which might give the possibility to explore different QCD regimes.

A key feature of the above asymmetries is that they are odd under naive time-reversal (T-odd), which corresponds to applying time reversal to all the momenta and spins involved in a scattering process, but without interchanging initial and final states. T-odd effects are absent at tree-level in perturbation theory, since they require a non-trivial absorptive amplitude, i.e., they require loop contributions. They originate, at higher order, from final (initial) state interactions. The term "T-odd" has been in use since before the advent of QCD, in particular in the context of searches for T-violation in electro-weak processes (e.g., [141, 142]): in cases where one could exclude or limit the effect of final state interactions, T-odd correlations would correspond to true violation of time-reversal symmetry. In the TMD framework, T-oddness can arise at leading twist from the properties of the Wilson lines in the definition of the TMD functions. This has been a central issue in studies of the three-dimensional nucleon structure, since the Sivers function would vanish without properly taking into account the role of Wilson lines [35, 57, 59]. In collinear factorization, T-odd asymmetries can be generated by twist-three correlations [143–146]. In leading-twist collinear factorization, on the other hand, T-odd effects have been the object of a longstanding search. Short after the first experimental successes of pQCD, these effects were indeed proposed as a clean test of the gluon self-coupling [147, 148]. The interest in T-odd phenomena in pQCD, being purely loop effects, is still present in modern times, and different predictions have been presented (see, e.g., [149, 150] and references therein). However, the observation of these asymmetries is challenging on the experimental side, due to their small size, so that a clear comparison to data has not been possible to date.

6.1 T-odd observables in perturbative QCD

Denoting as S_{fi} the scattering matrix element between an initial state i and a final state f, a *naive* time-reversal transformation corresponds to a time-reversal without interchange of initial and final states. Hence an observable is said to be T-odd if[1]

$$\left|S_{fi}\right|^2 \neq \left|S_{\tilde{f}\tilde{i}}\right|^2, \tag{6.1}$$

where $\tilde{i}(\tilde{f})$ is obtained from $i(f)$ by reversing momenta and spins. T-odd effects can be present also in the case of theories which are invariant under *true* time-reversal, for which:

$$\left|S_{fi}\right|^2 = \left|S_{\tilde{i}\tilde{f}}\right|^2. \tag{6.2}$$

This is easily understood by considering the reaction matrix T

$$S_{fi} \equiv \delta_{fi} + i(2\pi)^4\delta^{(4)}\left(P_f - P_i\right) T_{fi}, \tag{6.3}$$

---

[1] See, e.g., [151] or similar textbooks for a definition of scattering matrix and the reaction matrix, and, e.g., [141, 142] for a discussion of true time-reversal and T-oddness.

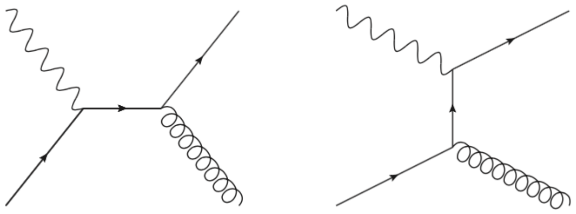


Figure 6.1: Tree-level diagrams for $\gamma^* + q \longrightarrow q + g$.

and the unitarity condition for the scattering matrix

$$T_{fi} - T^*_{if} = i \sum_X T^*_{Xf} T_{Xi} \delta^{(4)} \left(P_X - P_i \right) \equiv i\alpha_{fi}, \tag{6.4}$$

where in the last equation we introduced the notation α_{fi} for the absorptive part of the reaction amplitude. Eq. (6.4) can be rewritten as

$$T^*_{if} = T_{fi} - i\alpha_{fi}, \tag{6.5}$$

and, taking the square modulus of both sides

$$\left| T_{if} \right|^2 = \left| T_{fi} \right|^2 + \left| \alpha_{fi} \right|^2 + 2\mathrm{Im} \left(T^*_{fi} \alpha_{fi} \right). \tag{6.6}$$

T-reversal invariance, Eq. (6.2), implies $\left| T_{if} \right|^2 = \left| T_{\tilde{f}\tilde{i}} \right|^2$ (leaving aside the case $i = f$). Thus, if only electromagnetic and QCD interactions are considered, Eq. (6.6) gives an expression for naive T-violating terms:

$$\left| T_{\tilde{f}\tilde{i}} \right|^2 - \left| T_{fi} \right|^2 = \left| \alpha_{fi} \right|^2 + 2\mathrm{Im} \left(T^*_{fi} \alpha_{fi} \right). \tag{6.7}$$

If we consider–in view of our SIDIS calculation–the process $\gamma^* + q \longrightarrow q + g$, the leading order contributions to T_{fi} are the tree-level diagrams in Fig. 6.1, while the leading terms for the absorptive amplitude α_{fi} are given by the 1-loop diagrams in Fig. 6.2. The lowest-order contributions to T-odd effects come then from $2\mathrm{Im}\left(T^*_{fi} \alpha_{fi} \right)$, and are given by the interference of 1-loop and tree amplitudes. They are thus of order $O\left(\alpha_s^2\right)$. For the imaginary part to be non-zero, the loop amplitudes must have a relative phase with respect to the tree-level amplitude. This will reduce the number of contributing diagrams in Fig. 6.2.

Perturbative T-odd effects in QCD have been calculated for different processes. For $e^+e^- \longrightarrow 3$ jets annihilation, the $O\left(\alpha_s^2\right)$ calculation has been shown to vanish [152, 153]. Predictions for Drell-Yan can be found in [149, 150, 154], and for SIDIS

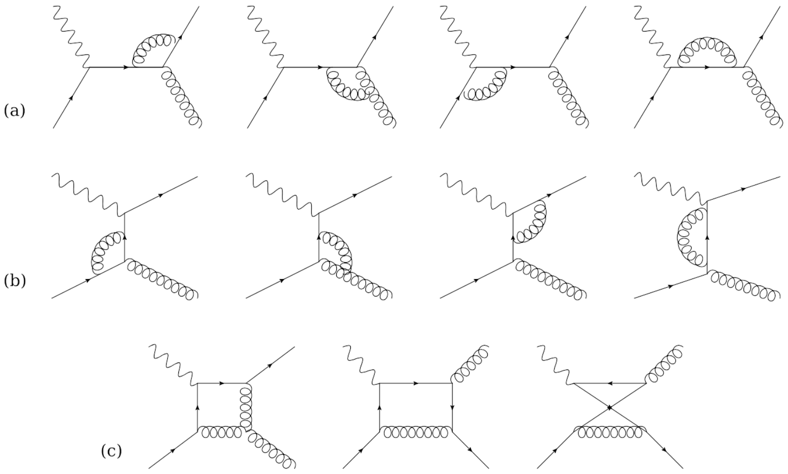


Figure 6.2: 1-loop diagrams for $\gamma^* + q \to q + g$. (a) s-channel (b) t-channel (c) box diagrams.

with polarized lepton beam in [148, 149, 155]. Analogous expressions for SIDIS in the case of polarized target are however lacking, except for the unpublished work reported in [156], where the absorptive amplitudes were derived from the e^+e^- results of [149] by applying crossing relations. In this work, we will follow the alternative strategy of directly calculating the relevant diagrams for the F_{UL} structure functions in SIDIS. This will allow us to cross check the results in [156], and give predictions for a future EIC.

6.2 Perturbative calculation of longitudinal single spin asymmetries

In this section we will discuss the longitudinal target single spin asymmetries (SSAs) in SIDIS. Experimentally, they are defined as

$$A_{UL}(\phi) \equiv \frac{\mathrm{d}\sigma^{\uparrow}(\phi) - \mathrm{d}\sigma^{\downarrow}(\phi)}{\mathrm{d}\sigma^{\uparrow}(\phi) + \mathrm{d}\sigma^{\downarrow}(\phi)}, \tag{6.8}$$

$\mathrm{d}\sigma^{\uparrow(\downarrow)}$ denoting the fully-differential cross section for positive (negative) helicity of the target[2]:

$$\mathrm{d}\sigma^{\uparrow(\downarrow)}(\phi) \equiv \frac{\mathrm{d}\sigma^{\uparrow(\downarrow)}}{\mathrm{d}x\mathrm{d}Q^2\mathrm{d}z\mathrm{d}P_{hT}^2\mathrm{d}\phi}. \tag{6.9}$$

---

[2] This means target polarization antiparallel (parallel) to the beam direction.

