Contributions to Transverse Momentum Dependent Parton Distribution Functions in Aspects of Theory, Phenomenology and Relation to Event Generators

DISSERTATION ZUR ERLANGUNG DES DOKTORGRADES DER NATURWISSEN-SCHAFTEN (DR. RER. NAT.) DER FAKULTÄT FÜR PHYSIK

DER UNIVERSITÄT REGENSBURG

vorgelegt von

Valentin Moos aus

Passau

im Jahr 2024

Promotionsgesuch wurde eingereicht am 4. Juni 2024
Die Arbeit wurde angeleitet von Dr. Pia Zurita

Prüfungsausschuss:
Prof. Dr. Vladimir Braun
Prof. Dr. John Schliemann
Prof. Dr. Jascha Repp

Datum Promotionskolloquium: 26. Juli 2024

ERKLÄRUNG

zur Anleitung der Dissertation

Die vorliegende Dissertation ist unter Anleitung von

Dr. Pia Zurita

der Universität Regensburg angefertigt worden.

Regensburg, den 05.06.2024

Valentin Moos

Contents

I Introduction

To test nature at small time and spatial scales, high energies and momenta are needed. This is a reason for why particle accelerators with increasing collision energies have been built in the last (almost) hundred years, with one of the first examples in the early 1930s [1] in which mercury ions could be accelerated over a voltage of 210 keV. Modern examples of colliders are the Hadron-Elektron-Ring-Anlage (HERA), the to-be-built Electron-Ion Collider (EIC), or the Large Hadron Collider (LHC) at CERN, the latter of which providing us with collision energies up to 13 TeV. With these experiments it is possible to validate theories like Quantum Chromo Dynamics (QCD) by comparing experimental results with theoretical predictions in scatterings of the constituents of hadrons, induced by high energy collisions of the later.

In the physical regime of high momentum transfers, the scattering involving hadrons can be described in perturbative QCD. At very high energies it is a good approximation to compute the scattering with an exchange of only a few gluons. This is the case, because each interaction with gluons is multiplied with the coupling constant g, which appears to be small at high energy transfer. The theoretical description of these scattering events in quantum field theory follows rather simple principles. Any computation in the high energy regime for QCD is thus typically a series, a function expanded in the coupling constant. This method is in principle the same as in Quantum Electrodynamics (QED), where the equivalent to g would be the electric charge e. One significant difference to that case is that the gluons are allowed to self-interact, in constrast to the photon which carries no electrical charge and can only interact with other electrically charged particles. Thus the coupling structure becomes more complicated in QCD, and the computation of higher corrections, i.e., terms with multiple interaction points and thus a larger power in the coupling g, are in principle straight forward to compute, but in reality it becomes very involved at higher powers very quickly.

We assume that at these high energies, the constituents (partons) of the colliding hadrons are free particles. The continuous potential of classical physics is replaced by a quantised exchange of energy and momentum among the partons, mediated by bosons. In the case of Electrodynamics, these bosons are photons – particles, which we can also observe in other physical regimes. In the strong interaction QCD the bosons are called gluons and, like the quarks, they cannot be detected at the small energy scales where matter is bound in hadrons. With gluons and quarks, we can associate a physical property called “colour”, of which three states can be occupied by quarks, and eight by gluons. These properties cannot be observed directly, but can be assumed to describe the formation of bound states. The hadrons are quantum mechanical states, which are invariant under rotations in the Hamilton space of colour states. Thus, these physically bound states are also labelled as “colourless”.

Because the quarks are not stable themselves, in nature they are bound to hadrons, and the scattering of quarks alone is not enough to describe hadron collisions. In order to obtain quarks as quasi-initial state for the scattering, in the parton model the partonic cross section – the scattering between the partons as quasi initial and final partices – is modified by an amplitude which contributes the probability of finding a certain parton inside a hadron. Then the physical cross section becomes a convolution of this probability distribution and the partonic cross section.

It was Bjorken who was able to obtain this naive model in a high energy limit for the process of deeply inelastic scattering (DIS) of an electron off a proton, arguing that, in an collision so brutal and imminent, the structure of the collision could only depend on dimensionles quantities in the reaction [2]. In 1990 the Nobel prize was awarded to Friedman, Kendall and Taylor for the experimental measurement of inelastic scattering events off pointlike constituents of matter (partons), and their measurements can be explained by Bjorken’s considerations. By now, many scattering processes have been formulated as a convolution between hadronic matrix elements of quark (and gluon) field operators, so-called parton distribution functions (PDFs), and a hard partonic cross section. This successful technique is called factorisation [3], as the physical cross section can be written as the product (convolution) of two terms. These two ingredients are very different in their nature. The partonic cross section can, in principle, be computed up to any desired order in the regime of perturbative physics, and it does depend on the examined process. On the other hand, the PDFs do not depend on the particular process; they are universal in the sense that they carry the information momentum distribution inside hadrons, which is independent from the reaction we put them into. However, they cannot be computed in perturbation theory.

The distribution of quarks (and gluons) inside a proton is the same if we scatter it with an electron, or another proton. It thus allows us to obtain the PDFs as a property of the hadron from a single process, and validate it by testing the theoretical prediction for another experiment (with another process) where the same distribution is relevant.

Nonetheless there is more than one approach to describe the distribution of partons inside a hadron. Even though the PDFs as a universal property of hadrons do not depend on the process, what does depend on the process is which kind of distributions of the partons contribute to the reaction. This does not only depend on the particles colliding, but crucially also on the question one is asking – by defining which variables are interesting. In the simplest example of DIS of an electron off a proton, information about the flavour and collinear momentum of the partons (PDFs) is enough to describe the process. The final state of the hadron is completely unobserved, and the partons momentum can only be probed by measuring the electrons scattering angle and energy-transfer.

Another process for which it is enough to consider the partons momenta to be collinear with the hadron, which characterises the PDFs, is the the by Drell and Yan (DY) suggested massive lepton pair production in dihadron collision. The lepton pair is the decay product of a short lived photon, the formation of which in the first place can be explained with a merging one quark of a hadron with its corresponding antiquark from the other hadron.

If we specify the question we are asking nature by changing the way we measure the scattering events, we find that it is useful to examine a more involved concept of parton distributions. For example, we can measure the total momentum of the produced lepton pair, and consider its component transverse to the collision axis. The transverse component can be generated by a real emission of energy in the collision, such that the quark-antiquark pair produces additional matter besides the lepton pair. In the case of large transverse momenta, indeed the description with collinear partons works very well [4]. For the observation of small transverse momenta of the lepton pair, it is required to consider intrinsic motion inside the hadron, taking place not only in the beam line, but in any direction. It is thus allowed for a parton to have a (typically small) component of transverse momentum, and if the transverse components of all partons add up to zero, the hadron as a whole propagates exactly along the beam line. Distributions taking into account this transverse motion are called transverse momentum dependent (TMD) PDFs. They can be understood as a generalization of a collinear PDFs, where now information about the momentums transverse component is taken into account.

The transverse momentum component is not the only generalization of collinear PDFs. Another aspect of the partons is their spatial distribution. As the hadrons collide at high velocities, viewed from the laboratory or from the collision partner, they are strongly Lorentz contracted. Their spatial extent in the beam direction is thus practically zero. The spatial distribution of the partons transverse to the beam line however is not affected by this Lorentz contraction, and is relevant for the physical process of deeply virtual Compton scattering (DVCS)[5]. The parton distributions in which also the spatial distribution is considered are called generalized parton distributions (GPDs).

The information of collinear momentum (PDFs), transverse momentum (TMDPDFs) and impact parameter (GPDs) is combined in so-called Wigner functions [6][7]. It is a naive understanding that TMDPDFs and GPDs can be derived by simply integrating over the transverse momentum or the impact parameter, respectively. The collinear PDFs would then the integration over both, such that only the collinear information remains. While not formally precise, this is a simple first picture to view the relations between the functions and their physical interpretation.

Even though the framework of TMD factorization is very successful in describing experimental data [8][9], TMDPDFs themselves can only be determined with a relatively large uncertainty, and to provide a realistic estimate for the theoretical uncertainties of TMDPDFs is an issue in modern phenomenology [10].

This thesis is focused on studying different aspects of TMD distributions. We consider theoretical computations, phenomenology of the non perturbative TMD distritutions, and their relation to software packages, called Monte Carlo event generators (MCEGs), which simulate scattering events as whole. In the following, we highlight those parts in **boldface** that have been worked on in detail in this thesis.

The factorization of observables implicitly defines operators for QCD which are important to study – because they contribute to measurable quantities via their related parton distributions. To this end we work out a solely perturbative aspect, by examining the TMD operator as a bare object. **We compute the ultra-violet (UV) divergent terms in one loop corrections to the TMD operators.** We present the corrections diagram by diagram, along which we discuss the developed methods to this computation.

Another major result **we present** in this work is **the extraction of TMD distributions** as non perturbatively calculable objects. In the standard approach, the TMDPDF is related to the collinear PDF by operator product expansion. To obtain the full function, the matching is modified by an ansatz. The ansatz is a model, parametrized in the variables x, the portion of collinear momentum carried by the parton relative to the hadron, and the transverse distance $\boldsymbol{b}$, which is the Fourier conjugated to the partons transverse momentum. The exact forms are obtained by fitting the ansatz against experimental data in the framework of TMD factorisation.

Like for collinear PDFs, the extraction of TMDPDFs from data has some history [11, 12] and is not a new approach. The novelty of the presented work is in two aspects. One aspect is the steadily increased perturbative accuracy [13]. One part of this accuracy is given as the power up to which the strong coupling is taken into account, i.e. after which order the expanded series are truncated. This regards the hard cross section, but also the evolution which TMD distributions obey, but due to the matching also the collinear PDF used, which is itself obtained at some perturbative order in the strong coupling. All this together defines the order at which the TMDPDFs are obtained. The extraction obtained in this work is at four loops in both the hard process scattering and the scale evolution of the TMDPDFs. The matching to collinear distributions is at three loops perturbative accuracy to a PDF set, which is obtained at two loops of accuracy [14]. We also consider fragmentation functions (FFs) in the TMD regime, for which we use the same matching procedure as for PDFs to connect our transverse momentum depenent FFs (TMDFFs) to the collinear fragmentation functions; the matching is to the same accuracy as it is for collinear PDFs, and the collinear FFs we use are obtained at two (pion) [15] and one (kaon) [16] loop(s).

The second aspect **we mention**, and perhaps **the physically** more **interesting novelty in this work,** is **the usage of a flavour dependent ansatz for the non perturbative TMDPDF model**. The TMDPDFs are flavour dependent already without this new approach, due to their relation to the flavour dependent collinear PDFs. With this ansatz, we are able to achieve a very good description of TMD relevant data of DY production with a very simple functional form. Using the flavour separation in our non perturbative ansatz, we are able to provide a very simple model for the TMDPDF and obtain a very good comparison to experimental data.

Even though the result for a fit to DY is very good, we observe a problem in treating quark and antiquark flavours separately in the region of small x. This issue comes to light once we consider the extracted TMDPDFs to describe the SIDIS process. We observe that the freedom in the original ansatz is a problem when describing SIDIS data. That is why for the further extractions we modify the parametrization, demanding that the quarks and their corresponding antiquark have the same behaviour in the region of small momentum fractions; the ansatz we then propose is flavour dependent nevertheless and retains its simple functional form.

Once we take into account these modifications on our ansatz for the TMDPDFs, **we** are able to **provide a good description of the SIDIS data**, which appeared to be a difficult task mostly due to the data located at the lowest energy resolution scales we access in our fit. Very interestingly, the same modification simultaneously improves the description of the experimental data on the pion-induced DY process, such that **we** can **provide a very good result for** such a fit as well, and along with this **an extraction of the pion TMDPDFs**. The improvements on both frontiers demonstrates the importance of this regulation in our ansatz, while the flavour separation in general remains a crucial aspect of the ansatz. All of this work has been concluded in the scope of this thesis, and we are able to provide first preliminary results for the extraction of TMDFFs from SIDIS and the pion TMDPDFs with our approach for the first time in this work.

In this thesis we also discuss the benefit TMD phenomenology can provide for algorithm design in Monte Carlo event generators (MCEGs). With MCEGs, software packages have been created which combine aspects of perturbative high energy physics, but also use intuition-based models to then model the generation of individual scattering events – the same processes that individually occur at particle colliders.

In (most) Monte Carlo Event generators collinear PDFs are used as an input. By sampling, respecting the parton distribution given, a quark with (or gluon) with a certain flavour and momentum is selected for a collision. The partons transverse momentum depends on various factors, one of which is the *primordial* k_T, the transverse momentum the parton has as a constituent of its hadron. To this end **we** discuss the performance of various MCEG software packages, and **use an intuition based approach – motivated by TMD phenomenology** – to suggest ways **to improve the generation routine for the transverse momentum component in MCEGs**.

This work is structured as follows. To give an overview of the field of phenomenology in particle physics, we discuss the fundamentals of perturbative QCD exemplary on the DIS process in chapter II. There we also review two relevant processes for TMD physics, the semi-inclusive DIS (SIDIS) and the DY process. For the latter reaction, we go through a derivation of the factorisation in the TMD regime, exemplary demonstrating how the cross section can be computed by perturbative means. The computation of the UV part of TMD operators can be found in chapter III, where we focus especially on the very recent results of gluon operators of twist-3. The phenomenological progress is contained in the chapter IV. There we presents a discussion on the framework (section A), the TMDPDF extraction ART23 (section B), such as the recently developed results, containing of a combined fit to we determine transverse momentum dependent fragmentation functions (TMDFFs) (section C), as well as an extraction of pion TMDPDFs and (section D). We review the data description of MCEGs for SIDIS processes and develop ansatzes to improve their performance as the last part of this work in chapter V.

II Relevant processes

In the following sections we explain the physical processes most relevant for this work, namely deeply inelastic scattering (DIS) of an electron off a proton, the process of semi-inclusive DIS (SIDIS) and the Drell-Yan process (DY).

For all mentioned processes we provide an analysis of the physical process, and explain how they are relevant for TMD physics. With the specifiy example of DIS we give a broad overview of the physics in high energy collisions. For the DY process we discuss the method of factorisation in more technical details. We demonstrate how the cross section for that process is derived in the kinematic regime of transverse momentum dependent distributions (TMDs). For all processes we list the factorisation formulae in terms of both collinear and TMD distributions.

A Deeply inelastic scattering

In this section we discuss the reaction of DIS, certainly one of the most relevant processes for perturbative QCD. DIS probes the hadron structure by examining the scattering of a lepton off a hadron. Like the Rutherford experiment is able to explore the substructure of atoms, DIS allows us to look inside nucleons. This is achieved by colliding hadrons at collision energies large enough to increase the chances an incoherent scattering between the lepton and the substructure of the hadron to occur - instead of a coherent scattering between lepton and hadron as a whole. The fundamental reaction allows for an ordering of detailed information about the structure of the hadron and its (more) fundamental constituents [2].

In the following we describe the process and illustrate how the electromagnetic scattering of a lepton off a nucleon is related to the strong force. We explain how the substructure of the hadron is encoded in structure functions and parton distribution functions (PDFs), and we show a factorised representation for the cross section of the DIS process. At the end of this section, we present phenomenological results for PDFs which we will later on require for TMDPDF determinations.

DIS is the reaction

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

between an initial state lepton l and a nucleon N to a final state lepton l' and an undefined final state of the hadron X, and the respective momenta are given in parenthesis. In the inelastic process the hadron in the initial state fragments due to the large momentum exchange between itself and the lepton transmitted by a boson. The momentum transfer can be mediated by either a photon (γ), a Z- or a $W^{\pm}$-boson. The process is depicted in figure 1, where the boson is depicted as photon(γ^*).

DIS is an inclusive reaction, and all the information about the event can be obtained by determination of the momentum of the outgoing lepton. The momentum distribution inside the unspecified hadronic remnant state X is not important, and the sum of all momenta inside X is constrained by conservation of total four-momentum in the process. I.e., two collisions with the same momenta in eq. (2.1), but a different decoposition of the momentum P' in potentially different hadrons are identical DIS reactions. We do not resolve of which hadrons the state X consists, and how the momentum P' is distributed among them.

The initial momenta are configured by the experiment. In many experimental setups the initial lepton either is an electron [17, 18] or a muon [19, 20]. If the exchanged particle is either γ or Z, both bosons are not charged and only momentum is transferred. Then the initial and the final state lepton are the same, and the scattered electron/muon can be detected directly to determine all momenta in eq. (2.1). For events in electron hadron collisions where a W-boson is exchanged, momentum and charge are transmitted in the reaction. The final state lepton then is

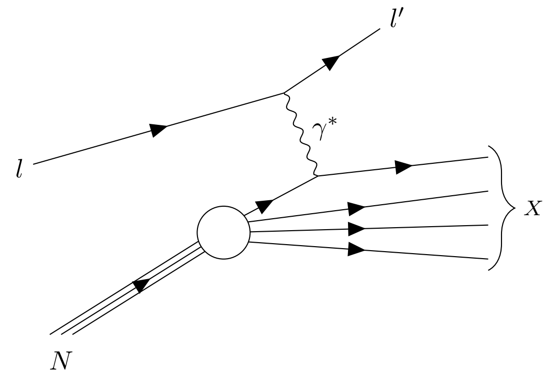


Figure 1: Schematic of DIS process. A lepton l scatters off one of the constituents of the hadron N via exchange of a virtual photon γ^*. The hadronic final state is denoted by X.

a neutrino which cannot be measured due to its weak interaction. These scatterings can be reconstructed by measuring the whole final state momentum P', such that the momentum of the neutrino can be determined by momentum conservation.

In what follows we restrict ourselves to the case where the in- and outgoing leptons are electrons and the boson is a photon. To this reaction the channel of a Z-boson contributes. However, we discard this contribution in the following[1]. The mass of the electron is negligible in this discussion. Furthermore we also take the hadron to be a proton, as it is the most common case. The reaction is analogous for other hadrons or leptons as colliding particles. We also restrict the following on unpolarized particles in the initial and final state.

We consider the reacting electron as a fundamental particle without substructure. The electron and the momentum transmitting photon are described precisely by quantum electro dynamics (QED), a well understood theory. Because of the definition of the reaction to an undefined hadronic final state, the proton only contributes in the initial state. To compute the process it thus is not necessary to describe the mechanism of the binding of the quarks back to physical bound states, after the initial hadron has been broken apart by the large momentum transfer onto a single quark.
This is why DIS can be pictured as the cleanest way of probing the structure of nucleons. To describe the reaction, a single hadronic state needs to be described. This gives us information about what we want to understand, while at the same time keeps the contribution of non-perturbative (NP) physics that need to be described at a minimal level. Due to the simple "question" formulated by the definition of DIS, the "answer" is as simple as it can be. Nevertheless it is non trivial to understand the process from a theoretical point of view.

The kinematic variables associated to an individual reaction, given as Lorentz invariants, are

$$q = k - k', \tag{2.2}$$
$$Q^2 = -q^2, \tag{2.3}$$
$$x_B = \frac{Q^2}{2q \cdot P}, \tag{2.4}$$
$$y = \frac{q \cdot P}{k \cdot P}, \tag{2.5}$$
$$\nu = q \cdot P, \tag{2.6}$$
$$W^2 = (q + P)^2. \tag{2.7}$$

[1]We neglect this contribution here simply for convenience; yet, even at high centre-of-mass energies of HERA, the pure Z-boson contribution can be neglected [17], and only the interference term is taken into account.

Here, q is the four momentum of the exchanged Boson and $Q^2 > 0$ its invariant mass (also called *virtuality* of the process). x_B is the Bjorken scaling variable. We discuss its meaning and relation to the parton model in the next section. In the laboratory frame, defined such that the proton is at rest, what typically is measured to reconstruct all kinematics are the final state energy of the electron (E') and its scattering angle (θ) relative to the initial electrons momentum. In the rest laboratory frame, with initial state electron of energy E, the interpretation of y/ν are the relative/absolute energy transfer from the electron to the proton. W is the invariant mass of the subsystem of the scattered nucleon.

The two variables measured in the hadrons rest frame can be also expressed by a combination of the two Lorentz scalars Q^2 and x_B. These are traditionally chosen as the two variables in which the DIS process is described.

The differential cross section for the scattering to any final hadron state is given by the relation between matrix element and differential cross section in any quantum field theory [21],

$$\mathrm{d}\sigma^{[e+N\to e+X]} = \frac{1}{2s}\frac{\mathrm{d}^3k'}{2\,|k'|}\int \mathrm{d}\Pi_X \left|\mathrm{i}\mathcal{M}\left(e(k)+N(P)\to e(k')+X(P)\right)\right|^2 \delta^{(4)}\left(P+q-P'\right). \tag{2.8}$$

Here $s = (P+k)^2$ is the invariant centre-of-mass energy, which is assumed to be much greater than the mass of the proton M. The final phase space integral over the states encoded in X is hidden in the integral over the phase space configuration $\mathrm{d}\Pi_X$.

Integrating over all hadronic final states X and the spin of the final state electron, and an averaging over the initial spin of the initial electron, the DIS cross section is

$$\mathrm{d}\sigma^{\mathrm{DIS}} = \frac{1}{2s}\frac{\mathrm{d}^3k'}{2|k'|}\frac{1}{2}\sum_{\mathrm{spin}}\sum_X\int \mathrm{d}\Pi_X \left|\mathrm{i}\mathcal{M}\left(eP\to eX\right)\right|(2\pi)^4\delta^4\left(P+q-P'\right). \tag{2.9}$$

Here the sum over X indicates the sum over any hadronic final state, and the integral indicates any momentum configuration corresponding to it.

The matrix element squared then can be organized such that there are two independent parts, a Lorentz structure for the leptonic part of the matrix element which origins from the rearrangement of the (massless) spinor fields, and the matrix elements in the hadronic tensor. This allows to present the differential cross section as the contraction of these two independent tensors,

$$\frac{\mathrm{d}\sigma^{\mathrm{DIS}}}{\mathrm{d}^3k'} = \frac{1}{2s}\frac{1}{2Q^4|k'|}L_{\mu\nu}W^{\mu\nu}(2\pi)^4\delta^4\left(P+q-P'\right). \tag{2.10}$$

The leptonic tensor in the considered spin average for initial electrons and sum over final electrons spin can be computed with Casimir's trick and for massless leptons is given by

$$L^{\mu\nu} = 2\left(k'^{\mu}k^{\nu} + k^{\mu}k'^{\nu} - k\cdot k' g^{\mu\nu}\right). \tag{2.11}$$

The hadronic tensor is

$$W^{\mu\nu}(p,q) = \frac{1}{2^2}\sum_X\int \mathrm{d}\Pi_X \left\langle P\right| J^{\dagger\mu}(0)\left|X\right\rangle\left\langle X\right| J^{\nu}(0)\left|P\right\rangle, \tag{2.12}$$

with a prefactor for the average over the protons spin and the helicity of the virtual photon. The operators J^μ are the electromagnetic interaction terms in the Lagrangian for quarks, divided by their charge and without contraction with the photon,

$$J^\mu = \overline{q}\gamma^\mu q. \tag{2.13}$$

The hadronic tensor can also be presented as [22]

$$W^{\mu\nu}(p,q) = \frac{1}{2}(2\pi)^2 \int \mathrm{d}^4 z e^{-\mathrm{i}qz} \langle P| \left[J^{\dagger\mu}(z), J^\nu(0) \right] |P\rangle\,, \tag{2.14}$$

where the brackets represent the commutator. This representation is more useful for the later discussed factorisation.

The hadronic tensor can be decomposed in a tensor of rank two, depending on the four momenta q, P and also the protons spin S.

Due to the QED Ward identity

$$q_\mu W^{\mu\nu} = q_\nu W^{\mu\nu} = 0, \tag{2.15}$$

the hadronic tensor can be parametrized into Lorentz structures obeying the above constraints. They depend on the Lorentz invariant combinations of the momenta P and q, where typically the hadrons mass $\sim P^2$ is neglected. Due to the Ward identity, the decomposition of the hadronic tensor in terms of structure functions can be rearranged to only the two functions $F_{1,2}$ and then is

$$W^{\mu\nu}(P,q) = F_1(x_B, Q^2)\left(g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2}\right) + F_2(x_B, Q^2)\left(P^\mu q^2 - P^\mu P\cdot q\right)\left(P^\nu q^2 - q^\nu P\cdot q\right). \tag{2.16}$$

The decomposition in eq. (2.16) is incomplete. There is an additional structure which is antisymmetric and reflects the dependence on the spin of the external state as well. This structure is omitted here because it does not contribute as it vanishes upon contraction with the symmetric leptonic tensor.

The differential cross section eq. (2.10) now can be written in terms of the structure functions and expressed with respect to the Lorentz invariants x_B and Q^2. For the discussed case of unpolarized DIS mediated only by a photon, the differential cross section is

$$\frac{\mathrm{d}\sigma^{\mathrm{DIS}}}{\mathrm{d}Q^2\mathrm{d}x_B} = \frac{4\pi\alpha^2}{Q^4}\left\{\left(x_B s - Q^2 - \frac{Q^2 x_B M^2}{s}\right) F_2(x_B, Q^2) + \frac{Q^4}{s} F_1(x_B, Q^2)\right\}. \tag{2.17}$$

From this form the structure functions can be obtained by analysing experimental data.

A.a Parton model and Bjorken limit

The parton model is a way to picture the proton or any hadron probed at sufficiently high energies. Its considerations require only fundamentals of special relativity. For a proton at rest, for an observer the signal time inside the proton is given by $\tau \sim R/c$ with the R the radius of the proton and c the speed of light. In analogy to the famous Gedankenexperiment by Einstein with a light-clock seemingly slowing down when moving relatively to an observer, also the signal time between partons is enhanced for an outside probe. In the theoretical limit where the proton travels at the speed of light, the interaction time $\tau \sim p_0 R/c$ becomes large enough to justify that any interaction between the partons is suppressed. The interactions come to a halt. Therefore, we approximate that the constituents of the proton do not interact with each other during the instance like event in which the photon strikes the proton.

This is the parton model [23][2]: In the picture the virtual photon strikes a *free* parton inside the proton. The proton is described as a collection of - for the short interaction time - non-interacting quarks and gluons. The gluon cannot interact with the photon directly, so we only take into account interactions with quarks and antiquarks in this limit. The cross section of the whole process then can be written as the sum of all partonic cross sections. This sum includes all quark flavours and the integral over the momentum fraction the quark carries with respect to the proton. As a result, in this model the differential cross section of the process between lepton and proton becomes the product of the partonic differential cross section and a probability density capturing the momentum distribution of quarks inside the proton f_f,

$$\frac{\mathrm{d}\sigma^{\text{DIS:parton model}}}{\mathrm{d}Q^2} = \sum_{f\in h} \int_0^1 \mathrm{d}x f_f(x) \sigma^{\text{partonic}}_{[e+q_f\to e+q_f]}(x, Q^2). \tag{2.18}$$

The partonic cross section is simply the interaction of the electron with a free parton, and the probability to find such a quark in the specific (momentum $\sim x$) configuration to enter the hard process is encoded in the function $f^f(x)$: The reaction is presented as the sum of the reactions of the substructure.

If we assume that the partons of the proton are quarks (and electrically chargeless gluons) the cross section by explicit computation of the partonic interaction in eq. (2.18) is [22]

$$\frac{\mathrm{d}\sigma^{\text{DIS:parton model}}}{\mathrm{d}\Omega \mathrm{d}E'} = \frac{\alpha^2}{4ME^2\sin^4(\theta/2)}\left[\frac{Mx_B}{\nu}\cos^2(\theta/2) + \sin^2(\theta/2)\right]\sum_f Q_f^2 f_f(x), \tag{2.19}$$

with the mass of the proton[3] M and the fraction of the elementary charge carried by the quark Q_f. For u type quark $Q_u = 2/3$, for d type quark $Q_d = -1/3$. Further the structure functions are now related to the parton densities:

$$F_2^{\text{parton}}(x_B, Q^2) = x_B \sum_f Q_f^2 f_f. \tag{2.20}$$

In the context of PDFs, x labels the momentum fraction between 0 and 1, carried by the parton. x_B as the Bjorken variable is defined as a fraction of invariants ($Q^2/2p\cdot q$). The Breit frame is defined as the reaction specific frame in which the electron elastically recoils and thus

$$q = (0, 0, 0, -Q). \tag{2.21}$$

Assuming massless partons, in this (unphysical) frame the definition of the momentum fraction x in eq. (2.18) matches x_B. This gives a physical interpretation to the ratio of invariants x_B.

The Bjorken limit is the formal limit of

$$\left\{Q^2, \nu\right\} \to \infty, \quad x_B = \frac{Q^2}{2\nu} = \text{fixed}. \tag{2.22}$$

Both scales must be significantly larger than the scales of bound hadrons such that the considerations of perturbative QCD are valid. The formal limit is substituted by the region $Q^2, \nu \gg \Lambda^2_{\text{QCD}}$, and we can expect that our computation is valid up to corrections $\sim \frac{\Lambda^2}{Q^2}, \frac{\Lambda^2}{\nu}$. It is argued by Bjorken that the scattering under these conditions cannot depend on any natural scale any longer, thus becomes dependent only on dimensionless quantities [2]. If one can describe

[2] The original paper from 1969 unfortunately is neither accessible via `iNSPIRE:HEP` nor via the Regensburger Katalog.

[3] This result is taken from [22] as indicated. Here we keep the term which is proportional to the mass M of the proton, which is neglected elsewhere.

the event in the invariants ν and Q^2, in the limit the structure functions depend only on a ratio of the two invariants $\sim x_B$,

$$F_{1,2}\left(\{Q^2,\nu\}\rightarrow\infty\right)=F_{1,2}\left(x_B=\frac{Q^2}{2\nu}\right). \tag{2.23}$$

This behaviour is called Bjorken scaling [2]. The assumption that the structure functions in this limit depend on this ratio of invariants is related to the point-like structure of partons, the motivation prior to the formulation of the parton model.

The last frame we want to discuss here is the *infinite momentum frame*. It helps to illustrate why exactly the light cone expansion of the currents in eq. (2.14) plays the dominant role in describing DIS. If the proton moves close to the speed of light and we neglect its mass, the momenta can be written down as

$$p=(p_0,0,0,p_0)\,,\quad q=(\nu/2p_0,\sqrt{Q^2},0,-\nu/2p_0),\quad p_0\approx\nu\rightarrow\infty \tag{2.24}$$

where the space is oriented such that the proton moves parallel to the z direction. From the Fourier integral eq. (2.14) it can further be argued, that in this limit the dominant region of the integral is given where by $p\cdot z\approx 0$. Elsewhere the contribution will be suppressed by a rapidly oscillating phase. This region is characterized by both

$$z_0\pm z_3\approxeq\nu^{-\frac{1}{2}},\quad z_1\approxeq\nu^{-\frac{1}{2}}. \tag{2.25}$$

The argument to also restrict the effectively contributing domain of the remaining (here: z_2) direction of the Fourier transform is, that otherwise a space-like distance in the commutator occurs in the commutator eq. (2.14); a contribution from this region violates locality [22].

Concluding from these considerations it follows that the dominant contribution to the hadronic matrix element is given by the region where the separation between the currents is light-like. This motivates the expansion of non local operators appearing in such a process along the light cone: The light-cone operator product expansion (OPE). A detailed analysis can be found in [22].

A.b Scale dependence of PDFs

In section A.a we discussed the important limits and contributions for high energy DIS, namely the Bjorken limit and the motivation for the parton model. All of the considerations there are at the leading order in QCD, i.e. quark-gluon interaction is neglected. This is true only in the limit eq. (2.22), which is never reached in experiment.

In this section we discuss what happens for scales at which the QCD effects cannot be neglected. They are important as corrections to the process themselves, but also for the properties of the in eq. (2.18) defined distributions. The dependence of the process on the physical resolution scale Q^2 can be seen as the continuous connection between two extrema. In the limit of $Q^2\rightarrow\infty$ we expect free partons. In the region of $Q^2\ll M^2$, on the other hand, the proton is better described as a fundamental particle without substructure. The wavelength of the virtual photon is too large to yield sensitivity to scales like the size of the nucleon and beyond.