Due to the structure of the SIDIS cross section discussed in Ch. 2, single spin asymmetries have the general form

$$A_{UL}(\phi) = \frac{A_{UL}^{\sin\phi}\sin\phi + A_{UL}^{\sin(2\phi)}\sin(2\phi)}{1 + A_{UU}^{\cos\phi}\cos\phi + A_{UU}^{\cos(2\phi)}\cos(2\phi)}, \tag{6.10}$$

where $A_{UL}^{\sin(n\phi)}$ are called *analyzing powers*. They can be calculated as

$$A_{UL}^{\sin(n\phi)} = \frac{\int \mathrm{d}\phi\ \sin(n\phi)\left[\mathrm{d}\sigma^{\uparrow}(\phi) - \mathrm{d}\sigma^{\downarrow}(\phi)\right]}{1/2\int \mathrm{d}\phi\left[\mathrm{d}\sigma^{\uparrow}(\phi) + \mathrm{d}\sigma^{\downarrow}(\phi)\right]}. \tag{6.11}$$

where $\phi \equiv \phi_h$ and $n = 1, 2$. In the collinear factorization framework, the SIDIS cross section can be actually written in terms of partonic cross section, as in Eq. (2.50). At leading order, the partonic subprocesses are 2-to-2 reactions: $\gamma^* + q(\bar{q}) \longrightarrow q(\bar{q}) + g$, $\gamma^* + g \longrightarrow q + \bar{q}$. In terms of 4-momenta, one has $q^\mu + p^\mu = p'^\mu + p_X^\mu$, where p'^μ is the momentum of the fragmenting particle, while p_X^μ is that of the unobserved particle in the final state. The partonic cross section, for a positive-helicity parton, has the expression

$$\frac{d\hat{\sigma}^{\uparrow}}{d\hat{x}\, dQ^2\, d\hat{z}\, dp_T'^2\, d\phi} = \frac{\pi\alpha^2 y^2}{4Q^8} L_{\mu\nu}\hat{W}^{\uparrow\mu\nu}\frac{1}{(2\pi)^3}\delta\left(\frac{p_T'^2}{Q^2} - \frac{\hat{z}(1-\hat{z})(1-\hat{x})}{\hat{x}}\right), \tag{6.12}$$

where the δ-function comes from the on-shellness condition $p_X^2 = 0$. $L_{\mu\nu}$ is the unpolarized lepton tensor in (2.13). The "hat" variables are defined as

$$\hat{x} = \frac{Q^2}{2p\cdot q}, \qquad \hat{z} = \frac{p\cdot p'}{p\cdot q}, \tag{6.13}$$

so that $p^\mu = (x/\hat{x})P^\mu$ and $p' = (\hat{z}/z)P_h^\mu$. The partonic Mandelstam variables are

$$\begin{aligned} \hat{s} &= (q+p)^2 = \frac{1-\hat{x}}{\hat{x}}Q^2, \qquad \hat{t} = (q-p')^2 = -\frac{1-\hat{z}}{\hat{x}}Q^2, \\ \hat{u} &= (p-p')^2 = -\frac{\hat{z}}{\hat{x}}Q^2. \end{aligned} \tag{6.14}$$

$\hat{W}^{\uparrow\mu\nu}$ in Eq. (6.12) is the parton-level version of the hadronic tensor

$$\hat{W}^{\uparrow\mu\nu} \equiv \langle p,\uparrow|J^\mu(0)|\,p'\,p_X\rangle\,\langle p'\,p_X|J^\nu(0)|\,p,\uparrow\rangle, \tag{6.15}$$

for an initial parton with positive helicity. As discussed above, we only need interference terms between tree-level and 1-loop amplitudes, thus remaining with

$$\begin{aligned} \hat{W}_{t/l}^{\uparrow\mu\nu} \equiv{}& \left\langle p,\uparrow\left|J_{tree}^{\mu}(0)\right|p'\,p_X\right\rangle\left\langle p'\,p_X\left|J_{loop}^{\nu}(0)\right|p,\uparrow\right\rangle + \\ &+\left\langle p,\uparrow\left|J_{loop}^{\mu}(0)\right|p'\,p_X\right\rangle\left\langle p'\,p_X\left|J_{tree}^{\nu}(0)\right|p,\uparrow\right\rangle. \end{aligned} \tag{6.16}$$

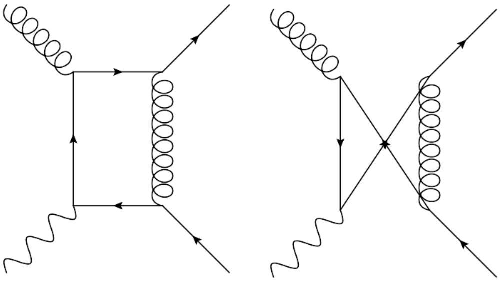


Figure 6.3: The 1-loop diagrams for $\gamma^* + g \longrightarrow q + \bar{q}$ contributing to T-odd asymmetries.

Moreover, since only the imaginary part of $T^*_{fi}\alpha_{fi}$ contributes to T-odd effects, Eq. (6.7), we need only consider terms in the loop amplitude that have a relative phase with respect to the tree amplitude. These are generated by the analytic continuation of logarithms in the loop integrals, e.g.:

$$\ln\left(-\frac{\mu^2}{\hat{s}+i\epsilon}\right) \longrightarrow \ln\left(\frac{\mu^2}{\hat{s}}\right) + i\pi. \tag{6.17}$$

For this reason, in the calculation with a quark in the initial state, Fig. 6.2, t-channel diagrams in row (b), and the third of the box diagrams in row (c), do not contribute to the asymmetries. Instead, in the initial gluon case, only the two box diagrams shown in Fig. 6.3 contribute.

To obtain the $\sin\phi$ and $\sin(2\phi)$ modulations of the cross section in the numerator of (6.11), we thus considered the partonic scattering amplitude

$$L_{\mu\nu}\hat{W}^{\uparrow\mu\nu}_{t/l}, \tag{6.18}$$

–restricting to the interference terms in $\hat{W}^{\uparrow\mu\nu}_{t/l}$, Eq. (6.16)– and we took 1/2 the discontinuity in the complex plane for $\epsilon \longrightarrow 0^{\pm}$ (ϵ being the infinitesimal imaginary part in Feynman propagators), to select the imaginary parts originating from analytic continuation:

$$\frac{1}{2}\left(\lim_{\epsilon\to 0^+} - \lim_{\epsilon\to 0^-}\right) L_{\mu\nu}\hat{W}^{\uparrow\mu\nu}_{t/l}. \tag{6.19}$$

Taking the positive helicity state in (6.16), this gives (half of) the numerator of the single spin asymmetries, (6.11), after convoluting the partonic cross section (6.12) with the helicity PDFs and the FFs:

$$\begin{aligned}\frac{d\Delta\sigma}{dx\, dQ^2\, dz\, dP_{hT}\, d\phi} &= \int d\hat{x}\, d\hat{z}\, dp'_T\, d\xi\, d\eta\, \delta\left(x - \xi\hat{x}\right)\, \delta\left(z - \eta\hat{z}\right) \delta\left(P_{hT} - \eta p'_T\right) \\ &\times \sum_{ab} h_1^b\left(\xi, Q^2\right) D_1^{h/a}\left(\eta, Q^2\right) \frac{d\Delta\hat{\sigma}^{ab}}{d\hat{x}\, dQ^2\, d\hat{z}\, dp'_T\, d\phi} \\ &= \sum_{ab} \int \frac{d\xi}{\xi}\, \frac{d\eta}{\eta^2}\, h_1^b\left(\xi, Q^2\right) D_1^{h/a}\left(\eta, Q^2\right) \frac{d\Delta\hat{\sigma}^{ab}}{d\hat{x}\, dQ^2\, d\hat{z}\, dp'_T\, d\phi},\end{aligned} \tag{6.20}$$

where we defined

$$\frac{d\Delta\sigma}{dx\, dQ^2\, dz\, dP_{hT}\, d\phi} \equiv \frac{d\sigma^{\uparrow} - d\sigma^{\downarrow}}{dx\, dQ^2\, dz\, dP_{hT}\, d\phi}. \tag{6.21}$$

To calculate (6.19), we used the *Tracer* package [157] for Dirac traces in d-dimensions involving γ_5, and *Package-X* [158] for the evaluation of loop integrals and their discontinuity. As a consequence of the Kinoshita-Lee-Nauenberg theorem [159, 160], at 1-loop order T-odd terms must be free of soft and collinear divergences, since there are no corresponding real emission diagrams to cancel them. Consistently, infrared divergences cancel out in our calculation. Our result is

$$\begin{aligned}\frac{d\Delta\hat{\sigma}^{ab}}{d\hat{x}\, dQ^2\, d\hat{z}\, dp'^2_T\, d\phi} &= \frac{\alpha^2\alpha_s^2}{2\pi Q^6} \left(A_1^{ab} \sin\phi + A_2^{ab} \sin 2\phi\right) \\ &\times \delta\left(\frac{p'^2_T}{Q^2} - \frac{\hat{z}(1-\hat{z})(1-\hat{x})}{\hat{x}}\right),\end{aligned} \tag{6.22}$$

where

$$\begin{aligned}A_1^{ab} &= \sqrt{1-y}(2-y)\frac{p'_T}{2Q\hat{x}} \left[\frac{1}{\hat{x}}\Delta h_1^{ab}(\hat{x},\hat{z}) + \left(\hat{z} + \frac{p'^2_T}{Q^2\hat{z}}\right) \Delta h_2^{ab}(\hat{x},\hat{z})\right], \\ A_2^{ab} &= -(1-y)\frac{p'^2_T}{Q^2\hat{x}}\Delta h_2^{ab}(\hat{x},\hat{z}),\end{aligned} \tag{6.23}$$

and, for the case of initial quark (or antiquark), the Δh_i functions are

$$\begin{aligned}
\Delta h_1^{qq}(\hat{x},\hat{z}) =& -\frac{\hat{x}^3(1-\hat{x}-\hat{z})}{(1-\hat{x})(1-\hat{z})} \\
& \times \left[\frac{1}{2}C_F C_A + C_F\left(C_F - \frac{C_A}{2}\right)\left(\frac{3-\hat{z}}{1-\hat{z}} + \ln(\hat{z})\frac{2}{(1-\hat{z})^2}\right)\right], \\
\Delta h_2^{qq}(\hat{x},\hat{z}) =& \frac{\hat{x}^3}{(1-\hat{x})(1-\hat{z})} \\
& \times \left[\frac{3}{2}C_F C_A + C_F\left(C_F - \frac{C_A}{2}\right)\left(\frac{1-3\hat{z}}{1-\hat{z}} + \ln(\hat{z})\frac{2(1-2\hat{z})}{(1-\hat{z})^2}\right)\right].
\end{aligned} \tag{6.24}$$

When the outgoing hadron originates from the fragmentation of the gluon, one has[3]

$$A_1^{gq}(\hat{z}) = -A_1^{qq}(1-\hat{z}), \qquad A_2^{gq}(\hat{z}) = A_2^{qq}(1-\hat{z}). \tag{6.25}$$

In the initial gluon case we obtain instead

$$\begin{aligned}
\Delta h_1^{qg} =& -\frac{\hat{x}^3}{\hat{z}(1-\hat{z})}\left(C_F - \frac{C_A}{2}\right)\left[2(1-\hat{z}) + \frac{2x}{\hat{z}} + \frac{\hat{z}-\hat{x}-1}{\hat{z}(1-\hat{z})} + \right. \\
& \left. - \frac{\hat{x}+\hat{z}-1}{(1-\hat{z})^2}\ln\hat{z} + \frac{\hat{x}-\hat{z}-1}{\hat{z}^2}\ln(1-\hat{z})\right], \\
\Delta h_2^{qg} =& -\left(C_F - \frac{C_A}{2}\right)\frac{\hat{x}^3}{\hat{z}(1-\hat{z})} \\
& \times \left[\frac{1-2\hat{z}+2\hat{z}^2}{\hat{z}(1-\hat{z})} + \frac{1}{(1-\hat{z})^2}\ln\hat{z} + \frac{1}{\hat{z}^2}\ln(1-\hat{z})\right],
\end{aligned} \tag{6.26}$$

and

$$A_1^{\bar{q}g}(\hat{z}) = -A_1^{qg}(1-\hat{z}), \qquad A_2^{\bar{q}g}(\hat{z}) = A_2^{qg}(1-\hat{z}). \tag{6.27}$$

The color factor $(C_F - C_A/2)$ suppresses initial gluon contributions with respect to those for the initial quark case. These will be further suppressed, in the observable, by the smallness of the helicity gluon PDF compared to the valence PDFs. Our results correspond to those in [156], apart from an overall sign for both the $\sin\phi$ and $\sin(2\phi)$ modulations. We believe this can be due to a sign error in Eq. (10.36) of [156] (to be compared to Eqs. (3.12-3.13) in [149]). A cross check is actually provided by the SIDIS beam asymmetries A_{LU} calculated in [148, 149]. Indeed, if one considers the neutrino scattering case in those calculations, the W-boson mediated interaction selects left-handed quarks, so that, even if the target

---

[3] The coefficient of $\sin\phi$ acquires a minus sign because the two outgoing particles have a relative angle of 180° in the transverse plane.

is unpolarized, the partonic matrix elements are the same as in our Eq. (6.16), with reversed helicity. By looking for instance at the functions F_8 and F_9 in Eq. (3.14) of [149], in the case of quark-initiated diagrams, one can indeed check that they correspond to our A_1 and A_2 functions with a reversed sign. . Clearly, the reasoning above does not allow comparisons in the initial gluon case, and the partonic functions are different when going from A_{UL} to A_{LU} asymmetries [40–44].

6.3 Phenomenological results

In this section we show our numerical estimates for $A_{UL}^{\sin\phi}$ and $A_{UL}^{\sin(2\phi)}$, obtained using the DSSV set for helicity PDFs [161, 162], and the DSS14 set for FFs [163], and setting both the renormalization and factorization scales equal to Q. For the denominator of the asymmetries we used a LO ($O(\alpha_s)$) expression, with the MMHT [5] set for the unpolarized PDFs.

As mentioned above, A_{UL} asymmetries were the first Single Spin Asymmetries observed in SIDIS [37]. Although the low P_{hT} range reached by the available data seems to suggest that a TMD formalism interpretation is more appropriate–and different studies have analyzed them in terms of leading and sub-leading twist TMDs[4]–we consider it interesting to compare also to our collinear-factorization predictions . Indeed, the two frameworks may not be so sharply separated in this case: what is relevant for distinguishing the TMD regime from the collinear one is the value of q_T/Q, and taking the average values of P_{hT} and z in [37–39] gives $q_T \simeq 1$ GeV, while Q is of the order of $1 \sim 2$ GeV. What is more, a quantitative study of collinear pQCD predictions is the first step towards a matching to TMDs.[5]

One issue when considering target spin asymmetries is given by the fact that the target is actually polarized along the beam direction, so that the experimental asymmetry, A_{UL}^{l}, receives contributions from both longitudinal and transverse asymmetries with respect to the direction of the virtual photon, A_{UL}^{q} and A_{UT}^{q} [164]. Combining with data from transversely polarized target, a reanalysis of the data of [37] has been done in [165], in order to provide a separation of the A_{UL}^{q}

---

[4] When decomposing the target spin with respect to the virtual photon direction, both the longitudinal and transverse components are present. Hence, one has to consider the competition between the twist-three combinations that make up the F_{UL} functions at TMD level—among which the most relevant involves the Collind fragmentation function $H_1^{\perp}$—and the twist-two F_{UT} modulations, which involve also the Sivers function $f_{1T}^{\perp}$. See [50] for a detailed decomposition of the structure functions in terms of TMDs, and [44] for an early discussion of the competing mechanisms.

[5] We notice that our collinear-factorized calculation for the $F_{UL}^{\sin 2\phi}$ structure function cannot match the one calculated in the TMD framework, since this is of order $O(\alpha_s)$ (see Table 2 of [51]), while we are dealing with 1-loop formulae. Instead, there is the possibility of matching for $F_{UL}^{\sin\phi}$.