Because in nature there are no free quarks, in any physical process we need to describe the emersion of a quark or gluon out of a hadron, and thus, as we describe the time evolution of a state in quantum mechanics, here we are required to describe the scale evolution from the original hadronic scale to the scale of the hard process where the parton is quasi free. We begin by recalling that by OPE the time ordered product of two electromagnetic currents

J in eq. (2.14) can be expanded as a series of local operators N_i with respective Wilson coefficients C_i,

$$TJ^\mu(x)J^\nu(0) = \sum_i C_i(x)N_i(0). \tag{2.26}$$

This is one technique that separates the propagation of quarks through the hadronic background field from the interaction with the photons at two vertices, illustrated in figure 2. It may be helpful to view this as the light-cone expansion of the object

$$J^\mu(x)J^\nu(0) \mapsto q(x)\gamma^\mu \left[\psi(x)\overline{\psi}(0)\right]\gamma^\nu q(0) = \sum_i q(x)C_i^{\mu\nu}(x)q(0), \tag{2.27}$$

where two of the fields that enter the operator are considered as slow fields (q), and the QCD quantum corrections are encoded in the coefficients C_i. It contains the propagation of the quark inside the background field of the proton. These two *external* quark fields) can be isolated from the corrections due to (loop) interactions of the quantum fields(ψ) and written into the forward amplitude for the proton. The proton then serves as source and sink for the interacting quark at two positions of light-like separation. This is the formal definition of PDFs, the forward matrix element light cone operator [24]

$$f_f(x) = \int \frac{\mathrm{d}z}{2\pi} e^{-\mathrm{i}xzpn} \langle P| \overline{\psi}(zn)^f \mathcal{P}\exp\left(-\mathrm{i}g\int_0^z \mathrm{d}tn^\mu A_\mu(tn)\right)\gamma^+\psi(0)^f |P\rangle \tag{2.28}$$

The appearance of the path ordered exponent can be seen as part of the definition to make this object gauge invariant. However, it also arises naturally by the full expansion eq. (2.27). This corresponds to interactions with the hadron via the exchange of gauge bosons. The light cone direction n here reflects the discussion of section A.a and g is the strong coupling parameter. The explicit Dirac structure $\gamma^+ = n^\mu\gamma_\mu$ is relevant in the discussed case of the unpolarized cross section.

We now sketch the appearance of on a virtual scale, and how the PDFs depend on it, while not going into many details. Technically there is dependence on a dimensionful parameter μ which enters by considering loop corrections. The corrections can be infrared (IR) or ultraviolet (UV) divergent. A common choice is to compute the loop corrections in $d = 4 - 2\epsilon$ dimensions. This regulates the divergences, and thus allows for the computation of individual parts. In the combination of all parts contributing to physical observables, like the cross section, or the coupling dependence on the scale, the dependence on the regularisation parameter ϵ cancels (order by order).

The scale dependence on μ remains as an artifact of perturbation theory. It is related to the separation of short distance physics (C_i) and long distance physics N_i in eq. (2.26); the bound state complexities then are contained in the matrix elements of the N_i [22]. We can picture the virtual scale as an artificial, but meaningful separation scale $\Lambda_{\mathrm{QCD}} < \mu < Q$. Processes with momentum scales below the separation scale μ can be taken to be part of the hadronic initial state, thus being part of the hadronic matrix element (∼PDF). Momenta larger than the separation scale are considered to be part of the hard partonic process.

To obtain the anomalous dimension $\gamma^{q\overline{q}}$ [25] for the operators in eq. (2.26), the loop corrections to the coefficients need to be computed order by order. Then we obtain the equations which govern the evolution of the parton densities. For this we refer to [22] or the previous reference.

From the anomalous dimension eventually the splitting kernels P can be derived, and the scale evolution of the PDF is given via the integral-equation

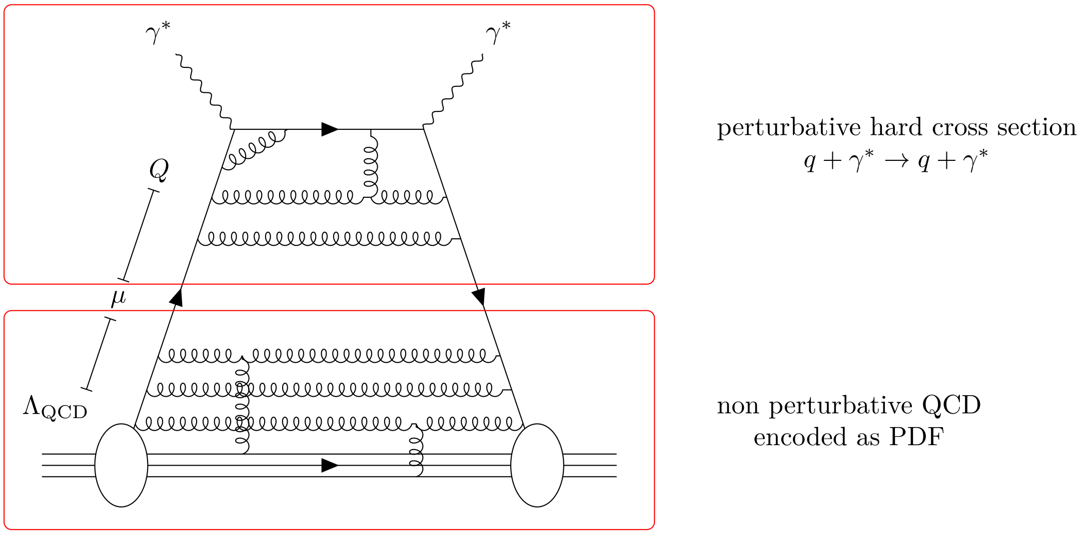


Figure 2: Schematic of DIS process via the forward scattering. A quark emerging from the hadron reacts with the virtual photon γ^*. The quark may radiate gluons at any level; gluon interaction also contributes with QCD corrections to the hard reaction. The separation between what is considered part of the hard process and the hadron propagation in the background is indicated by the upper and lower regions of the diagram.

$$\frac{\mu^2 \partial f_f^{\rm NS}(x,\mu^2)}{\partial \mu^2} = \frac{\alpha_s(\mu^2)}{2\pi} \int_x^1 \frac{\mathrm{d}y}{y} P_{NS}(x/y) f_f^{\rm NS}(y,\mu^2) \tag{2.29}$$

where P_{NS} is the splitting kernel. Notably, the evolution takes into account different values of momentum fraction larger than the value at which the derivative is evaluated. It corresponds to a quark emitting gluons and thus slowing down.[26] The combinations of flavours to which the above equation applies are the non singlet (NS) combinations

$$f_f^{\rm NS} = f_f - f_{\overline{f}} \tag{2.30}$$

which evolve on their own. The reason is that quarks and antiquarks are simultaneously generated or annihilated. Due to the interactions between quarks and gluons, the evolution of the singlet (S) combination

$$f^{\rm S} = \sum_f f_f \tag{2.31}$$

is mixed with the evolution of gluons (G)[26]. The evolution equation then is [22]

$$\frac{\mu^2 \partial}{\partial \mu^2} \begin{pmatrix} f^{\rm S}(x,\mu^2) \\ G(x,\mu^2) \end{pmatrix} = \frac{\alpha_s(\mu^2)}{2\pi} \int_x^1 \frac{\mathrm{d}y}{y} \boldsymbol{P}\left(\frac{x}{y}\right) \begin{pmatrix} f^{\rm S}(y,\mu^2) \\ G(y,\mu^2) \end{pmatrix}, \tag{2.32}$$

where the kernel now is a matrix:

$$\boldsymbol{P} = \begin{pmatrix} P_{qq} & P_{qG} \\ P_{Gq} & P_{GG} \end{pmatrix}. \tag{2.33}$$

The form for the individual splitting function follows from computation of corrections to the leading diagram of the process. The result thus obtained order by order in loop calculations, can be expanded as

$$\boldsymbol{P} = \sum_n \left(\frac{\alpha_s}{4\pi}\right)^n P_{qq}^{(n)}, \tag{2.34}$$

where the leading terms are

$$P_{\mathrm{NS}}^{(0)} = P_{qq}^{(0)} = C_F \left[\frac{1+x^2}{(1-x)_+} + \frac{3}{2}\delta(1-x)\right], \tag{2.35}$$

$$P_{qG}^{(0)} = \frac{x^2+(1-x)^2}{2} n_f, \tag{2.36}$$

$$P_{Gq}^{(0)} = C_F \left[\frac{1+(1-x)^2}{x} + \frac{3}{2}\delta(1-x)\right], \tag{2.37}$$

$$P_{GG}^{(0)} = 2C_A \left[\frac{x}{(1-x)_+} + \frac{1-x}{x} + x(1-x)\right] + \frac{11C_A - 2n_f}{6}\delta(1-x). \tag{2.38}$$

Here n_f denotes the number of flavours, C_F/C_A are the Casimir constants of the colour group in fundamental/adjoint representation, and the "plus"-distribution appearing in eq. (2.35)-2.38 must be understood as

$$\int_0^1 \mathrm{d}u w(u) v(u)_+ = \int_0^1 \mathrm{d}u \left[w(u) - w(1)\right] v(u) \tag{2.39}$$

for two sufficiently well behaving functions w, v with v divergent for $u \to 1$.

The equations eq. (2.29)-2.32 are known as DGLAP evolution equations, after the work of Dokshitzer, Gribov, Lipatov [27, 28] and Altarelli and Parisi [26]. They are crucial for determinations of PDFs. The leading order equation for the singlet eq. (2.29) can be solved analytically [29]. Other flavour combinations and higher order terms taken into account lead to evolution equations which cannot be solved analytically. Thus the evolution from a starting scale up to any other scale is implemented in programs numerically [30, 31].

A.c Determination of PDFs

In this segment we briefly summarize the discussed parts from previous sections and present the differential cross section in DIS without derivation. We discuss how the PDFs can be determined from data. Their knowledge is crucial for the extractions of TMDPDFs we present later in this work.

The terminology of factorising a cross section labels the decomposition of the differential cross section in a form that is similar to eq. (2.18). Thus one expects to write the differential cross section in a form where one identifies the hard cross section, being the process related to the a hard scale, and another part representing the hadronic origin of the parton. The latter then matches the constructed assumption in eq. (2.18), eq. (2.19) of describing the DIS process via parton distributions of $f^f(x)$. Indeed, for DIS at the leading order

$$\frac{\mathrm{d}^2\sigma^{\mathrm{DIS}}}{\mathrm{d}x\mathrm{d}Q^2} = \frac{2\pi\alpha^2}{Q^4} \sum_f f_f(x) Q_f^2 \left(1+(1-y)^2\right), \tag{2.40}$$

with all kinematic variables defined at the beginning of section A. As in figure 2, the hadronic state is source and sink of the fermion field. The scale dependence of f_f is omitted here.

To determine the form of the PDF, a functional form can be assumed. With the assumed PDF form, the (theoretical) cross section can be computed. Then via various methods, the parameters or parametrization are varied until the values are found with which the function describes the data the best. Due to the evolution equations eq. (2.29)-2.32, the parametrizations only need to be determined for some chosen initial scale. The accepted nomenclature is "the PDF is extracted at a certain scale".

In order to find the best parameters to a given PDF form, a χ^2-test function is introduced[32] with

$$\chi^2 = \sum_{i,j=1}^{N_{\mathrm{data}}} (m_i - t_i)\, V_{ij}^{-1}\, (m_j - t_j)\,. \tag{2.41}$$

The labels m_i and t_i represent the value of the experimental datapoint and its theory prediction, respectively. The matrix V_{ij} is the correlation matrix. It depends on the errors in the experimental measurement. If all datapoints are independent (uncorrelated), the matrix only contains diagonal entries.

The test function describes the quality of the fit. Typically the expression is given relative to the number of included datapoints N_{pt} in the fit. With this measure, works that do not use identical data can be compared. A good fit is one that yields a $\chi^2/N_{\mathrm{pt}} \sim 1.0$[4]. Even though this is a good measure to describe the quality of the fit, it is often not enough to look at this number only. If the data is very precise, small deviations from the measured value can produce a large contribution to the χ^2 function. Therefore it is sensible use e.g. graphical comparison between theory prediction and the data in addition to the test function to judge the quality of the fit.

The smaller the test function, the less the difference between theory and experiment. It is thus the goal of phenomenology to find the best parametrization for the non perturbatively calculable PDFs, to obtain good agreement with the experimental data.

In order to minimize the χ^2 function several methods can be applied. A commonly used technique is *gradient descent*. A review of this method and adaptions thereof can be found in [33]. The default algorithm consists of two steps:

1. Compute the gradient of the function in parameter space at the given position numerically.

2. Change the parameters towards the direction of the negative Gradient.

As long as the gradient is non zero, these two steps are be repeated, and the parameters should converge to a minimum of the χ^2 test function. In the review [33] ideas are listed how the algorithm can be improved to find a minimum in the parameter space with a reduced amount of computation steps. Given that in this task we work in a highly multidimensional space, this can be immensely helpful for functions which are very *expensive*[5] to compute.

There have been extractions made without the formulation of a functional form for collinear PDFs, instead with the usage of neural networks [32]. First of all, this demonstrates that the methodology exists to extract PDFs using a neural network. A second important aspect of this technique: not to constrain the non-perturbative function to a specific shape allows the function to stay more flexible, thus more adaptive to what the data drives it to. This reduces the bias that comes along with restricting to a model put by hand. On the downside, the use of neural networks brings a bias related to its architecture. We do not go into further details about the technical advantages of various methods, but we state that this purely technical aspect itself is an active field of research. Its efficiency is one of the aspects that limits what can practically be done in modern phenomenology.

In the following we give an overview of the PDF obtained in the fit `MSHT20PDF` [14] at NNLO accuracy. We choose to present this result because it is used in the TMDPDF extraction discussed later. The extraction for the

[4]More precisely, it is normalization not to the number of datapoints but to the "degrees of freedom", which is number of data minus the number of parameters. In the situations we discuss the number of datapoints is several hundreds and the number of parameters is ~ 10 order of magnitude, and can be neglected for this reason.

[5]Expensive in the sense that it requires a large amount of computing time.

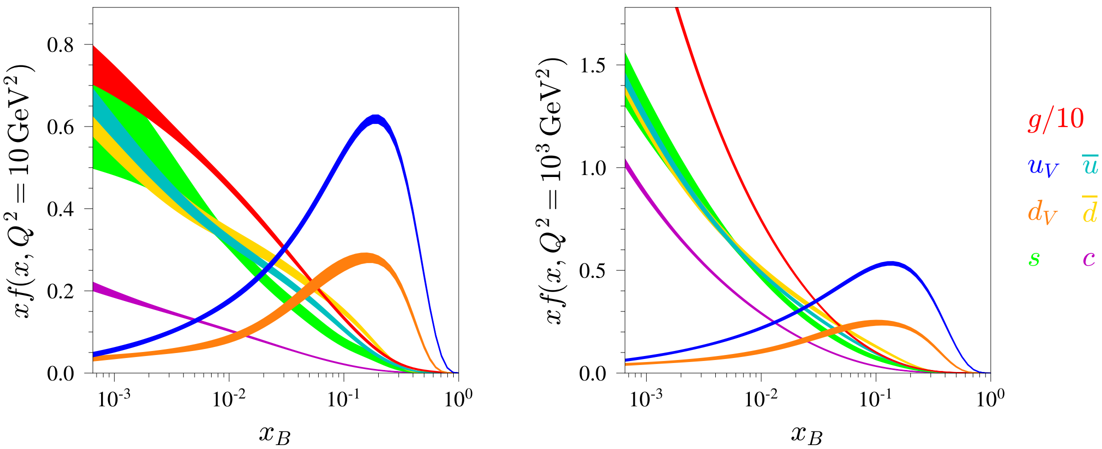


Figure 3: MSHTPDF20 with uncertainties for different flavour combinations at two points of Q^2. The valence flavour is defined as $q_V = q - \overline{q}$.

PDF in that work is obtained using the ansatz

$$xf(x, Q_0^2) = A(1-x)^{\eta} x^{\delta} \left[1 + \sum_{i=1}^{n} a_i T_i^{\mathrm{CH}}\,(y(x)) \right], \tag{2.42}$$

in which $T_i^{\mathrm{Ch}}(y)$ are Chebsyshev polynomials in $y = 1 - 2x^{0.5}$. Q_0^2 is here chosen to be $1\,\mathrm{GeV}$. The parameters A, η, δ may depend on the flavour, and from these flavour specifications the fit then has to determine a total of 52 parton parameters [14]. The fit presented there is obtained from data from DIS at HERA [14] and other experiments (e.g. SLAC), and also from hadron hadron collisions at the Tevatron, the LHC and other experiments [14]. The resulting parton distributions are shown for $Q^2 = 10\,\mathrm{GeV}^2$ and $Q^2 = 10^3\,\mathrm{GeV}^2$ in figure 3.