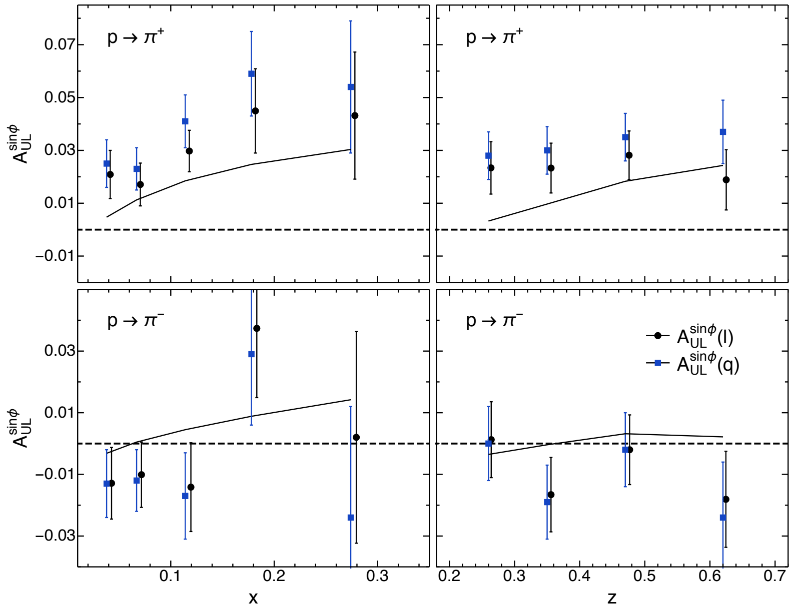


Figure 6.4: Comparison of our theoretical predictions for $A_{UL}^{\sin\phi}$ to HERMES data from [165]. These data were obtained from a reanalysis of those in [37], to estimate the asymmetry for a polarization along the virtual photon direction (blue squares), from the measured asymmetry (black dots).

contribution. In Fig. 6.4 we show a comparison of those data with our predictions for the $\sin\phi$ asymmetry, calculated at the mean values of x, z, Q^2 and P_{hT} for each point reported in Table 1 of [165].

In Fig. 6.5 we compare the analyzing powers for both $\sin\phi$ and $\sin(2\phi)$ using the original data from [37] (hence without corrections for the polarization direction), while in Fig. 6.6 we compare to asymmetries off a deuteron target, from [39]. For these sets of data, we calculate our predictions at the declared mean values of x and Q^2 for each point, while assuming for $\langle z\rangle$ and $\langle P_{hT}\rangle$ the same values as in the binning operated by [165], as they are not declared here, and the binning choices are very similar. Although theoretical curves generally tend to underestimate the asymmetries, especially for the "corrected" data in Fig. 6.4, we notice that the trend of data is reproduced in most cases, as well as the difference in magnitude between the $\sin\phi$ and $\sin(2\phi)$ components, and between π^+ and π^- and the two different targets. The situation thus appears to be different from the unpolarized structure function $F_{UU}^{\cos 2\phi}$ analyzed in [166]: there, the perturbative $\mathcal{O}(\alpha_s)$ prediction at HERMES kinematics was shown to be negligible compared to

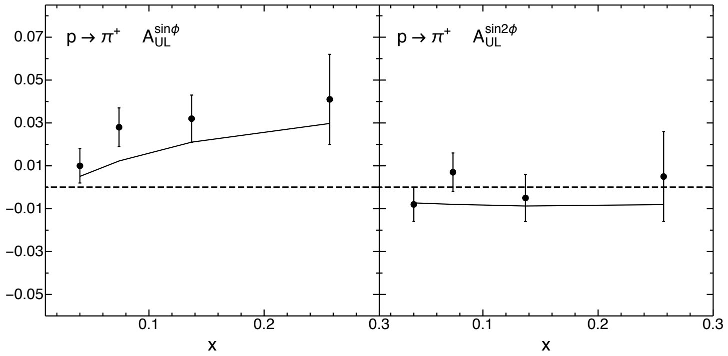


Figure 6.5: $A_{UL}^{\sin\phi}$ and $A_{UL}^{\sin 2\phi}$ asymmetries for production of positive pions off a proton target [37], compared to our calculations.

higher-twist effects.

We will now show some estimates for a future Electron-Ion Collider (EIC). One crucial difference compared to fixed-target experiments is that lower values of x are explored. Since the helicity PDFs strongly suppress asimmetries outside the valence region, this makes $\sin\phi$ and $\sin(2\phi)$ modulations harder to detect in a collider regime. In Fig. 6.7 (Left) we show the steep x-dependence of the analyzing powers for a center-of-mass energy of 140 GeV, at given values of z, Q^2 and P_{hT}. The kinematic values have been chosen by looking at projected data in [167]. We have checked that, for lower values of the invariant mass, $Q^2 \lesssim 10\,\mathrm{GeV}^2$ (hence $10^{-4} \lesssim x \lesssim 10^{-2}$), where the majority of events is expected, asymmetries are generally lower than $\sim 10^{-3}$. Fig. 6.7 (Right) shows the z-dependence of our predictions, which is moderate through most of the range.

Due to this reasons (strong suppression of asymmetries at low x, mild z-dependence) in the following we show results integrated over large bins in z and Q^2, but differential in x and P_{hT}^2:

$$A_{UL\,\mathrm{int}}^{\sin(n\phi)} \equiv \frac{\int_{z_{min}}^{z_{max}} \mathrm{d}z \int_{Q^2_{min}}^{Q^2_{max}} \mathrm{d}Q^2 \int \mathrm{d}\phi\ \sin(n\phi)\left[\mathrm{d}\sigma^{\uparrow}(\phi) - \mathrm{d}\sigma^{\downarrow}(\phi)\right]}{1/2 \int_{z_{min}}^{z_{max}} \mathrm{d}z \int_{Q^2_{min}}^{Q^2_{max}} \mathrm{d}Q^2 \int \mathrm{d}\phi \left[\mathrm{d}\sigma^{\uparrow}(\phi) + \mathrm{d}\sigma^{\downarrow}(\phi)\right]}, \tag{6.28}$$

and we concentrate on the higher-x part of the spectrum. As can be seen from Fig. 6.8, analyzing powers are close to 0 for π^- production, while for π^+ they reach values around 0.01 only in the highest accessible P_{hT} range (see again [167] for estimates of SIDIS cross sections at an EIC).

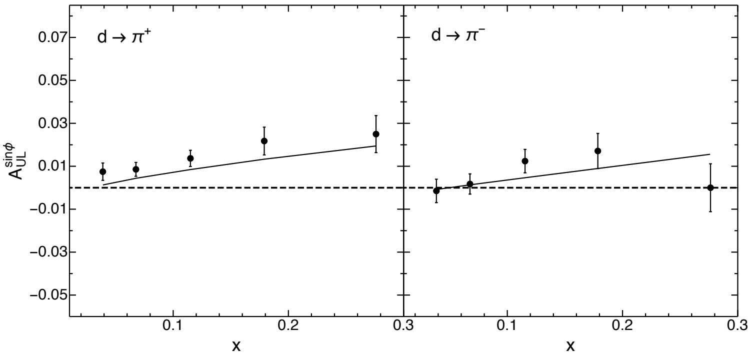


Figure 6.6: $A_{UL}^{\sin\phi}$ asymmetry for pion production off a deuteron target [39], compared to our calculations.

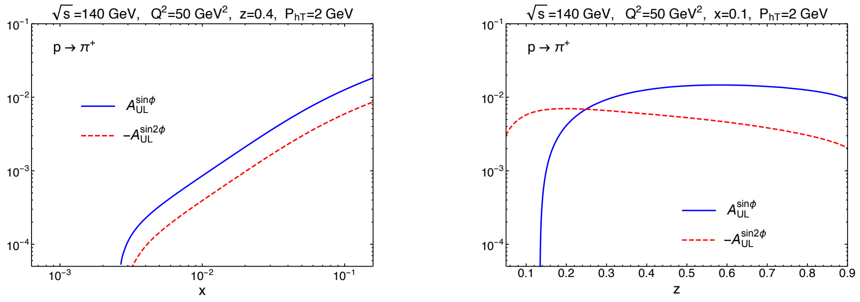


Figure 6.7: x-dependence (**Left**) and z-dependence (**Right**) of the analyzing powers for positive pion production from a proton target at a future EIC. The $\sin(2\phi)$ analyzing power is multiplied by -1. In both figures, $\sqrt{s} = 140$ GeV, $Q^2 = 50$ GeV^2, and $P_{hT} = 2$ GeV. In the left panel, z is fixed to 0.4, while in the right panel $x = 0.1$.