B Semi-inclusive DIS

In terms of the colliding particles and underlying hard reaction, SIDIS is the same process as DIS. In DIS, the hadronic final state is completely inclusive, and its substructure unobserved. By conservation of energy and momentum it is clear what the hadronic state as a whole is in the kinematic phase space after measuring the outgoing electron. In SIDIS the hadronic final state is considered and in addition to the scattered electron the final state of the struck quark is part of the specified final state of the reaction. The total hadronic remnant state consists of a cascade of particles which form from the initial protons remnants after the hard collision. The relevant final state in SIDIS is the hadron which forms from the quark in the hard reaction. To this end, the final state hadron must be detected. In experiment, this state is selected by fulfilling kinematic requirements [20].

The process is given by the reaction

$$l(k) + N(P) \to l(k') + h(p_h) + X(P'), \tag{2.43}$$

where h is the state associated to the final state of the quark of the hard (partonic) scattering process. The relevant invariant kinematics used to describe the scattering are those of the underlying DIS process eq. (2.2)-eq. (2.7) and also

$$z = \frac{P \cdot p_h}{P \cdot q}, \tag{2.44}$$

which, in the rest frame of the target N, is the fraction of the energy of the relevant final state hadron relative to the energy carried by the photon that strikes the quark and induces the reaction.

B.a Collinear factorisation of SIDIS

The knowledge of collinear fragmentation functions (FFs) is necessary to describe and extract transverse momentum dependent FFs (TMDFFs) in the currently used approach [12, 8], in the same way as collinear PDFs are necessary to extract TMDPDFs.

Thus in the following we give a short overview of the collinear factorisation of SIDIS, including the definition and a few aspects of collinear FFs, taken from [34]. The collinear FFs are typically obtained from single-inclusive electron-positron annihilation (SIA), SIDIS and single-inclusive hadron production (SIH) [34, 35]. The processes SIA and SIH are not directly important to this work, thus are not further discussed and references on experimental results for these processes are given in[36, 37]. The differential cross section for single hadron production in SIDIS is given by

$$\frac{\mathrm{d}\sigma^h}{\mathrm{d}x\mathrm{d}y\mathrm{d}z_H} = \frac{2\pi\alpha^2}{Q^2}\left[\frac{1+(1-y)^2}{y}2F_1^{N/h}(x,z,Q^2) + \frac{2(1-y)}{y}F_L^{N/h}(x,z,Q^2)\right], \tag{2.45}$$

The appearing structure functions at NLO are

$$\begin{aligned} 2F_1^{N/h}(x,z,Q^2) &= \sum_q Q_q^2\Big\{f^q(x,Q^2)d_q^h(z,Q^2) + \frac{\alpha_s(Q^2)}{2\pi}\Big[f_q \otimes C_{qq}^1 \otimes d_{q\to h} \qquad (2.46)\\ &\quad + f_q \otimes C_{gq}^1 \otimes d_{g\to h} + G \otimes C_{qg}^1 \otimes d_{q\to h}\Big]\left(x,z,Q^2\right)\Big\} \\ F_L^{N/h}(x,z,Q^2) &= \frac{\alpha_s(Q^2)}{2\pi}\sum_q Q_q^2\Big[f_q \otimes C_{qq}^L \otimes d_{q\to h} \qquad (2.47)\\ &\quad + f_q \otimes C_{gq}^L \otimes d_{g\to h} + G \otimes C_{qg}^L \otimes dq \to h\Big](x,z,Q^2). \end{aligned}$$

The function $d_{q\to h}$ is the (collinear) fragmentation function of a quark with flavour q into a hadron h, and the same notation is used for the gluon FF. The sums in the above equations include antiquarks. The representation of the functions $C_{ij}^{1/L}$ may be found in [34]. As it can be seen, the QCD corrections allow that gluon and quark densities mix. A gluon in the initial and final state of the hard reaction appears only at NNLO (and higher orders) accuracy.

The FFs are formally defined as the product of matrix element [38, 12]

$$d_{i\to h}(z) = \frac{p^+}{2\pi z}\int \mathrm{d}w e^{\mathrm{i}wpn/z}\sum_X \langle 0|\,\psi^i(wn)\,|h,X\rangle\,\langle h,X|\,\overline{\psi}^i(0)\,|0\rangle\,, \tag{2.48}$$

where the sum is over the unspecified remnant state of the proton, i.e. any hadronic state, and ψ^i is a quark field operator with flavour i. For a gluon it would be replaced accordingly, which also applies to the PDF.

In section A.a we discussed the interpretation of the PDF as the probability to find a parton of type i inside the initial state hadron N. It depends on the hard scale Q, thus this probability also takes into account the evolution from the hadronic scale Λ_{QCD} to the scale of the hard process. The fragmentation function encodes the opposite

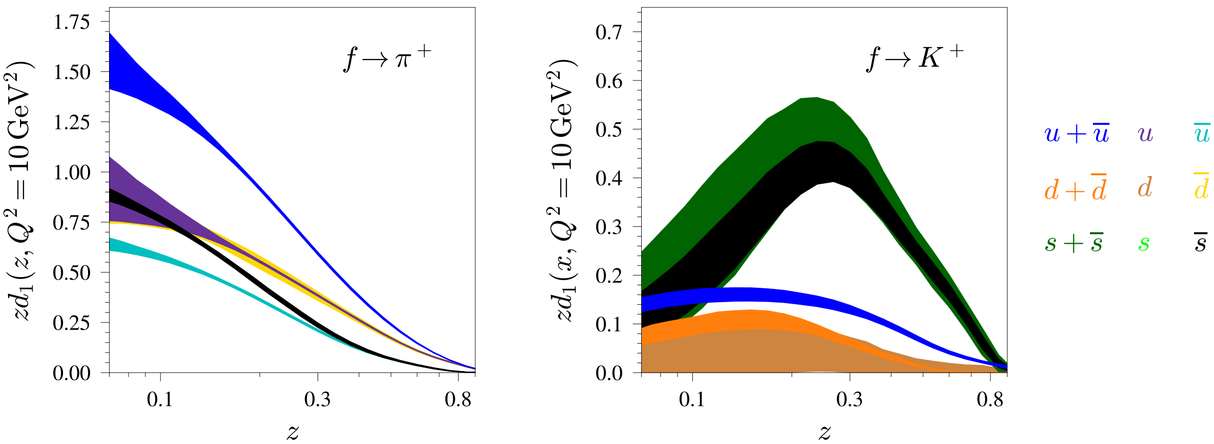


Figure 4: Fragmentation functions from DSS collaboration for pions and kaons at $Q^2 = 10\,\mathrm{GeV}^2$ for a selection of flavour combinations as indicated in the plots. The same FFs are used in the extraction of TMDFFs later in this work.

scale evolution. It describes the probability of a parton i at the hard scale Q to hadronise into a hadron h at the scale Λ_{QCD}.

In the later we require the knowledge of collinear FFs for pions and kaons. The FFs we use are extracted at approximated NNLO and NLO, respectively for pions [15] and kaons [16]. These phenomenological works use ansatzes similar to the in [39] presented functional form

$$d_{i\to\pi^+}(z, Q_0) = N_i z^{\alpha_i} \sum_{j=1}^{3} \gamma_{ij}(1-z)^{\beta_{ij}}. \tag{2.49}$$

The index i specifies the flavour of the parton to fragment into the hadron. N_i a normalization factor and α_i, β_{ij} further parameters. In the FFs we use, the ansatz is modified by a more involved normalization factor [15, 16]. The resulting FFs for the pion and the kaon are shown in figure 4. The virtuality scale of $Q^2 = 10\,\mathrm{GeV}^2$ depicted there is relevant for the TMD extractions at the range of $2\,\mathrm{GeV}^2 < Q^2 < 81\,\mathrm{GeV}^2$ in the discussed in section C.

B.b TMD factorisation of SIDIS

In this section we present the SIDIS cross section for the TMD regime reviewed already in [12]. For a more involved discussion we also refer to [40, 41].

It is useful for notation and for the factorisation to use light-cone coordinates, and thus define two light-like vectors n, $\overline{n}$. In SIDIS, they are associated with the two relevant hadrons, the initial state hadron and the specified final state hadron.[6] Their momenta can then be decomposed in the plane of the two light-like vectors,

$$P^\mu = P^+\overline{n}^\mu + \frac{M^2}{2P^+}n^\mu, \qquad p_h^\mu = p_h^- n^\mu + \frac{m^2}{2p_h^-}\overline{n}^\mu. \tag{2.50}$$

The coefficients are clear once we specify the normalization of the light-cone vectors with $n \cdot \overline{n} = 1$ and the masses of the hadrons $P^2 = M^2$ and $p_h^2 = m^2$.

[6] In DY we associate one light vector to each of the two colliding hadrons, both initial states. In section C.a the concept of light-cone decomposition is briefly discussed as a more abstract scheme.

It is further useful to define the combinations

$$\gamma = \frac{2Mx_B}{Q}, \tag{2.51}$$

$$\varsigma = \gamma\frac{m}{zQ}, \tag{2.52}$$

$$\varsigma_\perp = \zeta\frac{m_\perp}{zQ} \tag{2.53}$$

with all variables as in previous sections.The transverse direction is defined as the direction transverse to both the incoming proton and the virtual photon. The transverse components are then implicitly defined by the transverse metric tensor

$$g_\perp^{\mu\nu} = g^{\mu\nu} - \frac{1}{M^2Q^2 + (P\cdot q)^2}\left[Q^2P^\mu P^\nu + P\cdot q\left(P^\mu q^\nu + q^\mu P^\nu\right) - M^2 q^\mu q^\nu\right]. \tag{2.54}$$

Thus the transverse momentum and transverse mass are given by

$$p_T^2 = p_h^\mu p_h^\nu g_{\perp,\mu\nu}, \tag{2.55}$$

$$m_\perp^2 = m^2 + \boldsymbol{p_T}^2, \tag{2.56}$$

and we use boldface to specify the metric used for the transverse momenta

$$\boldsymbol{p_T}^2 = p_\mu p_\nu g_\perp^{\mu\nu}. \tag{2.57}$$

It must be noted, that this definition of the transverse direction is different from the one used deriving the TMD factorisation. The TMD factorisation is valid from small momenta q_T which are defined in the plane of both hadrons. The relation between q_T and the measured transverse momentum $\boldsymbol{p_T}$ defined via eq. (2.54) is

$$q_T^2 = \frac{\boldsymbol{p_T}^2}{z^2}\frac{1+\gamma^2}{1-\varsigma^2}. \tag{2.58}$$

It must be noted that q_T is **not** defined via the eq. (2.54) but with the components transverse to the hadron-hadron plane used in TMD factroization framework [12].

In [40] the full differential TMD SIDIS cross section

$$\frac{\mathrm{d}\sigma}{\mathrm{d}x\mathrm{d}y\mathrm{d}\psi\mathrm{d}z\mathrm{d}\phi\mathrm{d}p_T^2} = \frac{\alpha^2 y}{8zQ^4}2MW^{\mu\nu}L_{\mu\nu}, \tag{2.59}$$

is discussed with the dependence on ψ (the azimuthal angle of the lepton) and ϕ (the azimuthal angle of the final state hadron). The leptonic tensor is the same as for DIS eq. (2.11). The hadronic tensor a priori is

$$W^{\mu\nu} = \int\frac{\mathrm{d}^4x}{(2\pi)^4}e^{-\mathrm{i}x\cdot q}\sum_X \langle P|\, J^{\mu\dagger}(x)\, |h(p_h), X\rangle \,\langle h(p_h), X|\, J^\nu(0)\, |P\rangle\,. \tag{2.60}$$

The hadronic tensor can be factorised and expressed via distributions, matrix elements of the implicitly defined TMD operators. The factorised expression up to power corrections of $1/Q^2$ is given by [12]

$$
\begin{aligned}
W^{\mu\nu} = -2z_s \sum_f Q_f \left|C_V(Q^2,\mu_F)\right|^2 \int \mathrm{d}^2 b e^{-\mathrm{i}b\cdot q} \\
\times \Big[g_\perp^{\mu\nu} F_{1,f\leftarrow H}(x_s,b;\mu_F,\zeta_1) D_{1,f\to h}(z_s,b;\mu_F,\zeta_2) \\
+ \left(g_\perp^{\mu\nu} b^2 - 2b^\mu b^\nu\right) \frac{mM}{4} h^\perp_{1,f\leftarrow N}(x_s,b;\mu_F,\zeta_1) H^\perp_{1,f\to h}(z_s,b;\mu_F,\zeta_2)\Big].
\end{aligned}
\tag{2.61}
$$

In this expression, the hard partonic cross section is factorised and contained in the hard coefficient C_V. It reflects the partonic quasi free process at the scale Q with the separation scale μ_F. It can be computed order by order in α_s in perturbation theory. In our work, we take the expression given in the literature [13]. The function $F_{1,f\leftarrow N}$ is the TMDPDF, where two spatially separate quark fields are contained inside the protons state N

$$
F_{1,f\leftarrow N}(x,\boldsymbol{b};\mu,\zeta) = \int \frac{\mathrm{d}w}{2\pi} e^{-\mathrm{i}wxp_+} \langle N|\, \overline{q}(wn+b)\, [wn+b, \infty n+b]\, \frac{\gamma^+}{2}\, [\infty n+, 0]\, q(0)\, |N\rangle\,, \tag{2.62}
$$

and the function $D_{1,f\to h}$ is the unpolarized TMDFF

$$
\begin{aligned}
D_{1,f\to h}(z,\boldsymbol{b};\mu,\zeta) = \frac{1}{2zN_C} \int \frac{\mathrm{d}w}{2\pi} e^{\mathrm{i}wp_+/z} \sum_X \langle 0|\, \frac{\gamma^+}{2}\, [-\infty n+b, wn+b]\, q(wn+b)\, |h(p_h),X\rangle \\
\times \langle h(p_h),X|\, \overline{q}(0)\, [0,-\infty n]\, |0\rangle\,.
\end{aligned}
\tag{2.63}
$$

Here in both terms we disregard the transverse segment of the Wilson line at infinity. The expressions for $h_1^\perp$ (Boer-Mulders function), $H_1^\perp$ (Collins function) can be found in the literature [12, 42]. In the literature [8, 12] the second term is omitted in the cross section for the extraction, i.e. the corresponding terms are neglected. Also in this work these terms are dropped and not taken into account for the (later) extraction.

The distributions in eq. (2.61) depend on the factorisation scale μ_F and additionally on the scale ζ. We discuss the scale evolution for TMDPDFs at a later point, but state that in order to obtain the physical point the scales must be chosen such that $\zeta_1\zeta_2 = q^4$. In SIDIS factorisation eq. (2.61) the distributions depend on the variables

$$
x_S = -x\frac{2}{\gamma^2}\left(1-\sqrt{1+\gamma^2\left(1+\frac{q_T^2}{Q^2}\right)}\right), \tag{2.64}
$$

$$
z_S = z\frac{x_S}{x}\frac{1+\sqrt{1-\varsigma^2}}{2\left(1+\frac{q_T^2}{Q^2}\right)}. \tag{2.65}
$$

The distributions in eq. (2.61) depend on the modulus $\boldsymbol{b}$ and not on the rotation of b. The boldface notation is to be understood as

$$
\boldsymbol{b}^2 = -b^2. \tag{2.66}
$$

Hence the angular integral can be taken and the hadronic tensor can be presented as

$$
W^{\mu\nu} \approx -\frac{z_s}{\pi}\sum_f g_\perp^{\mu\nu} \mathcal{W}^f_{f_1D_1}(Q,|q_T|,x_s,z_s), \tag{2.67}
$$

$$
\mathcal{W}^f_{F_1D_1}(Q,q_T,x_s,z_S) = \left|C_V(Q^2,\mu_F)\right|^2 \int_0^\infty \mathrm{d}b J_0(\boldsymbol{b}q_T) F_{1,f\leftarrow N}(x_S,b;\mu_F,\zeta_1) D_{1,f\to h}(z_S,b;\mu_F,\zeta_2). \tag{2.68}
$$

J_0 is the Bessel function of the first kind. The equal sign states that this neglects the contribution the second term in eq. (2.61) and is valid up to power corrections $\left(q_T^2/Q^2\right)$.

With all quantities defined and the TMDPDF and TMDFF hidden inside the structure eq. (2.68) then we give the formula for the unpolarized cross section in SIDIS as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}x\mathrm{d}z\mathrm{d}Q^2\mathrm{d}\boldsymbol{p_T}^2} = \frac{\pi}{\sqrt{1-\varsigma_\perp^2}}\frac{\alpha_{EM}^2}{Q^4}\frac{y^2}{1-\varepsilon}\frac{z_S}{z} \times \sum_f e_f^2\left(1-\frac{q_T^2}{Q^2}\frac{\varepsilon-\gamma^2/2}{1+\gamma^2}\right)\mathcal{W}^f_{F_1D_1}(Q,|q_T|,x_S,z_S). \tag{2.69}$$

C Drell Yan process

If we discuss the scattering of two hadrons, it is more difficult to define an observable that can be used to reconstruct the kinematics of a scattering of constituents of the respective hadrons. It was the idea of Drell and Yan [43] to propose the examination the production of a lepton pair by a hadron-hadron collision. The picture is that a quark of one hadron hits its corresponding antiquark in the other hadron. The two combine (annihilate) into a virtual photon[7]. This photon then materializes; it decays into anything, quarks or leptons.

The process we are interested in is the production of the virtual photon. The virtual photon itself is not observable. It must be reconstructed by its decay products. If the photon decays into quarks, they will hadronize and it is difficult to recreate the virtual boson, and its (four) momentum. This process is equally interesting, but the hadronic final state content is difficult to separate from the other remnants of the collision. These are hadrons as well and could even interact during their formation.

A lepton pair is a more clear signal. A measured lepton pair $l+\bar{l}$ of invariant mass that can be traced back to a common vertex means the lepton pair is the decay product of a photon of identical mass. The photon itself carries the momentum of the annihilated quark pair. This is how we can measure the sum of momenta of colliding quark pairs in hadron-hadron collisions.

The reaction to the Drell-Yan process (DY) is

$$h_1(p_1) + h_2(p_2) \to l(k_1) + \bar{l}(k_2) + X, \tag{2.70}$$

where we consider the production of a lepton pair $l+\bar{l}$ and an undetected arbitrary (hadronic) final state X. The final state momentum of X is not explicitly denoted and constrained by momentum conservation. DY is an important process for TMD physics, because it is the process where TMDPDFs can be determined the best. The momentum transverse to the scattering plane is related to the transverse momentum carried by the initial quarks[8]. Due to the definition of the process, there are two distributions (one for each hadron) describing the amplitude of a respective quark inducing the (hard) process. For identical hadrons we thus have only one amplitude which needs to be extracted from data - since it cannot be computed directly in perturbation theory. This is how DY is the cleanest process to determine TMDPDFs as it does not mix distributions.

The DIS process itself cannot provide information about the transverse momentum because the final hadronic state is not detected. SIDIS cures this problem by detecting the hadron the hard parton hadronises into. However, for the description of the TMD related process, both the TMDPDF and a TMDFF contribute. In that sense it is less clean, because it depends simultaneously on TMDPDFs and TMDFFs.

[7]In the discussed case of quark-antiquark collision they combine into a photon, or instead a Z-boson, if the energy is large enough. In general, also the production of a $W^{\pm}$-boson can be considered. In this section, we only treat the case of a virtual photon. The strategy for the factorisation in the general case is the same.

[8]This is a direct relation at leading order. Corrections of real emission of radiation in the hard process violates the one-to-one correspondance.

In the following we describe a factorisation theorem for the differential cross section for a lepton pair production in hadron-hadron collision in the TMD regime in most details.

C.a Light-cone coordinates

Let us consider the process of two hadrons moving at large momenta towards each other. The masses of the hadrons are negligible compared to the relativistic energy. Thus we consider them to be massless. With each of the two colliding hadrons we associate a light-like direction,

$$p_1^\mu \sim \overline{n}^\mu, \qquad p_2^\mu \sim n^\mu, \tag{2.71}$$

where n, $\overline{n}$ are light-like vectors, $n^2 = \overline{n}^2 = 0$, and they are normalized as a convention to $n \cdot \overline{n} = 1$. These two vectors define the scattering plane, and any four vector k^μ can be decomposed into its projection to that plane and the (remaining) transverse component:

$$k^\mu = k^+ \overline{n}^\mu + k^- n^\mu + k_T^\mu, \tag{2.72}$$

where the $+/-$ indices are shorthand notations for a product with $n/\overline{n}$. This notation is known as light-cone coordinates, such that we can write eq. (2.72) as a three component vector

$$k = \left(k^+, k^-, k_\perp\right). \tag{2.73}$$

C.b Power counting

In order to establish a factorisation theorem valid in the regime of the production of bosons with small transverse momenta, a consistent way of weighting fields according to their kinematic contribution is required. Such a formalism is developed in soft collinear effective field theory (SCET) [44]. It is an intuition-based approach which allows for a consequent treatment of contributions. The weighting formalism associates a power of a *small* scale parameter (λ) with the components of the fields, and also space. In this way, the factorisation has been established in the literature [44] and we sketch this approach in some detail in the following section. The result will be the proper TMD operator, relevant for the factorisation theorem at leading power.

To isolate the individual fields according to their contribution, and the kinematic scale we are interested in is

$$\lambda = \frac{q_T}{Q}, \tag{2.74}$$

where q_T is the transverse component of the momentum q associated with the virtual boson, and Q its (Minkowski) modulus.

We associate to fermion fields from different origins different counting to their momentum decomposition. Explicitly, this means we define *hard-collinear* (hc) fields, which scale as

$$p_{hc} \sim Q\left(1, \lambda^2, \lambda\right) \tag{2.75}$$

in their light-cone vector decomposition eq. (2.72). To help understanding this definition via scaling properties, we need to state that we consider the reaction in the centre-of-mass frame (CM).

By a *hc* field we think of a parton of the *hc* hadron moving fast in the $\overline{n}$ direction. As part of that hadron, this field is expected to have a momentum decomposition in light-coordinates with dominant "+", suppressed transverse and a even more suppressed "$-$" component. The counterpart to these *hc* fields are the anti-(hard-collinear) fields

$\overline{hc}$ from the hadron moving in opposite direction. Their momenta are associated with the n direction, and the fields have the scaling rule

$$p_{\overline{hc}} \sim Q\left(\lambda^2, 1, \lambda\right). \tag{2.76}$$

The precise relation of linear and quadratic suppression can be motivated by considering a Lorentz boost in the z direction, along which the hadrons are colliding. The power ordering of the fields in the CM is denoted in eq. (2.75)eq. (2.76). We now want to demonstrate that this definitions are consistent with a two fold suppression of the $-$ component in the hc field (and vice-versa in the $\overline{hc}$ field). We can consider a Lorentz boost from the centre-of-mass frame to the rest frame of the $\overline{hc}$ hadron. The components of a field are expected to transform with

$$B^{\mathrm{CM}\to\overline{hc}}:\ (k^-, k^+, k_\perp) \mapsto \left(\frac{k^-}{\lambda}, \lambda k^+, k_\perp\right). \tag{2.77}$$

The boost B enhances the "$-$" component by $1/\lambda$ while suppressing the "$+$" component is equally suppressed by λ. The transverse direction is by definition transverse to z and unaffected by the boost. We can look at the $\overline{hc}$ field in the above relation to motivate the behaviour of the boost. In the rest-frame of the incoming fields, all field components are equivalent in all directions ($\sim \lambda$). The behaviour can consitently be considered with a boost in the opposite direction. The inverse boost of eq. (2.77) transforms the fields from the CM frame to the hc frame (or from the $\overline{hc}$ frame to the CM frame). In this way we motivate the scaling rules defined in eq. (2.75) eq. (2.76). They are the point of departure for the kinematic power ordering procedure. This scheme relates the scaling behaviour from the hc fields to $\overline{hc}$, and shows that the scaling rule with leading and quadratically suppressed terms is consistent.

The quark fields can be separated into "good" and "bad" components by two projectors

$$\mathcal{P}_1 = \frac{\gamma^+\gamma^-}{2}, \qquad \mathcal{P}_2 = \frac{\gamma^-\gamma^+}{2}, \qquad \mathcal{P}_1 + \mathcal{P}_2 = 1, \tag{2.78}$$

which together cover the whole Dirac space. The good (ξ) components are those which are dominant, the bad (η) fields are power suppressed, as shown in the following. The quark field components are defined as

$$\begin{aligned} \xi_n &= P_1 q_n, \qquad & \xi_{\overline{n}} &= P_2 q_{\overline{n}}, \\ \eta_n &= P_2 q_n, \qquad & \eta_{\overline{n}} &= P_1 q_{\overline{n}}. \end{aligned} \tag{2.79}$$

The quark fields from the (anti) collinear hadron obey the QCD equations of motion (EoM) for the fields associated to the (anti) collinear region. Because the quarks are taken as massless, the EoMs are

$$\not{D}\left[A_n\right] q_n = 0, \qquad \not{D}\left[A_{\overline{n}}\right] q_{\overline{n}} = 0. \tag{2.80}$$

The EoMs for the individual components mix the good and bad components of the background quark fields,

$$\begin{aligned} \gamma^-\mathcal{D}_+\left[A_n\right]\xi_n &= -\not{D}_\perp[A_n]\eta_n, \qquad & \gamma^+\mathcal{D}_-\left[A_n\right]\eta_n &= -\not{D}_\perp[A_n]\xi_n, \\ \gamma^+\mathcal{D}_-\left[A_n\right]\xi_{\overline{n}} &= -\not{D}_\perp[A_{\overline{n}}]\eta_{\overline{n}}, \qquad & \gamma^-\mathcal{D}_+\left[A_{\overline{n}}\right]\eta_{\overline{n}} &= -\not{D}_\perp[A_{\overline{n}}]\xi_{\overline{n}}. \end{aligned} \tag{2.81}$$

The components of the derivatives are related to momentum components, which have different power counting themselves. The above relation follows from the EoM for the given field and they are exact. The power counting for fields related above thus has to match. This leads to the observation that good and bad components have different power counting, namely the bad components are power suppressed by one power relative to the good components:

$$\eta \sim \lambda\xi. \tag{2.82}$$

In order to obtain factorisation at the leading power, it is justified to solely take into account the good components of the background fields.

The counting rules for the gluon fields also directly follow from the EoMs, as the gluon field has to be of equal power as the derivative, and thus the components have powers as

$$\begin{aligned} A_n^+ &\sim \lambda^2, & A_n^- &\sim 1, & A_n^\perp &\sim \lambda, \\ A_{\overline{n}}^+ &\sim 1, & A_{\overline{n}}^- &\sim \lambda^2, & A_{\overline{n}}^\perp &\sim \lambda. \end{aligned} \tag{2.83}$$

These relations are not required for the factorisation at the leading power, but at first corrections to the (λ) power expansion [45].

The quark fields can be decomposed into (anti-)collinear and background fields (ψ)

$$q = q_n + q_{\overline{n}} + \psi, \tag{2.84}$$

and with this the current operator splits as

$$\begin{aligned} J^\mu = \overline{q}\gamma^\mu q \mapsto\ & \overline{q}_{\overline{n}}\gamma^\mu q_n + \overline{q}_n\gamma^\mu q_{\overline{n}} \\ & + \overline{\psi}\gamma^\mu q_n + \overline{\psi}\gamma^\mu q_{\overline{n}} + \overline{q}_n\gamma^\mu\psi + \overline{q}_{\overline{n}}\gamma^\mu\psi + \overline{\psi}\gamma^\mu\psi \\ & + \overline{q}_n\gamma^\mu q_n + \overline{q}_{\overline{n}}\gamma^\mu q_{\overline{n}}. \end{aligned} \tag{2.85}$$

The leading contribution to the cross section consists of the first line, in which we have operators in which the (anti-)collinear fields are directly connected. We keep the leading expression only, and expand the field by its leading two contributions in λ. With the coordinate expansion around the dominant field contribution, the collinear field at position y is

$$\begin{aligned} q_n(y) &= q_n(y_-) + y_\perp^\alpha \partial^\alpha q_n(y_-) + y_+\overline{\partial}^- q_n(y_-) + \dots \\ &= q_n(y_- + y_\perp). \end{aligned} \tag{2.86}$$

where higher derivatives are omitted. The transverse derivative adds one power in λ, which is compensated by an additional $y_\perp$ scaling with λ^{-1}. This scaling is given by the invariance of $q \cdot y \sim 1$, and the scaling rule for y must be [44, 45].

$$\left(y^-, y^+, y_\perp\right) \sim Q^{-1}\left(1, 1, \lambda^{-1}\right). \tag{2.87}$$

Hence full transverse expansion contributes at the same power in λ. The derivative in "$-$" direction adds two powers and thus is suppressed. The transverse direction contributes in full expansion without changing the contributing power. The effective current to leading powers then is

$$\overline{q}_{\overline{n}}(y_+\overline{n} + y_\perp)\gamma^\mu q_n(y_- n + y_\perp) + \overline{q}_n q_n(y_- n + y_\perp)\gamma^\mu q_{\overline{n}}(y_+\overline{n} + y_\perp) + \mathcal{O}(\lambda^4), \tag{2.88}$$

because the leading field combinations already have power counting λ^2, and the transverse components of one field have a minimal contribution of λ if considered at all. The fields then can be expanded into good and bad components and the effective current operators then are

$$\overline{q}_{\overline{n}}\gamma^\mu q_n + \overline{q}_n\gamma^\mu q_{\overline{n}} = \overline{\xi}_n\gamma_\perp^\mu\xi_{\overline{n}} + \overline{\xi}_{\overline{n}}\gamma_\perp^\mu\xi_n \\ + n^\mu\overline{\xi}_n\gamma^-\eta_{\overline{n}} + \overline{n}^\mu\overline{\xi}_{\overline{n}}\gamma_+\eta_n + n^\mu\overline{\eta}_{\overline{n}}\gamma^-\xi_n + \overline{n}^\mu\overline{\eta}_n\gamma_+\xi_{\overline{n}} \\ + \overline{\eta}_n\gamma_\perp^\mu\eta_{\overline{n}} + \overline{\eta}_{\overline{n}}\gamma_\perp^\mu\eta_n. \tag{2.89}$$

The first line of the right-hand-side (RHS) is of order λ^2 and involves good components only. The additional contributions contain mixed components ($\sim \lambda^3$) or are constituted solely of bad components ($\sim \lambda^4$).

The operator inside the hadronic matrix element consists of two currents at different position. Then the full contributing operator to leading power is

$$J^\mu(y)J^\nu(0) = \overline{\xi}_n(y_-n + y_\perp)\gamma_\perp^\mu\xi_{\overline{n}}(y_+\overline{n} + y_\perp)\overline{\xi}_{\overline{n}}(0)\gamma_\perp^\nu\xi_n(0) \\ + \overline{\xi}_{\overline{n}}(y_+\overline{n} + y_\perp)\gamma_\perp^\mu\xi_n(y_-n + y_\perp)\overline{\xi}_n(0)\gamma_\perp^\nu\xi_{\overline{n}}(0). \tag{2.90}$$

The other two terms do not contribute, as they involve combinations that cannot contribute to a forward scattering. The terms are required to have one (anti)fermion field from each hadron. This expression is useful in the following, as it points out the leading contributing field combinations in the process.