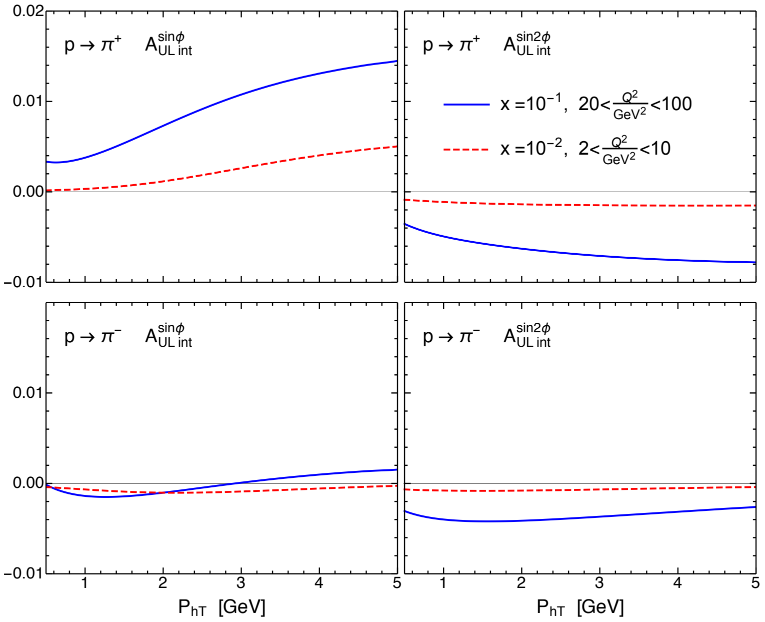


Figure 6.8: Predictions for analyzing powers $A_{UL\,\mathrm{int}}^{\sin(n\phi)}$ at an EIC with $\sqrt{s} = 140$ GeV, integrated over z and Q^2, for two different x values (full lines: $x = 10^{-1}$, dashed lines: $x = 10^{-2}$). In all cases the integration range for z goes from 0.05 to 0.8, while the one for Q^2 is $20 - 100$ GeV^2 or $2 - 10$ GeV^2, respectively for $x = 10^{-1}$ or $x = 10^{-2}$.

The results shown above are for a center-of-mass energy of 140 GeV, around the highest values foreseen for the EIC. However, for given values of x, Q^2 and z, asymmetries depend on $\sqrt{s}$ only through the y-dependent prefactors of the structure functions, like A_1 and A_2 in (6.22), and this dependence is canceled to a large extent when taking the ratio to the unpolarized cross section, in (6.11). We checked this explicitly by comparing predictions at $\sqrt{s} = 45$ GeV and 140 GeV. There will be differences in the kinematic span and in the available luminosity, as well as the different expected number of events (see, for instance, Figs. 1-2 of [167]).

Chapter 7

Conclusions

In this thesis we presented investigations on different aspects of transverse-momentum dependent observables in QCD. After reviewing the general properties of SIDIS and Drell-Yan cross sections (Ch. 2) and the evolution and formal properties of TMDs (Ch. 3), we concentrated on the lower q_T spectrum of both SIDIS and Drell-Yan (Ch. 4). This is a crucial kinematic region, since it gives access to three-dimensional information on the momentum distribution of quarks. In Sec. 4.1, we have evidenced how the case of SIDIS multiplicities is controversial: the majority of data lies beyond what–on formal grounds–can be considered the TMD region, but, at the same time, their shape is well described by gaussian-like parameterizations, as shown in previous parton-model (and leading order) fits. The apparent contradiction is due to the fact that lowest-order solutions for TMDs–i.e., solutions with the hard factor and Wilson coefficients at zeroth order in α_s–are not properly normalized and thus hide difficulties in reproducing the data, which only show up at higher order. Indeed, when going at order α_s, theoretical predictions underestimate measurements by a factor ~ 2. We provided formal arguments to confirm that this is precisely due to the above issue, i.e., that the considered data are mostly beyond the TMD range. In particular, one argument is that the TMD-factorized cross section is, consistently, normalized to the (infrared-regularized) q_T-integral of the asymptotic term, which is naturally smaller than the total integrated cross section. The smallness of this integral is quantified by the theoretical ratio given by Eq. (4.24),

$$\frac{\frac{\mathrm{d}\sigma^h}{\mathrm{d}x\mathrm{d}Q^2\mathrm{d}z}}{\int \mathrm{W}\big|_{O(\alpha_s^0)+O(\alpha_s^1)}}, \tag{7.1}$$

which reproduces, to a good approximation, the normalization factors needed to recover the discrepancy between theory and data. In conclusion, TMD factorization in its standard form is not able to reproduce SIDIS data at the presently available kinematics. The reason is most probably that there is little or no room for the application of the TMD formalism, which requires $q_T \ll Q$. Moreover, perturbative calculations that underpin the TMD formalism beyond lowest order should be applicable only in the region $q_T \gg \Lambda_{QCD}$. Clearly, when Q approaches Λ_{QCD} or M, the whole formalism is questionable. At the same time, the fact that lowest-order results can successfully reproduce the data probably indicates that intrinsic transverse momentum plays a prominent role.

For the Drell-Yan process, we have presented an analysis of TMDs with the inclusion of data from LHC, for the first time without multiplying theoretical predictions by any artificial normalization factor. It is a delicate procedure, since some of these data, in particular those from ATLAS, have an unprecedented precision. It is found that the maximum available perturbative accuracy (N^3LL) is needed in order to have an acceptable description. Moreover, LHC data appear to be compatible with a vanishing intrinsic transverse momentum. Their addition in a fit is however crucial for reducing the present uncertainties on TMDs, especially at low x. Indeed, they can be thought of as setting an upper limit on the size of nonperturbative effects. With a 2-gaussian (slightly different from a gaussian) as the nonperturbative function, a good simultaneous description of data at different energies is obtained. These are however first explorations of the newest data in the standard TMD formalism,[1] and will be followed by more detailed studies, where different parameterizations will be compared [168].

As anticipated above, a recent work [33] has found discrepancies in the theoretical description of SIDIS data at intermediate to high q_T, i.e., beyond the TMD region: predictions from collinear factorization have been shown to largely underestimate differential multiplicities measured by COMPASS [83]. In Ch. 5, we showed an analogous study for Drell-Yan at low energy. Fixed-target data are widely used in TMD analyses (see, e.g., [19, 20, 25, 29, 30, 74]). However, analyses of points at higher q_T are lacking. We showed that, analogously to the SIDIS case, these points are largely above theory. The effect appears to depend on the center-of-mass energy, as PHENIX data ($\sqrt{s} = 200$ GeV) show a fair agreement with calculations, and also data from Tevatron and LHC (for Z boson production) are known to be correctly reproduced. We discussed possible mechanisms that might contribute to the discrepancies: on the perturbative side, threshold resummation is found to play some role, especially for the lower energy data. A non negligible contribution is also found for intrinsic transverse momentum, whose effect we estimated through a phenomenological model, such to simulate power corrections to collinear factorization. However, the size of these phenomena is not yet sufficient to correct for the mismatch with data. Future studies smight work

---

[1] Other studies were provided in [29, 30], where however a modified prescription was used, different from the standard CSS formalism.

on the possible connections between Drell-Yan and SIDIS. Keeping in mind the possibility of experimental errors, especially in the presently available Drell-Yan data, a more formal analysis of power corrections, and–as suggested in [33]–a re-analysis of parton distribution functions and fragmentation functions, are possible directions.

In Ch. 6, we have presented calculations in collinear-factorization of single spin asymmetries in SIDIS, for the case of polarized target. This corresponds to having a high-q_T prediction for the F_{UL} structure functions (see Ch. 2), while to date it was only calculated in the TMD framework, if we exclude the unpublished work reported in [156], where the asymmetries where indirectly calculated by applying crossing symmetry to e^+e^- amplitudes. In this thesis we showed a direct calculation of the relevant Feynman diagrams at 1-loop, pointing out a sign error in both the $\sin\phi$ and $\sin 2\phi$ modulations of [156]. Azimuthal modulations for a future Electron-Ion Collider (EIC) are predicted to have a strong x dependence, and to be almost negligible for π^- produced off a proton target, whereas, in the case of π^+, a positive (negative) analyzing power for $\sin\phi$ ($\sin 2\phi$) is foreseen, reaching values around 1% at the highest measurable x and P_{hT}. In the fixed-target regime, it is interesting to observe that our calculations show an approximate agreement with the size and shape of the asymmetries measured by HERMES, which are commonly interpreted in terms of TMDs. This signals that matching corrections to TMD factorization should not be overlooked in studies of longitudinal spin asymmetries.