C.c Factorisation of the cross section

The differential cross section for a lepton pair production is a priori

$$\mathrm{d}\sigma^{\mathrm{DY}} = \frac{1}{2s}\frac{\mathrm{d}^3k\mathrm{d}^3k'}{2\,|k|\,(2\pi)^3 2\,|k'|\,(2\pi)^3}\int \mathrm{d}\Pi_X \left|\mathrm{i}\mathcal{M}\left(N_1 + N_2 \to l(k) + \bar{l}(k') + X(p_X)\right)\right|^2 \delta^{(4)}\left(P_1 + P_2 - q - p_X\right), \tag{2.91}$$

with the momentum of the lepton pair

$$q = k + k'. \tag{2.92}$$

Further, we do not want to resolve the momentum distribution of the lepton pair. What matters in this analysis is the combined momentum q and, for our particular purpose, its transverse component q_T. We can integrate out the angular dependence of one lepton from eq. (2.91) and keep the modulus and direction of the combined lepton pair undetermined. We obtain the cross section, differential in terms of the momentum components of the virtual photon as [44]

$$\mathrm{d}\sigma^{\mathrm{DY}} = \frac{4\pi\alpha^2}{3q^2 s}\frac{\mathrm{d}^4q}{(2\pi)^4}\int \mathrm{d}^4y^4 e^{-\mathrm{i}y\cdot q}\left(-g_{\mu\nu}\right)\left\langle N_1(P_1)N_2(P_2)\right| J^{\mu\dagger}(x)J^\nu(0)\left|N_1(P_1)N_2(P_2)\right\rangle. \tag{2.93}$$

In this computation, the principal leptonic structure is identical to that that from DIS eq. (2.11), consisting of a lepton pair in initial and final states and a two fermion lines. Thus, in this formula the metric tensor is the result of the leptonic tensor and the integral over the momentum decomposition inside the lepton pair. The hadronic structure as a forward matrix element is obtained via the optical theorem. It can be disentangled from the cross section as the hadronic tensor

$$W^{\mu\nu} = \int \mathrm{d}^4y e^{-\mathrm{i}y\cdot q}\left\langle N_1(P_1)N_2(P_2)\right| J^{\mu\dagger}(x)J^\nu(0)\left|N_1(P_1)N_2(P_2)\right\rangle. \tag{2.94}$$

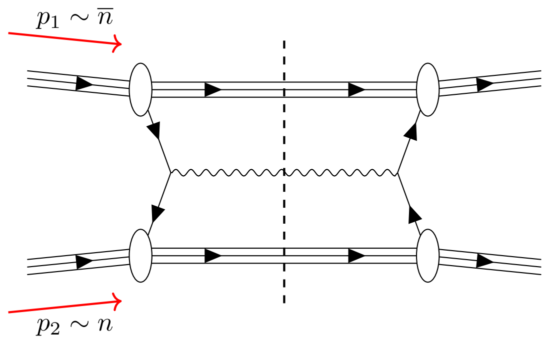


Figure 5: Schematic of optical theorem for Drell-Yan process in hadron-hadron scattering. The vertical line in the centre indicates the final state. The whole diagram then depicts the forward scattering with a virtual photon production. The arrows to the left indicate the (light-like) directions of the incoming hadrons.

After the leptonic and hadronic part have been separated, the factorisation of the process proceeds to this term into hadronic (non perturbative) and partonic (quasi free QCD) contributions. To proceed, we follow the work in [44] and, by that means, use the results from SCET. The factorisation is formally completed by replacing the current operator by

$$J^\mu = C_V(-q^2, \mu_F)\overline{\chi}_{\overline{hc}} S_{\overline{n}}^\dagger \gamma^\mu S_n \chi_{hc}. \tag{2.95}$$

The definition of the fields is

$$\chi_n = W_n^\dagger \xi_n, \qquad \chi_{\overline{n}} = W_{\overline{n}}^\dagger \xi_{\overline{n}}, \tag{2.96}$$

and they are related to the QCD fermion fields and gauge link connecting them to their source the initial(or final[9]) state hadron. The $S_{n/\overline{n}}^{(\dagger)}$ denote soft Wilson lines. Their contribution can be condensed to a vacuum matrix element which contributes a unity factor to the cross section [44].

The current in the hadronic tensor now is replaced by its leading power term eq. (2.90), and at the same the transformations of eq. (2.95) are applied. The Dirac structure can be Fierz decomposed [46] as

$$(\chi_{hc}\overline{\chi}_{hc})_{ij} = \overline{\chi}_{hc}\chi_{hc}\delta_{ij} + \overline{\chi}_{hc}\gamma_\mu\chi_{hc}\gamma^\mu_{ij} + \cdots, \tag{2.97}$$

where the full expansion in the complete basis of the Dirac space $\{\mathbb{1}, \gamma^\mu, \gamma_5, \gamma^\mu\gamma_5, \sigma_{\mu,\nu}\gamma_5\}$ is truncated. To the unpolarised cross section, only the terms $\sim \gamma^\mu$ contribute [44].

Inserting these contributions only, the hadronic tensor becomes

$$W^{\mu\nu} = \int \mathrm{d}y^4 e^{-\mathrm{i}y\cdot q} \langle N_1| \overline{\chi}_n(x)\gamma_\alpha\chi_n(0) |N_1\rangle \langle N_2| \overline{\chi}_n(0)\gamma_\beta\chi_n(x) |N_2\rangle \operatorname{Tr}\left[\gamma^\alpha\gamma_\perp^\mu\gamma^\beta\gamma_\perp^\nu\right] + (\overline{q} \leftrightarrow q), \tag{2.98}$$

with the trace carrying the spinor algebra. It needs to be normalized by the number of spin states and the colour states of the quarks N_c. The second term which is not explicitly denoted is the result of the second term in eq. (2.90). This additional term simply acknowledges that the (anti)quark can originate from the first or second hadron.

Now the matrix elements of the hadronic tensor are in a form of two separate matrix elements of the SCET fields χ. This rewriting is justified by considering the operator of collinear fields to be an operator fully absorbed

[9] After application of the optical theorem we have forward matrix elements, in which the same hadronic states appear on each side. Due to the orientation, in DY both lines point to the initial position of the hadron before the scattering at $-\infty\overline{n}$ for N_1(n for N_2)

by the initial and final collinear state. The behaviour for the anti-collinear fields is analogue. In other words, the operator combination is a diagonal block matrix in the representation of the states $\{N_1(P_1), N_2(P_2)\}$.

The trace in eq. (2.98) contracted with the remnant from the leptonic tensor $(-g^{\mu\nu})$ projects out all transverse components. The collinear tensor can be split into the terms $n^\alpha \overline{n}^\beta$ and $n^\beta \overline{n}^\alpha$. Because of the definitions of good components in eq. (2.79) the first of these contributes and the second term vanishes. The cross section then becomes

$$\begin{aligned} \mathrm{d}\sigma^{\mathrm{DY}} = &\frac{4\pi\alpha^2}{3N_c q^2 s}\frac{\mathrm{d}q^4}{(2\pi)^4}\int \mathrm{d}^4 y e^{-\mathrm{i}y\cdot q}\left|C_V(-q^2,\mu_F)\right|^2 \\ &\times \sum_f \langle N_1|\,\overline{\chi}_{\overline{n}}(y)\frac{\gamma^+}{2}\chi_{\overline{n}}(0)\,|N_1\rangle\,\langle N_2|\,\overline{\chi}_n(0)\frac{\gamma^-}{2}\chi_n(y)\,|N_2\rangle + (\overline{q}\leftrightarrow q). \end{aligned} \tag{2.99}$$

The space integral can be decomposed into (anti-)collinear and transverse directions. The first matrix element only depends on the transverse and the "−" component of the spatial separation, due to power counting eq. (2.86). The second matrix element, on the contrary, depends on the transverse and the "+" component. The integrals along the light directions then can be included in the definition of the TMDPDFs as in eq. (2.62).

The matrix elements themselves do not depend on the orientation of $y_\perp$. In the following, we replace the label $y_\perp$ by b which is typically used for the transverse distance the TMD notation. The final result for the cross section, which we give here, is the form differential in $Q^2 = q^2$, q_T^2 and y, with the latter being the rapidity

$$y = \ln\left(\frac{q_+}{q_-}\right), \tag{2.100}$$

associated with the direction of the photon along the light cone direction. For $y = 0$, the photon carries transverse momentum only. We note that the components q_0 and q_3 are equivalent to the components q_+ and q_-. If we are interested in the modulus but not the azimuthal direction of the virtual photon, we can integrate out the angle. We obtain [44]

$$\frac{\mathrm{d}^3\sigma^{\mathrm{DY}}}{\mathrm{d}Q^2\mathrm{d}y\mathrm{d}q_T^2} = \frac{4\pi\alpha^2}{3N_c Q^2 s}\frac{1}{(4\pi}\int \mathrm{d}^2 b e^{-\mathrm{i}b\cdot q_T}\left|C_V(-q^2,\mu_F)\right|^2 \tag{2.101}$$

$$\times \sum_f Q_f^2 F_{1,f\leftarrow N_1}(x_1,\boldsymbol{b},\mu_F)F_{1,f\leftarrow N_1}(x_2,\boldsymbol{b},\mu_F) + (\overline{q}\leftrightarrow q). \tag{2.102}$$

Leading order the symbols $x_{1,2}$ are interpreted as the momentum fractions of the colliding quarks,

$$x_1 = \sqrt{\frac{q_+q_-}{s}}e^y, \quad x_2 = \sqrt{\frac{q_+q_-}{s}}e^{-y}. \tag{2.103}$$

They appear in the implicit integral over light-cone components and are related to the rapidity, which can be seen by decomposition of the momentum of the virtual photon into light cone components

$$q = q_+\overline{n} + q_-n + q_T. \tag{2.104}$$

We approximate that the $+(-)$ component is given solely from a momentum fraction x of the (anti)collinear hadron,

$$q_+ = x_1 P_1^+, \qquad q_- = x_2 P_2^-, \tag{2.105}$$

as the other components contribute power suppressed only.

The result of this section is eq. (2.101). There are some details we omitted to discuss. Thus far, we did not discuss the matching coefficient C_V. It is (naively) clear that it includes the higher loop corrections to the partonic scattering. In our computation we take the result from the literature [13]. Notably, the form of the matching coefficient is identical for DY and SIDIS as the separation of the hard scattering from the quark current is identical [12]. This is clear since the partonic matrix element is a photon exchange between an electron-type fermion line and a quark-type fermion line. In SIDIS, $q^2 = \hat{t}$, in DY $q^2 = \hat{s}$ in mandelstam variables for the partonic process. The difference is that the sign of the hard scale that enters the matching coefficient in DY is opposite to the sign in SIDIS.

We do not explain in great detail how the Wilson line W and the soft terms (S) arise precisely. For these – we here consider them as technical, but important – details, we refer to [47].

For the sake of completeness we denote that in the same framework (SCET) also the collinear factorisation is worked out in the literature [47]; the cross section for the DY reaction of unobserved/integrated transverse momentum reads

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}Q^2\mathrm{d}y} = \frac{4\pi\alpha^2}{3N_cQ^2s}\sum_{ij}\int \mathrm{d}x_1\mathrm{d}x_2\tilde{C}_{ij}(x_1,x_2,s,Q,\mu_F)f_{i\leftarrow N_1}(x_1,\mu_F)f_{j\leftarrow N_2}(x_2,\mu_F). \tag{2.106}$$

The expression depends on the collinear PDFs f, and they and the hard scattering kernels $\tilde{C}_{ij}$ depend on the factorisation scale. Their expression at NLO can be found in the reference. The number of colours is denoted by N_c, and s denotes the mandelstam variable.

C.d Scale dependence of a TMD distribution

The TMDPDFs obey a factorisation scale evolution governed by two coupled differential equations. These are determined from eq. (2.101) in which the full expression must be scale independent. It follows that

$$\begin{aligned}\left\{\frac{\mathrm{d}}{\mathrm{d}\ln(\mu)}\left|C_V(-q^2,\mu^2)\right|^2\right\}\int \mathrm{d}^2be^{-\mathrm{i}qb}F_1(x_1,\boldsymbol{b},\mu)F_2(x_1,\boldsymbol{b},\mu)&\\ = -\left|C_V(-q^2,\mu^2)\right|^2\int \mathrm{d}^2be^{-\mathrm{i}qb}\left\{\frac{\mathrm{d}}{\mathrm{d}\ln(\mu)}F_1(x_1,\boldsymbol{b},\mu)F_2(x_1,\boldsymbol{b},\mu)\right\},&\end{aligned} \tag{2.107}$$

with $F_{1,2}$ a general TMD distribution. This way we obtain the factorisation scale evolution as a relation between the matrix elements and the hard coefficient.

The matching coefficient is a purely perturbative object. The scale dependence of the matching coefficient can be computed order by order and it is collected in two functions,

$$\frac{\mathrm{d}}{\mathrm{d}\ln(\mu)}C_V(-q^2,\mu) = \left[\Gamma^F_{\mathrm{cusp}}(\alpha_s)\ln\left(\frac{-q^2}{\mu^2}\right)+\gamma^V(\alpha_s)\right]C_V(-q^2,\mu). \tag{2.108}$$

The function $\gamma^V(\alpha_s)$ is the quark anomalous dimension[25]. We can simplify the above evolution for the TMDPDF by assuming the transformation rule is a local (in b space) property and remove the transverse integral. We then only look at the scale connection of the functions. Further, we assume the transformation rules for quark and antiquark to be identical, thus also for both TMD distributions involved. Then the evolution of a single TMDPDF with respect to the factorisation scale can be expressed with eq. (2.107) as

$$\frac{\mathrm{d}}{\mathrm{d}\ln(\mu)}F(x,\boldsymbol{b};\mu) = \left[\Gamma^F_{\mathrm{cusp}}(\alpha_s)\ln\left(\frac{\mu^2}{-q^2}\right)-\gamma^V(\alpha_s)\right]. \tag{2.109}$$

The product of both transverse position dependent matrix elements is independent of the auxiliary scales needed for the analytic regularisation of the matching kernels to collinear operators in the limit of small transverse distances. As explained in [44], for the TMD factorisation we have two scales: q^2 and $q_T^2 \sim \lambda^2$. There it is discussed that the product of the sum of hard collinear graphs with the sum of anti-hard collinear graphs, entering the cross section as a product, must be independent of the auxiliary scales. We refer to that text for a more details. A consequence of this is the relation

$$[F_1(x_1, \boldsymbol{b}, \mu) F_2(x_2, \boldsymbol{b}, \mu)]_{q^2} = F_1(x_1, \boldsymbol{b}, \mu) F_2(x_2, \boldsymbol{b}, \mu) \left(\frac{\boldsymbol{b}^2 q^2}{4e^{-2\gamma_E}} \right)^{-2\mathcal{D}(\boldsymbol{b},\mu)}, \tag{2.110}$$

which removes the implicit dependence on q^2 of the TMDPDFs and brings it into a prefactor. Its exponent $\mathcal{D}$ is independent of $x_{1,2}$. This function is identified as the rapidity anomalous dimension (RAD) or Collins-Soper (CS) kernel [48].

It might appear as arbitrary to complicate the problem of describing the matrix elements even further by allowing for the dependence on an additional scale. The advantage of this prescription is that the TMD distribution can be understood and examined independently from the physical cross section. It promotes the TMD distribution to an universal object, that does not depend on the exterior momentum scales any longer. This step completes the factorisation; we obtain a TMDPDF which depends on an artificial scale ζ and the factorisation scale μ, and it does not depend on the scales of the physical process. From the above relation we then deduce that for a single TMDPDF

$$F(x, \boldsymbol{b}; \mu, q^2) = F(x, \boldsymbol{b}; \mu, \zeta) \left(\frac{\boldsymbol{b}^2 \zeta}{4e^{-2\gamma_E}} \right)^{-\mathcal{D}(\boldsymbol{b},\mu)}, \tag{2.111}$$

The left-hand-side (LHS) does not depend on the artificial scale ζ which is evaluated at q^2. From this equation, we find the scale dependence on this second energy scale with

$$\frac{\mathrm{d}}{\mathrm{d}\ln\zeta} F(x, \boldsymbol{b}; \mu, \zeta) = -\mathcal{D}(\boldsymbol{b}, \mu) F(x, \boldsymbol{b}; \mu, \zeta). \tag{2.112}$$

For two TMDPDFs one can allow both of them to have different values in this scale[10], ζ_1 and ζ_2, but in order to come back to the physical point and enter the cross section, their product must restore $\zeta_1\zeta_2 = (q^2)^2$ to satisfy eq. (2.112). A typical choice for the physical[11] point is $\zeta_1 = \zeta_2 = q^2$.

The equations eq. (2.109) eq. (2.112) define the evolution in the two-scale plane for the TMD distributions. Both equations are joined result in the Collins-Soper equation

$$\frac{\mathrm{d}}{\mathrm{d}\ln\mu} \mathcal{D}(\boldsymbol{b}, \mu) = -\Gamma_{\text{cusp}}(\mu), \tag{2.113}$$

which constrains the scale evolution of the RAD to the perturbative Γ_{cusp}.