T-odd asymmetries in collinear factorization are also an interesting subject per se, as they are a pure loop effect in QCD. Despite being hunted for decades, they still lack clean experimental comparisons. For the future, it will be interesting to extend our study to analogous asymmetries in hadron collisions at the Relativistic Heavy Ion Collider (RHIC).

Acknowledgments

These three years of research, lectures and travels have been a period of deep personal growth, and this is thanks to the continuous contact with amazing people.

To start with, I want to thank Alessandro Bacchetta. I think it would be impossible, for anyone who met him, not to be impressed by his outstanding human qualities, and not to be influenced by his incredible positivity. Besides his valuable teachings about specific scientific topics and how to conduct research, he showed me some of the most important lessons a person can learn: how it is possible to be humble and self-confident at the same time; what it means to really respect—and listen to—other people, thus building constructive discussions; how to accept my own and other people's shortcomings and have a good laugh; how to leverage the positive side of an apparent obstacle to overcome it...

A similar, astonishing, attitude I found in Barbara Pasquini. She has been patiently advising me over the years, since my bachelor studies, from time to time knocking at my door just to see how I was feeling. She has been always so incredibly sweet... and funny!

My thesis would never have come to an end without the help of the amazing researchers and professors here in the theory corridor, that is, without the experience and wise advice of Valerio Bertone, Giuseppe Bozzi, Miguel Echevarria, Christian Pisano, Marco Radici. I am deeply grateful to you all!

My sincere thanks go to prof. Lucio Andreani, who has been always taking care of everything concerning the doctoral school, with the greatest dedication and heart.

I cannot imagine how my PhD life would have been without my fellows: Luca, Filippo, Stefano, Simone, Francesco Giovanni, Nicola, Francesco... thanks for all the lovely moments. And special thanks to my super office mate, Chiara: the office has become a brighter place, since you came with your contagious laugh, your incredible empathy and sweetness.

Meeting colleagues around the world, and sharing ideas and problems with them, is a fundamental part of research, and a source of human enrichment. I am really

glad that I met people like Kyle, Andrea Signori and all the other young students and researchers that have been so inspiring to me. Special thanks go to my collaborator in Tübingen, Maurizio, for being a unique example of an authentic person, always so welcoming, and moved by pure curiosity and enthusiasm.

It has been a great experience to meet prof. Werner Vogelsang, a person of vast knowledge and great humility, who makes young students feel encouraged, and patiently discuss with them in a friendly mood. I am deeply grateful to him for giving me the opportunity of a beautiful research stay.

Un ringraziamento speciale va alla mia famiglia, mamma, papà e Fabio, per avermi fatto apprendere e interiorizzare ciò che dà un senso ai rapporti umani, senza il quale la vita, e ogni nostra attività, perde di sapore: l'affetto sincero.

Un pensiero e un abbraccio non può che andare a tutte le persone conosciute a Pavia, al coro, agli amici jazzisti, e agli amici di una vita: Aurora, Bose, Diego, Martino, Rocco, Saypen... sono la prova di quanto i legami possano durare e maturare nel tempo, alla scoperta sempre più approfondita delle persone care.

Bibliography

[1] E. D. Bloom *et al.*, Phys. Rev. Lett. **23**, 930 (1969).

[2] M. Breidenbach *et al.*, Phys. Rev. Lett. **23**, 935 (1969).

[3] K. Cichy and M. Constantinou, Adv. High Energy Phys. **2019**, 3036904 (2019), arXiv:1811.07248.

[4] S. Dulat *et al.*, Phys. Rev. **D93**, 033006 (2016), arXiv:1506.07443.

[5] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, Eur. Phys. J. **C75**, 204 (2015), arXiv:1412.3989.

[6] NNPDF, R. D. Ball *et al.*, Eur. Phys. J. **C77**, 663 (2017), arXiv:1706.00428.

[7] S. Bhattacharya, A. Metz, and J. Zhou, Phys. Lett. **B771**, 396 (2017), arXiv:1702.04387.

[8] S. Bhattacharya, A. Metz, V. K. Ojha, J.-Y. Tsai, and J. Zhou, (2018), arXiv:1802.10550.

[9] A. V. Belitsky and A. V. Radyushkin, Phys. Rept. **418**, 1 (2005), arXiv:hep-ph/0504030.

[10] D. E. Soper, Phys. Rev. Lett. **38**, 461 (1977).

[11] R. P. Feynman, R. D. Field, and G. C. Fox, Phys. Rev. **D18**, 3320 (1978), [,739(1978)].

[12] S. D. Drell and T.-M. Yan, Phys. Rev. Lett. **25**, 316 (1970), [Erratum: Phys. Rev. Lett.25,902(1970)].

[13] A. S. Ito *et al.*, Phys. Rev. **D23**, 604 (1981).

[14] D. Antreasyan *et al.*, Phys. Rev. Lett. **45**, 863 (1980).

[15] Physics Letters B **76**, 356 (1978).

[16] J. C. Collins, D. E. Soper, and G. F. Sterman, Nucl. Phys. **B250**, 199 (1985).

[17] C. Davies, B. Webber, and W. Stirling, Nuclear Physics B **256**, 413 (1985).

[18] F. Landry, R. Brock, G. Ladinsky, and C. P. Yuan, Phys. Rev. **D63**, 013004 (2001), arXiv:hep-ph/9905391.

[19] F. Landry, R. Brock, P. M. Nadolsky, and C. P. Yuan, Phys. Rev. **D67**, 073016 (2003), arXiv:hep-ph/0212159.

[20] A. V. Konychev and P. M. Nadolsky, Phys. Lett. **B633**, 710 (2006), arXiv:hep-ph/0506225.

[21] P. Nadolsky, D. R. Stump, and C.-P. Yuan, Phys. Rev. D **61**, 014003 (1999).

[22] P. M. Nadolsky, D. R. Stump, and C. P. Yuan, Phys. Rev. **D64**, 114011 (2001), arXiv:hep-ph/0012261.

[23] A. Signori, A. Bacchetta, M. Radici, and G. Schnell, Journal of High Energy Physics **2013**, 194 (2013).

[24] M. Anselmino, M. Boglione, J. O. Gonzalez H., S. Melis, and A. Prokudin, Journal of High Energy Physics **2014**, 5 (2014).

[25] A. Bacchetta, F. Delcarro, C. Pisano, M. Radici, and A. Signori, JHEP **06**, 081 (2017), arXiv:1703.10157.

[26] J. Collins, Cambridge, UK: Univ. Pr. (2011) 624 p (2013).

[27] M. G. Echevarría, A. Idilbi, and I. Scimemi, Journal of High Energy Physics **2012**, 2 (2012).

[28] M. Echevarria, *Definition and properties of Transverse Momentum Distributions*, PhD thesis, Universidad Complutense de Madrid, 2013.

[29] I. Scimemi and A. Vladimirov, Eur. Phys. J. **C78**, 89 (2018), arXiv:1706.01473.

[30] V. Bertone, I. Scimemi, and A. Vladimirov, JHEP **06**, 028 (2019), arXiv:1902.08474.

[31] W. Bizoń *et al.*, Journal of High Energy Physics **2018**, 132 (2018).

[32] M. Boglione, J. O. Gonzalez Hernandez, S. Melis, and A. Prokudin, JHEP **02**, 095 (2015), arXiv:1412.1383.

[33] J. O. Gonzalez-Hernandez, T. C. Rogers, N. Sato, and B. Wang, Phys. Rev. **D98**, 114005 (2018), arXiv:1808.04396.

[34] D. W. Sivers, Phys. Rev. **D41**, 83 (1990).

[35] S. J. Brodsky, D. S. Hwang, and I. Schmidt, Phys. Lett. **B530**, 99 (2002), arXiv:hep-ph/0201296.

[36] A. Bacchetta and M. Radici, Phys. Rev. Lett. **107**, 212001 (2011).

[37] The HERMES Collaboration, A. Airapetian *et al.*, Phys. Rev. Lett. **84**, 4047 (2000).

[38] HERMES Collaboration, A. Airapetian *et al.*, Phys. Rev. D **64**, 097101 (2001).

[39] A. Airapetian *et al.*, Physics Letters B **562**, 182 (2003).

[40] M. Boglione and P. Mulders, Physics Letters B **478**, 114 (2000).

[41] A. Efremov, K. Goeke, and P. Schweitzer, Physics Letters B **522**, 37 (2001).

[42] A. V. Efremov, K. Goeke, and P. Schweitzer, The European Physical Journal C - Particles and Fields **32**, 337 (2003).

[43] A. Efremov, K. Goeke, and P. Schweitzer, Physics Letters B **568**, 63 (2003).

[44] B.-Q. Ma, I. Schmidt, and J.-J. Yang, Phys. Rev. D **66**, 094001 (2002).

[45] J. V. Guerrero, J. J. Ethier, A. Accardi, S. W. Casper, and W. Melnitchouk, JHEP **09**, 169 (2015), arXiv:1505.02739.

[46] J. V. Guerrero and A. Accardi, Phys. Rev. D **97**, 114012 (2018).

[47] CTEQ, R. Brock *et al.*, Rev. Mod. Phys. **67**, 157 (1995).

[48] J. Collins and J.-W. Qiu, Phys. Rev. D **75**, 114014 (2007).

[49] T. C. Rogers and P. J. Mulders, Phys. Rev. D **81**, 094006 (2010).

[50] A. Bacchetta *et al.*, JHEP **02**, 093 (2007), arXiv:hep-ph/0611265.

[51] A. Bacchetta, D. Boer, M. Diehl, and P. J. Mulders, JHEP **08**, 023 (2008), arXiv:0803.0227.

[52] R. Meng, F. I. Olness, and D. E. Soper, Phys. Rev. D **54**, 1919 (1996).

[53] P. M. Nadolsky, D. R. Stump, and C. P. Yuan, Phys. Rev. **D61**, 014003 (2000), arXiv:hep-ph/9906280, [Erratum: Phys. Rev.D64,059903(2001)].

[54] X. Ji, J.-P. Ma, and F. Yuan, Phys. Rev. D **71**, 034005 (2005).

[55] P. J. Mulders and R. D. Tangerman, Nucl. Phys. **B461**, 197 (1996), arXiv:hep-ph/9510301, [Erratum: Nucl. Phys.B484,538(1997)].

[56] D. Sivers, Phys. Rev. D **43**, 261 (1991).

[57] J. Collins, Nuclear Physics B **396**, 161 (1993).

[58] J. C. Collins, Acta Phys. Polon. **B34**, 3103 (2003), arXiv:hep-ph/0304122.

[59] J. C. Collins, Phys. Lett. **B536**, 43 (2002), arXiv:hep-ph/0204004.

[60] D. Boer and P. J. Mulders, Phys. Rev. **D57**, 5780 (1998), arXiv:hep-ph/9711485.

[61] D. Boer and W. Vogelsang, Phys. Rev. **D74**, 014004 (2006), arXiv:hep-ph/0604177.

[62] S. Arnold, A. Metz, and M. Schlegel, Phys. Rev. D **79**, 034005 (2009).

[63] R. D. Tangerman and P. J. Mulders, Phys. Rev. D **51**, 3357 (1995).

[64] S. M. Aybat and T. C. Rogers, Phys. Rev. **D83**, 114042 (2011), arXiv:1101.5057.

[65] J. Collins and T. C. Rogers, Phys. Rev. D **96**, 054011 (2017).

[66] M. G. Echevarria, I. Scimemi, and A. Vladimirov, Journal of High Energy Physics **2016**, 4 (2016).

[67] A. Bacchetta, M. G. Echevarria, P. J. G. Mulders, M. Radici, and A. Signori, JHEP **11**, 076 (2015), arXiv:1508.00402.

[68] J. Collins and T. Rogers, Phys. Rev. D **91**, 074020 (2015).

[69] I. Scimemi and A. Vladimirov, Journal of High Energy Physics **2018**, 3 (2018).

[70] K. G. Wilson, Phys. Rev. **179**, 1499 (1969).

[71] P. B. Arnold and R. P. Kauffman, Nuclear Physics B **349**, 381 (1991).

[72] G. Bozzi, S. Catani, D. de Florian, and M. Grazzini, Nuclear Physics B **737**, 73 (2006).

[73] J. Collins *et al.*, Phys. Rev. **D94**, 034014 (2016), arXiv:1605.00671.

[74] U. D'Alesio, M. G. Echevarria, S. Melis, and I. Scimemi, JHEP **11**, 098 (2014), arXiv:1407.3311.

[75] M. G. Echevarria, T. Kasemets, J.-P. Lansberg, C. Pisano, and A. Signori, Phys. Lett. **B781**, 161 (2018), arXiv:1801.01480.

[76] G. Lustermans, J. K. L. Michel, F. J. Tackmann, and W. J. Waalewijn, (2019), arXiv:1901.03331.

[77] G. Bozzi, S. Catani, G. Ferrera, D. de Florian, and M. Grazzini, Nucl. Phys. **B815**, 174 (2009), arXiv:0812.2862.

[78] W. Bizon *et al.*, (2018), arXiv:1805.05916.

[79] The H1 Collaboration and C. Adloff et al., The European Physical Journal C - Particles and Fields **12**, 595 (2000).

[80] ZEUS Collaboration *et al.*, Zeitschrift für Physik C Particles and Fields **70**, 1 (1996).

[81] HERMES Collaboration, A. Airapetian *et al.*, Phys. Rev. D **87**, 074029 (2013).

[82] C. Adolph *et al.*, The European Physical Journal C **73**, 2531 (2013).

[83] COMPASS Collaboration, M. Aghasyan *et al.*, Phys. Rev. D **97**, 032006 (2018).

[84] D. de Florian, M. Stratmann, and W. Vogelsang, Phys. Rev. D **57**, 5811 (1998).

[85] V. Bertone, PoS **DIS2017**, 201 (2018), arXiv:1708.00911.

[86] I. Scimemi and A. Vladimirov, (2019), arXiv:1912.06532.

[87] M. G. Echevarria, A. Idilbi, Z.-B. Kang, and I. Vitev, Phys. Rev. **D89**, 074013 (2014), arXiv:1401.5078.

[88] G. Moreno *et al.*, Phys. Rev. **D43**, 2815 (1991).

[89] PHENIX, C. Aidala *et al.*, Phys. Rev. **D99**, 072003 (2019), arXiv:1805.02448.

[90] CDF, T. Affolder *et al.*, Phys. Rev. Lett. **84**, 845 (2000), arXiv:hep-ex/0001021.

[91] CDF, T. Aaltonen *et al.*, Phys. Rev. **D86**, 052010 (2012), arXiv:1207.7138.

[92] D0, B. Abbott *et al.*, Phys. Rev. **D61**, 032004 (2000), arXiv:hep-ex/9907009.

[93] D0, V. M. Abazov *et al.*, Phys. Rev. Lett. **100**, 102002 (2008), arXiv:0712.0803.

[94] ATLAS, G. Aad *et al.*, JHEP **09**, 145 (2014), arXiv:1406.3660.

[95] ATLAS, G. Aad *et al.*, Eur. Phys. J. **C76**, 291 (2016), arXiv:1512.02192.

[96] CMS, S. Chatrchyan *et al.*, Phys. Rev. **D85**, 032002 (2012), arXiv:1110.4973.

[97] CMS, V. Khachatryan *et al.*, JHEP **02**, 096 (2017), arXiv:1606.05864.

[98] LHCb, R. Aaij *et al.*, JHEP **08**, 039 (2015), arXiv:1505.07024.

[99] LHCb, R. Aaij *et al.*, JHEP **01**, 155 (2016), arXiv:1511.08039.

[100] LHCb, R. Aaij *et al.*, JHEP **09**, 136 (2016), arXiv:1607.06495.

[101] S. Catani and M. Grazzini, Phys. Rev. Lett. **98**, 222002 (2007), arXiv:hep-ph/0703012.

[102] S. Catani, L. Cieri, G. Ferrera, D. de Florian, and M. Grazzini, Phys. Rev. Lett. **103**, 082001 (2009), arXiv:0903.2120.

[103] R. D. Ball *et al.*, JHEP **04**, 125 (2013), arXiv:1211.5142.

[104] G. D'Agostini, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **346**, 306 (1994).

[105] A. Bacchetta *et al.*, Phys. Rev. D **100**, 014018 (2019).