It is important to observe that in eq. (2.113) there is no dependence on $\boldsymbol{b}$ on the right hand side. If the relation holds exactly, the function $\mathcal{D}$ depends on $\boldsymbol{b}$ only via $L_\perp = \ln(\mu^2 \boldsymbol{b}^2)$. The function can be expanded as a perturbative series, in which the μ dependence enters via $L_\perp$ but also $\alpha_s(\mu)$. It has been worked out in the literature [49], and the perturbative truncation of the series works in the regime where both $\alpha_s(\mu)$ and $L_\perp$ are small quantities. For a scale choice such that $(\alpha_s(\mu) L_\perp(\mu)) \sim 1$, the expansion can be reorganised to a resummed expression [49].

[10] We stress that this mentioned freedom is valid for ζ only and not for the factorisation scale μ.

[11] The physical point is the scale value for ζ where the cross section is described. For a single TMD distribution, a physical point is ill defined as explained in the above discussion, and in more detail in the literature [44].

III TMD operators

This chapter deals with TMDs as mathematical objects via the definition of TMD operators. We define a general TMD distribution as the matrix element of the TMD operator, which can itself be analysed, expanded and related to other operators [50][42]. Properties of the operator are more universal than properties of the distributions, as the former apply disregarding the external states. We can study them disentangled from the cross section and distributions. In the context of this, we present a calculation of one loop corrections to the TMD operator.

A Definitions

The computation involves corrections to what we call twist*-1 and twist*-2 operators for both the case of a quark or a gluon field located on the end of the semi-infinite Wilson line. This will be clarified later, and we use the superscript * to distinguish between this and regular twist counting (dimension - spin). The TMD operator for gluons at twist-2 consists of two gluons separated by a light-like distance $z_1 n$ and a transverse distance $\boldsymbol{b}$. In between, a staple of straight Wilson lines connects them, resulting in a gauge invariant definition of the operator. For gluons it reads

$$\mathcal{U}(z_1, \boldsymbol{b}) = F_{\mu+}(z_1 n)\,[z_1 n, \pm\infty n]\,[\pm\infty n, \pm\infty n + \boldsymbol{b}]\,[\pm\infty n + \boldsymbol{b}, \boldsymbol{b}]\,F_{\nu+}(\boldsymbol{b}), \tag{3.1}$$

where n denotes a light-cone vector, $F_{\mu\nu}$ is the field-strength tensor for the gluon field

$$F^a_{\mu\nu}(x) = \partial_\mu A^a_\nu(x) - \partial_\nu A^a_\mu(x) + g f^{abc} A^b_\mu(x) A^c_\nu(x). \tag{3.2}$$

For quarks the TMD operator is the same structure with quark fields at the end points of the stapled gauge link. Denoted with open spinor indices, it is

$$\mathcal{U}^q_{ij}(z_1, \boldsymbol{b}) = \bar{q}_i(z_1 n)\,[z_1 n, \pm\infty n]\,[\pm\infty n, \pm\infty n + \boldsymbol{b}]\,[\pm\infty n + \boldsymbol{b}, \boldsymbol{b}]\,q_j(\boldsymbol{b}). \tag{3.3}$$

Contributing to the factorisation the leading important terms are those with the spinors contracted in a certain way. For the quark case, the terms we are interested in are contracted with the gamma matrix $\gamma^\mu n_\mu$.

For twist-3, an additional field strength tensor appears along the connecting Wilson lines, leaving two options for where to insert the term: before or after taking the Wilson line along the transverse distance at infinity. The third possibility to insert the field strength along the transverse gauge link is valid, but the physical field vanishes at infinity $F^{\mu\nu}(\infty) = 0$. We denote an appearance of the field before infinity as twist-(2+1), and after infinity as twist-(1+2). With this notation, by a twist-(2+1) gluon TMD operator we refer to

$$\mathcal{U}(\{z_1, z_2\}, \boldsymbol{b}) = F_{\mu+}(z_1 n)\,[z_1 n, z_2 n]\,F_{\nu+}(z_2 n)\,[z_2 n, \pm\infty n]\,[\pm\infty n, \pm\infty n + \boldsymbol{b}]\,[\pm\infty n + \boldsymbol{b}, \boldsymbol{b}]\,F_{\nu+}(\boldsymbol{b}), \tag{3.4}$$

where z_2 describes the position of the intermediate field, ordered as $z_1 \leqslant z_2 \leqslant \pm\infty$. The TMD operator for twist-2(3) can be written as the function of two (three) separate positions $z_1, z_2(, z_3)$. The matrix element of this operator is translation-invariant, thus we set the second (third) positional argument to zero without loss of generality.

A quark TMD operator then differs by replacing the gluon fields at the ends of the stapled Wilson lines by quark fields, and an insertion of a Dirac structure. In addition, the representation of the colour group differs between the gluon and the quark operator. For the quark case, the colour group is in the fundamental representation; for the gluon case, the group is in the adjoint representation. The different representation gives different results for the colour factor - labeling the computation of colour terms only in a given expression.

The colour structure is relevant for the computation of the one-loop corrections. Resolving the colour connection along the *path ordered* exponent in more detail, the TMD operator eq. (3.1) is the product of matrices

$$\begin{aligned}\mathcal{U}(z_1 n, \boldsymbol{b}) &= F^a_{\mu+}(z_1 n)\,[z_1 n, \pm\infty n]^{ab}\,[\pm\infty n, \pm\infty n + \boldsymbol{b}]^{bc}\,[\pm\infty n + \boldsymbol{b}, \boldsymbol{b}]^{cd}\,F^d_{\nu+}(\boldsymbol{b}) \\ &= F^a_{\mu+}(z_1 n)\,[\pm\infty n, \pm\infty n + \boldsymbol{b}]^{ad}\,F^d_{\nu+}(\boldsymbol{b}),\end{aligned} \tag{3.5}$$

where the second line is in light-cone gauge for the background field ($A_+ = 0$) turning the straight Wilson lines along the light cone into unit matrices.

We define the "half" of the TMD operator, i.e. the terms up to the first semi-infinite Wilson line as an individual object, that now has an open colour index:

$$\mathcal{X}^a_\mu\,(z_1) = F^e_{\mu+}(z_1 n)\,[z_1 n, \pm\infty n]^{ea}\,, \tag{3.6}$$

$$\mathcal{X}^a_{\mu,\nu}\,(z_1, z_2) = F^e_{\mu+}(z_1 n)\,[z_1 n, z_2 n]^{eb}\,F^d_{\nu+}(z_2 n) T^d_{bc}\,[z_2 n, \pm\infty n]^{ca}\,, \tag{3.7}$$

for a twist*-1 and a twist*-2 half operator. The colour matrix in eq. (3.7) at the second field or any field inserted along the Wilson line is in adjoint representation and hence can be rewritten

$$T^b_{ac} = \mathrm{i} f^{abc}, \tag{3.8}$$

with f^{abc} representing one of the structure constants of the group. For the quark case the colour matrices are in fundamental representation. The quark operators differ from gluon operators by the replacement

$$T^b_{ac} \mapsto \lambda^b_{ac}, \tag{3.9}$$

with λ^b a Gell-Mann matrix and a quark field at the finite end of the Wilson line instead of a field strength tensor.

B Background field method

The computation to follow is performed with the background field method (BFM) [51]. Effectively the BFM splits the quark and boson fields in two components: slow and fast fields, classical and quantum fields, background and dynamical fields; all synomyms for the decomposition of the fields. In the notation of [45], in which a part of this computation was already published, we separate

$$q(x) \to q(x) + \psi(x), \tag{3.10}$$

$$A(x) \to A(x) + B(x), \tag{3.11}$$

where q (ψ) then represents the background (dynamical) quark field, and A (B) denotes the background (dynamical) gluon field.

The Lagrangian in the BFM for the interaction between background and dynamical fields is given in [45]. It reads

$$\mathcal{L}_{\text{eff}} = g\left(\bar{q}\not{B}\psi + \bar{\psi}\not{A}\psi + \bar{\psi}\not{B}q\right) - g f^{abc} A^a_\mu \left[2\partial_\nu B^b_\mu B^c_\nu - \partial_\mu B^b_\nu B^c_\nu - \frac{1+\alpha}{\alpha}\partial_\nu B^b_\nu B^c_\mu\right] \tag{3.12}$$

$$- \frac{g^2}{2} f^{abr} f^{rcd}\left[A^a_\mu A^b_\nu B^c_\mu B^d_\nu + A^a_\mu B^b_\nu A^c_\mu B^d_\nu + A^a_\mu B^b_\nu B^c_\mu A^d_\nu + \frac{1}{\alpha} A^a_\mu B^b_\mu A^c_\nu B^d_\nu\right]. \tag{3.13}$$

This neglects the contribution of *ghost* fields, but includes everything needed for the calculation at one loop. The necessary ingredients are: antiquark-quark-gluon interaction (in three terms each containing one background and two dynamical field), as well as $(1+2)$- and $(2+2)$-gluon interaction. The parameter α is the free gauge parameter for the dynamical gluon fields.

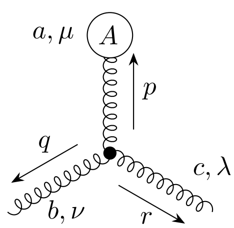
For the dynamical fields the gauge choice is $\alpha = 1$, in which the term involving three gluon fields can be rewritten to

$$\mathcal{L}_{(1+2)g} = -f^{abc} A_\iota^a \partial_\alpha B_\beta^b B_\gamma^c v_{\iota\alpha\beta\gamma}, \tag{3.14}$$

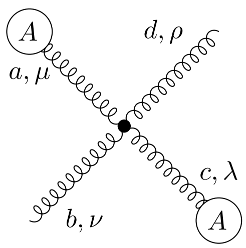
$$v_{\iota\alpha\beta\gamma} = \left(2g^{\iota\beta} g^{\alpha\gamma} - g^{\iota\alpha} g^{\beta\gamma} - 2g^{\iota\gamma} g^{\alpha\beta}\right). \tag{3.15}$$

This rewriting is trivial, but appears to be convenient for the further calculation. With the first part of this expression contractions and their colour factors can be directly computed. The result can be combined, respecting the colour factor, and the Lorentz structure given by $v_{\iota\alpha\beta\gamma}$ is consistently computed in *Mathematica* at a later point in the calculation. We think it is convenient to use the action given through eq. (3.13) directly, but it is also possible to derive the Feynman rules and work with them.

The Feynman rules for the $(1+2)$- and $(2+2)$-gluon vertex can be computed via the connected correlation functions,



$$g f^{abc} \left[2g_{\mu\xi} p_\nu + g_{\nu\xi} (r-q)_\mu - 2g_{\mu\nu} p_\xi\right] \tag{3.16}$$



$$\begin{aligned} -\mathrm{i}g^2 \Bigg[& f^{abx} f^{xcd} \left(g_{\mu\lambda} g_{\nu\rho} - g_{\mu\rho} g_{\lambda\nu} + \frac{1}{a} g_{\mu\nu} g_{\lambda\rho}\right) \\ & + f^{adx} f^{xbc} \left(g_{\mu\nu} g_{\lambda\rho} - g_{\mu\lambda} g_{\nu\rho} - \frac{1}{a} g_{\mu\rho} g_{\lambda\nu}\right) \\ & + f^{acx} f^{xbd} \left(g_{\mu\nu} g_{\lambda\rho} - g_{\mu\rho} g_{\lambda\nu}\right) \Bigg] \end{aligned} \tag{3.17}$$

The momenta in eq. (3.16) are all outgoing, and for the $2+2$ vertex the momenta are irrelevant. This is the reason why we omit them in the expression, and also why the expression is identical in position space. The expressions match those from [51] (with the gauge parameter $\alpha = 1$).

In order to compute the one-loop correction to the TMD operator, we use the corresponding terms of the effective Lagrangian eq. (3.13). The corrections are computed for quark- and gluon- TMD operators, and in each case, for what we denote as twist-(1+1) and twist-(2+1). It means we compute the corrections to the operators of twist*-1 and twist*-2.

The quark TMD operators depend on a Γ structure inserted between the quark fields. We compute the corrections to the quark TMD operators for the *good*[12] components. This means we restrict ourselves to the situation where

$$\gamma^+ \Gamma = \Gamma \gamma + = 0. \tag{3.18}$$

[12]For a motivation compare to the definition in chapter II.C.b.

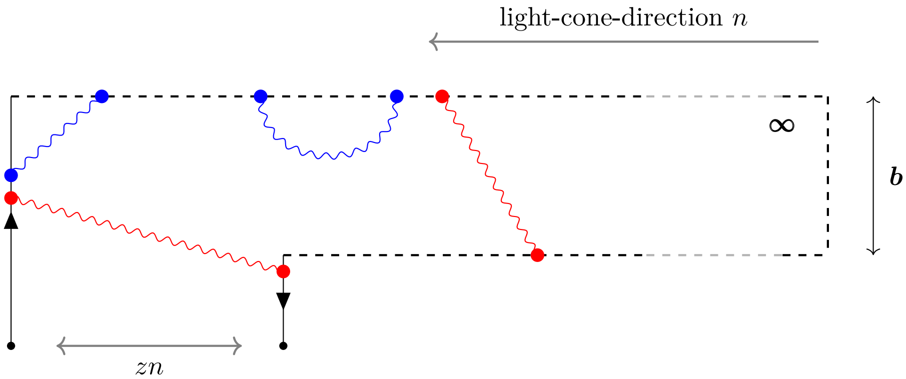


Figure 6: Schematic of the quark TMD operator with quantum corrections that do (blue) and do not (red) contribute to the UV divergent pole. The dashed lines indicate the gauge links. The quark fields at the end of the Wilson line are shown as fermion propagators to display corrections. The separation between the quark fields is indicated as zn, and the transverse distance by $\boldsymbol{b}$. Quantum corrections do not preserve the spatial light-cone separation of the external fields.

I.e. $\Gamma \propto \gamma^+$

to the left and to the right[13] . Since n is a light-like vector $(\gamma^+)^2 = n^2 = 0$, all contributions in which this factor appears to the left/right are eliminated.

For the gluon TMD operators there is no *Dirac* structure involved. An additional colour matrix between the fields inserted in a gauge invariant way is related to higher twist.

C One loop corrections

In this section we compute corrections to the TMD operator which correspond to the ultra-violet (UV) divergent part. We compute all the necessary diagrams that allow for interaction at only one of the two semi-infinite Wilson lines inside the TMD operator. Corrections which connect the two far end Wilson lines, or even the transverse Wilson line at infinity, naturally include propagation over the transverse distance $\boldsymbol{b}$, self-regulating the UV divergence. This is illustrated in figure 6.

In the following we present some exemplary computations of diagrams, along which we develop the methods to compute them. A full list of all expressions corresponding to all contractions for all diagrams, including their colour structure, can be found in appendix A. The technical computation is done with the program *Mathematica* and a self-developed approach, parts of which are strongly influenced by ideas of A. Vladimirov. For the implementation of four vector algebra we use the package `FeynCalc` [52].

A few general remarks: The computation is done in light cone gauge[14] ($A_+ = 0$) which allows to neglect the non-interacting part of the Wilson lines. The divergent integral is computed in dimensional regularization, and we solely extract the $1/\epsilon$ pole part. For most diagrams it can be seen with some experience which contributions matter. The propagators in d dimensions for the quark and gluon field we obtain from [53]. They are

$$\overbrace{\psi(x)\bar{\psi}}(0) = \frac{\Gamma\left(\frac{d}{2}\right)}{2\pi^{\frac{d}{2}}} \frac{\mathrm{i}\not{x}}{(-x^2 + \mathrm{i}\varepsilon)^{\frac{d}{2}}}, \tag{3.19}$$

$$\overbrace{A^a_\mu(x) A^b_\nu}(0) = \frac{-1}{4\pi^{\frac{d}{2}}} \frac{\Gamma\left(\frac{d}{2} - 1\right) g_{\mu\nu}\delta^{ab}}{(-x^2 + \mathrm{i}\epsilon)^{\frac{d}{2}-1}}. \tag{3.20}$$

[13] In light-cone notation $\gamma^+ = n_\mu \gamma^\mu$.

[14] For background fields. Dynamical gluons have a different gauge choice, such that we obtain the regular propagators for them.