[106] G. Bozzi, S. Catani, G. Ferrera, D. de Florian, and M. Grazzini, Phys. Lett. **B696**, 207 (2011), arXiv:1007.2351.

[107] T. Becher, M. Neubert, and D. Wilhelm, JHEP **02**, 124 (2012), arXiv:1109.6027.

[108] S. Catani, L. Cieri, G. Ferrera, D. de Florian, and M. Grazzini, Phys. Rev. Lett. **103**, 082001 (2009), arXiv:0903.2120.

[109] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, Comput. Phys. Commun. **182**, 2388 (2011), arXiv:1011.3540.

[110] S. Dulat *et al.*, Phys. Rev. **D93**, 033006 (2016), arXiv:1506.07443.

[111] NuSea, E. A. Hawker *et al.*, Phys. Rev. Lett. **80**, 3715 (1998), arXiv:hep-ex/9803011.

[112] J. C. Webb, *Measurement of continuum dimuon production in 800-GeV/C proton nucleon collisions*, PhD thesis, New Mexico State U., 2003, arXiv:hep-ex/0301031.

[113] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, Eur. Phys. J. **C63**, 189 (2009), arXiv:0901.0002.

[114] NNPDF, R. D. Ball *et al.*, Eur. Phys. J. **C77**, 663 (2017), arXiv:1706.00428.

[115] H.-L. Lai *et al.*, Phys. Rev. **D82**, 074024 (2010), arXiv:1007.2241.

[116] R. D. Ball *et al.*, Nucl. Phys. **B867**, 244 (2013), arXiv:1207.1303.

[117] D. Antreasyan *et al.*, Phys. Rev. Lett. **47**, 12 (1981).

[118] D. Antreasyan et al., Phys. Rev. Lett. **48**, 302 (1982).

[119] S. Gavin *et al.*, Int. J. Mod. Phys. **A10**, 2961 (1995), arXiv:hep-ph/9502372.

[120] A. Szczurek and G. Slipek, Phys. Rev. **D78**, 114007 (2008), arXiv:0808.1360.

[121] K. Kovarik *et al.*, Phys. Rev. **D93**, 085037 (2016), arXiv:1509.00792.

[122] D. de Florian and W. Vogelsang, Phys. Rev. **D71**, 114004 (2005), arXiv:hep-ph/0501258.

[123] P. Hinderer, F. Ringer, G. Sterman, and W. Vogelsang, (2018), arXiv:1812.00915.

[124] N. Kidonakis and V. Del Duca, Phys. Lett. **B480**, 87 (2000), arXiv:hep-ph/9911460.

[125] N. Kidonakis and A. Sabio Vera, JHEP **02**, 027 (2004), arXiv:hep-ph/0311266.

[126] R. J. Gonsalves, N. Kidonakis, and A. Sabio Vera, Phys. Rev. Lett. **95**, 222001 (2005), arXiv:hep-ph/0507317.

[127] N. Kidonakis and R. J. Gonsalves, Phys. Rev. **D89**, 094022 (2014), arXiv:1404.4302.

[128] C. Muselli, S. Forte, and G. Ridolfi, JHEP **03**, 106 (2017), arXiv:1701.01464.

[129] D. de Florian, A. Kulesza, and W. Vogelsang, JHEP **02**, 047 (2006), arXiv:hep-ph/0511205.

[130] J. Huston *et al.*, Phys. Rev. **D51**, 6139 (1995), arXiv:hep-ph/9501230.

[131] U. D'Alesio and F. Murgia, Phys. Rev. **D70**, 074009 (2004), arXiv:hep-ph/0408092.

[132] L. Apanasevich *et al.*, Phys. Rev. **D59**, 074007 (1999), arXiv:hep-ph/9808467.

[133] E. Laenen, G. F. Sterman, and W. Vogelsang, Phys. Rev. **D63**, 114018 (2001), arXiv:hep-ph/0010080.

[134] E. Laenen, G. F. Sterman, and W. Vogelsang, Phys. Rev. Lett. **84**, 4296 (2000), arXiv:hep-ph/0002078.

[135] H.-N. Li, Phys. Lett. **B454**, 328 (1999), arXiv:hep-ph/9812363.

[136] M. A. Kimber, A. D. Martin, and M. G. Ryskin, Eur. Phys. J. **C12**, 655 (2000), arXiv:hep-ph/9911379.

[137] G. F. Sterman and W. Vogelsang, Phys. Rev. **D71**, 014013 (2005), arXiv:hep-ph/0409234.

[138] J. F. Owens, Rev. Mod. Phys. **59**, 465 (1987).

[139] M. Boglione, S. Melis, and A. Prokudin, Phys. Rev. D **84**, 034033 (2011).

[140] M. Anselmino, M. Boglione, J. O. Gonzalez Hernandez, S. Melis, and A. Prokudin, JHEP **04**, 005 (2014), arXiv:1312.6261.

[141] G. Karpman, R. Leonardi, and F. Strocchi, Phys. Rev. **174**, 1957 (1968).

[142] F. Cannata, R. Leonardi, and F. Strocchi, Phys. Rev. D **1**, 191 (1970).

[143] A. Efremov and O. Teryaev, Sov. J. Nucl. Phys. **36**, 140 (1982).

[144] A. Efremov and O. Teryaev, Sov. J. Nucl. Phys. **39**, 962 (1984).

[145] J. Qiu and G. Sterman, Phys. Rev. Lett. **67**, 2264 (1991).

[146] J.-w. Qiu and G. F. Sterman, Phys. Rev. **D59**, 014004 (1999), arXiv:hep-ph/9806356.

[147] A. D. Rujula, R. Petronzio, and B. Lautrup, Nuclear Physics B **146**, 50 (1978).

[148] K. Hagiwara, K.-i. Hikasa, and N. Kai, Phys. Rev. D **27**, 84 (1983).

[149] J. G. Korner, B. Melic, and Z. Merebashvili, Phys. Rev. **D62**, 096011 (2000), arXiv:hep-ph/0002302.

[150] R. Frederix, K. Hagiwara, T. Yamada, and H. Yokoya, Phys. Rev. Lett. **113**, 152001 (2014).

[151] L. H. Ryder, (1985).

[152] J. Körner, G. Kramer, G. Schierholz, K. Fabricius, and I. Schmitt, Physics Letters B **94**, 207 (1980).

[153] K. Fabricius, I. Schmitt, G. Kramer, and G. Schierholz, Phys. Rev. Lett. **45**, 867 (1980).

[154] K. Hagiwara, K.-i. Hikasa, and N. Kai, Phys. Rev. Lett. **52**, 1076 (1984).

[155] K. Hagiwara, K.-i. Hikasa, and N. Kai, Phys. Rev. Lett. **47**, 983 (1981).

[156] M. Aicher, *Threshold resummation effects on the parton distribution function of the pion and time-reversal-odd single-spin asymmetries*, PhD thesis, University of Regensburg, 2011.

[157] J. Collins, Computer Physics Communications **74**, 265 (1993).

[158] H. H. Patel, Computer Physics Communications **197**, 276 (2015).

[159] T. Kinoshita, J. Math. Phys. **3**, 650 (1962).

[160] T. D. Lee and M. Nauenberg, Phys. Rev. **133**, B1549 (1964).

[161] D. de Florian, R. Sassot, M. Stratmann, and W. Vogelsang, Phys. Rev. Lett. **101**, 072001 (2008).

[162] D. de Florian, R. Sassot, M. Stratmann, and W. Vogelsang, Phys. Rev. D **80**, 034030 (2009).

[163] D. "de Florian, R. Sassot, M. Epele, R. J. Hernández-Pinto, and M. Stratmann, Phys. Rev. **D91**, 014035 (2015), arXiv:1410.6027.

[164] M. Diehl and S. Sapeta, The European Physical Journal C - Particles and Fields **41**, 515 (2005).

[165] A. Airapetian *et al.*, Physics Letters B **622**, 14 (2005).

[166] V. Barone, A. Prokudin, and B.-Q. Ma, Phys. Rev. **D78**, 045022 (2008), arXiv:0804.3024.

[167] E. C. Aschenauer, I. Borsa, R. Sassot, and C. Van Hulse, Phys. Rev. D **99**, 094004 (2019).

[168] A. Bacchetta *et al.*, (2019), arXiv:1912.07550 (submitted to JHEP).