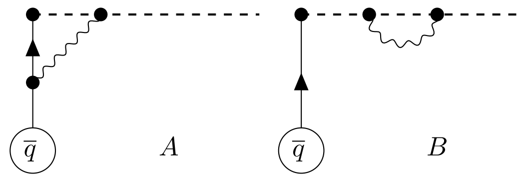


Figure 7: One loop correction to the quark TMD operator of twist*-1. Black dots indicate the location of quantum fields (vertices), the encircled $\overline{q}$ represents the remaining quark background field.

The relevant Γ factor develops as follows. The propagators combined with derivatives acting on them (in most diagrams) via eq. (3.14) determine the power of the denominator. By joining the propagators at the cost of introducing integration variables all Γ factors from propagators are replaced by a single Γ function. The argument of the Γ function is the sum of the previous arguments, equal to the total power of the denominator. The angular integral in d dimensions reduces the argument of the Γ effectively by $d/2$ plus one per pair of coordinates in the numerator. In all diagrams we thus find that, by collecting the correct power of coordinate expressions (x^λ, y^λ) in the numerator of a d dim. integral, we can always restore the factor of $\Gamma(\frac{d}{2}-2)$ - as in the end we are interested in the result for $d \to 4$. The whole computation is contained in d dimensions using the above formulae for the propagators. We isolate and solely compute terms which contribute to the expression $\sim \Gamma(\frac{d}{2}-2)$ and use

$$d = 4 + 2\epsilon, \tag{3.21}$$

$$\Gamma(\frac{d}{2}-2) = \frac{1}{\epsilon} \tag{3.22}$$

for the final result.

C.a Quark twist 1

The corrections to twist*-1 can be computed ad-hoc from the diagrams in figure 7 with standard methods. The fermion-boson interaction term is given in eq. (3.13) and quantum fields need to be inserted along the Wilson line as depicted in the diagrams. The position of the inserted dynamical gluon B is integrated over along the gauge link. In the case of Diagram A the limits for the single integration are $z_1, \pm\infty$. Diagram A, denoted as $\mathcal{Q}_A$, corresponds to the expression

$$\mathcal{Q}_A = \left(\mathrm{i}g \int \mathrm{d}^d x \overline{q}(x) \not{B}(x) q(x)\right) \overline{q}(zn)\, [zn, -\infty n] \left[\mathrm{i}g \int_{-\infty}^{z} \mathrm{d}\sigma B^+(\sigma n)\right]. \tag{3.23}$$

One contraction contributes[15], and the colour factor for this diagram can be computed respecting the order of the operator: The colour structure of the quark is convoluted with the structure of the gluon field

$$B_+(\sigma) = B_+^a(\sigma)\lambda_{bc}^a. \tag{3.24}$$

The d-dim. integral can be taken, where we again stress that the presented result corresponds only to the UV divergent part and *good* components.

The bare expression for diagram A reads

$$\mathcal{Q}_A = \frac{2\alpha_s}{4\pi\epsilon} C_F \int_{-\infty}^{z} \mathrm{d}\sigma \int_0^1 \mathrm{d}u \overline{u} \partial_+ \overline{q}(u\sigma n), \tag{3.25}$$

[15] This is not obvious; as seen later, for other diagrams there are several contributions(contractions) which do not simply add up to a symmetry factor.

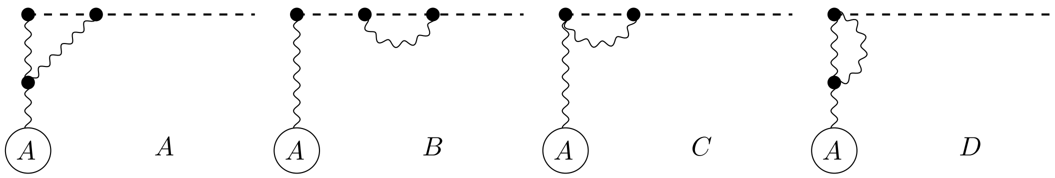


Figure 8: One loop correction to the gluon TMD operator of twist*-1. Black dots indicate the location of quantum fields (vertices), the encircled A represents the remaining gluon background field.

and it still contains the infinite integral along the Wilson lines contour. To regulate it, we make use of the so-called δ-regulator [45], with which both integrals can be eliminated simultaneously. Finally, the diagrams are

$$\mathcal{Q}_A = \left[1 - \ln\left(\frac{-\mathrm{i}\hat{p}_+^z}{\delta}\right)\right] \overline{q}(zn), \tag{3.26}$$

$$\mathcal{Q}_B = 0. \tag{3.27}$$

A formal derivative on the argument of the external quark field appearing in the logarithm. The sign of the derivative depends on the $\pm\infty$ inside the definition of the TMD operator, which in turn depends on the considered process [12].

Any "self-interaction" of a single Wilson line, as it appears in diagram B, yields zero. These terms are proportional to the light cone vector squared $n^2 = 0$ due to the propagator of the corresponding gluon. Therefore this and many of the appearing diagrams do not contribute.

C.b Gluon twist 1

In this subsection we show the contributing diagrams to the gluon TMD operator of twist-1 eq. (3.6), see figure 8. The gluon TMD corrections involve the same contours as those for the quark operator. Here the fermion propagators are replaced by a gluons, and the (anti-)quark-gluon interaction is replaced by the three-gluon interaction term in eq. (3.13). This results in a derivative to the propagator. In general, it appears that the gluon diagrams have the benefit of an absent Dirac algebra, but this comes at the cost of derivatives acting on the propagators (enormously) complicating the calculation. The additional diagrams $\mathcal{G}_C, \mathcal{G}_D$ result from the non-abelian structure in eq. (3.2), due to which two gluons can couple to the field located at the end of the Wilson line. The results of the diagrams are

$$\mathcal{G}_A = \frac{-g^2 C_A}{16\pi^{\frac{d}{2}}\epsilon}\left(1 - \ln\left(\frac{-\mathrm{i}p_+^z}{\delta_L}\right)\right)\mathcal{X}_\mu^A(z), \tag{3.28}$$

$$\mathcal{G}_B = 0, \tag{3.29}$$

$$\mathcal{G}_C = \frac{-g^2 C_A}{16\pi^{\frac{d}{2}}\epsilon}\left(1 - \ln\left(\frac{-\mathrm{i}p_+^{z_2}}{\delta_L}\right)\right)\mathcal{X}_\mu^A(z), \tag{3.30}$$

$$\mathcal{G}_D = \frac{g^2 C_A}{4\pi^{\frac{d}{2}}\epsilon}\mathcal{X}_\mu^A(z) \tag{3.31}$$

C.c Quark twist 2

Here we discuss the one loop correction to the quark TMD operator of twist*-2. First we analyse the structure of the corrections. The most interesting contours that appear as corrections are shown in figure 9. Diagram C and

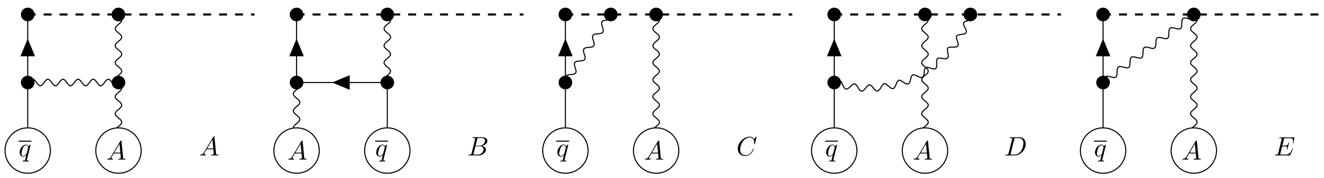


Figure 9: (Incomplete) set of one loop corrections to the quark TMD operator of twist*-2. Black dots indicate the location of quantum fields (vertices), the encircled q (A) represents the remaining quark (gluon) background field.

D both correspond to diagram A for the quark twist*-1. However, they cannot by simply joined to a single diagram due to the different colour structure. The same happens for diagrams with a gluon between the Wilson line and the three-gluon vertex.

There are also corrections that already appear for the gluon twist*-1 correction, plus trivial corrections. For a more detailed discussion we refer to [45].

The diagrams A and B are an ordinary loop correction to the legs of the external fields. The expression for diagram A (denoted as $\mathbb{Q}_A$) is

$$\mathbb{Q}_A = \overline{\psi}^A(z_1)F^D_{\mu+}(z_2)\lambda^D_{AC}\Big(+ \mathrm{i}g\int \mathrm{d}^d x\overline{q}^X\gamma^\nu B^Z_\nu\lambda^Z_{XY}\psi^Y\Big)\Big(- \mathrm{i}g\int \mathrm{d}^d yA^{A'}_\iota\partial_\alpha B^{B'}_\beta B^{C'}_\gamma\Big)v^{\iota\alpha\beta\gamma}_{A'B'C'} \tag{3.32}$$

$$+\overline{\psi}^A(z_1)F^D_{\mu+}(z_2)\lambda^D_{AC}\Big(+ \mathrm{i}g\int \mathrm{d}^d x\overline{q}^X\gamma^\nu B^Z_\nu\lambda^Z_{XY}\psi^Y\Big)\Big(- \mathrm{i}g\int \mathrm{d}^d yA^{A'}_\iota\partial_\alpha B^{B'}_\beta B^{C'}_\gamma\Big)v^{\iota\alpha\beta\gamma}_{A'B'C'}, \tag{3.33}$$

and has two distinct terms due to the vertex term for (1+2) gluons. The arguments of the vertex terms are implicitly clear. The first term has one propagator without derivatives, and one propagator on which two derivatives act. The second term has two propagators, and one derivative acts on each of them. Due to the vertex structure there are two terms that cannot be added up by introducing a symmetry factor like it can be done for symmetric vertices.

By joining the three propagators in eq. (3.32), three Feynman integration variables $\{\phi,\chi,\psi\}$ are introduced. The integration variables are each bound to the domain $[0,1]$ and fulfil

$$\phi+\chi+\psi = 1. \tag{3.34}$$

We find it convenient to identify the combinations

$$\Lambda = \phi\chi+\chi\psi+\psi\phi, \qquad \phi' = \frac{\chi\psi}{\Lambda}, \qquad \chi' = \frac{\psi\phi}{\Lambda}, \qquad \psi' = \frac{\phi\chi}{\Lambda}, \tag{3.35}$$

which satisfy eq. (3.34) as well as the original set of variables. The (two dimensional) Jacobian $\mathcal{J}$ for this transformation $\mathcal{F}$ is

$$\mathcal{F} : (\phi,\chi,\psi) \mapsto (\phi',\chi',\psi')\,, \tag{3.36}$$

$$\mathcal{J}\left(\mathcal{F}\right) = \frac{\phi\chi\psi}{\Lambda^3}. \tag{3.37}$$

This type of combinations also appear in the twist-2 gluon corrections.

With the methods discussed, we perform the calculation of the diagrams for the quark twist*-2 operator. We do not conclude their computation in here, and refer to [45] where we published our results.

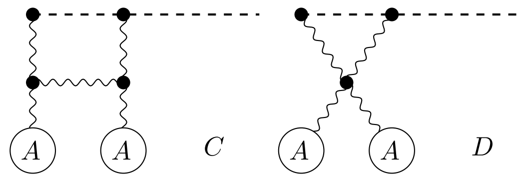


Figure 10: (Incomplete) set of one loop corrections to the gluon TMD operator of twist*-2. Black dots indicate the location of quantum fields (vertices), the encircled A represents the remaining gluon background field.

C.d Gluon twist*-2

In this section we demonstrate our developed formalism by computing the corrections to the gluon twist*-2. The results to this are novel to our best knowledge. They have been developed in collaboration with Óscar del Río of the Complutense University of Madrid. The most interesting diagrams that appear are shown in figure 10. In the following these (two) diagrams are analysed and calculated. All other diagrams are a (modified) repetition of gluon twist*-1 and quark twist*-2 diagrams, and thus the remaining diagrams and results are listed in appendix A.

Diagram $\mathbb{G}_C$. This diagram appears to be the most complicated diagram in this computation. it has the same contour as the quark diagram $\mathbb{Q}_A$, but its computation is much more complicated due to the derivatives acting on the propagator, due to which a manifold of terms need to be considered. Therefore this section deals with the detailed analysis and computation of the $\mathbb{G}_C$ split into steps.

To begin with the calculation, a common colour factor can be extracted for all contractions, differing only in relative signs. The expressions that need to be computed then are

$$
\begin{aligned}
\mathbb{G}_C = \mathrm{i}f^{A'D'E}\frac{C_A}{2}\frac{(-\Gamma(\frac{d}{2}-1))^3(\mathrm{i}g)^2}{(4\pi^{\frac{d}{2}})^3}\int \mathrm{d}^dx\mathrm{d}^dyA_\iota^{A'}A_\kappa^{D'}v^{\iota\alpha\beta\gamma}v^{\omega\lambda\eta\theta}& \\
\Big\{\partial_{x^\alpha}\partial_{y^\lambda}\Delta_{\beta\eta}(x-y)\partial_{z^\mu}\Delta_{\gamma M}(z_1n-x)\partial_{z^\nu}\Delta_{\theta N}(z_3n-y)& \\
+\Delta_{\gamma\theta}(x-y)\partial_{z^\mu}\partial_{x^\alpha}\Delta_{\beta M}(z_1n-x)\partial_{z^\nu}\partial_{y^\lambda}\Delta_{\eta N}(z_3n-y)& \\
-\partial_{x^\alpha}\Delta_{\beta\theta}(x-y)\partial_{z^\mu}\Delta_{\gamma M}(z_1n-x)\partial_{z^\nu}\partial_{y^\lambda}\Delta_{\eta N}(z_3n-y)& \\
-\partial_{y^\lambda}\Delta_{\eta\gamma}(x-y)\partial_{z^\mu}\partial_{x^\alpha}\Delta_{\beta M}(z_1n-x)\partial_{z^\nu}\Delta_{\theta N}(z_3n-y)\Big\}&
\end{aligned}
\tag{3.38}
$$

The components of this calculation are as follows. We apply a standard procedure to each of the terms above, which consists of

1. joining the propagators by introducing Feynman integration variables,
2. shifting the position space integration to bring the d dimensional integral into a symmetric form,
3. expanding the remaining (external field(s)) A in terms of small positions [53],
4. collecting terms that contribute to the divergent term $\Gamma(\epsilon)$.

Diagram C is symmetric as indicated in figure 11. Therefore we take into account a symmetry factor of 2 at the end of our computation. A more technical aspect of exploiting this symmetry the solidness of the computation at each step. The property can be checked at each step in the calculation, unless the symmetry is violated.

Step two is the shift of the integration variable, such that the denominator in the loop integral has quadratic form. Analysing the structure of diagram C it can be seen that the denominator D_C of the loop integral takes the form

$$D_C = \left[-\chi(x - z_2 n)^2 - \psi(y - z_1 n)^2 - \varphi(x - y)^2\right]^m , \tag{3.39}$$

where χ, ψ, φ are integration variables. The powers m in the denominator that can appear are $\{\frac{3d}{2}, \frac{3d}{2} \pm 1\}$[16]. To take a single of the spatial integrals can be done by decoupling the respective variable from the others by finding the correct shift. This step violates the diagrams symmetry, by choosing a preferred variable to integrate out first. The symmetry must be restored by taking the second integral, such that, in principle, there is no problem. However, due to the asymmetric shifts from a single integration, it appears to be very difficult to reorganise the terms to again find symmetric expressions. Due to the amount of terms that arise and the (non-commuting) shifts from the individual integrations, we abandon this direct approach as it is unpractical.

We propose to shift the variables in a symmetric manner, to decouple them from the light cone vector n only, and keep a mixing term $\sim xy$ in the denominator. The expression is symmetric, and we can take both integrals at a time in a symmetric way. In the variables of figure 11 and eq. (3.39), and the conjugated variables from eq. (3.35), the shift that symmetrically decouples the integration variables x and y from the light-cone direction n is

$$x \mapsto x + z_1\chi' n + z_2\overline{\chi}' n \qquad y \mapsto y + z_1\psi' n + z_2\overline{\psi}' n. \tag{3.40}$$

With this shift, the denominator takes the desired form

$$D_C \mapsto \left[-(\phi + \chi)x^2 - (\phi + \psi)y^2 - 2\phi xy\right]^m . \tag{3.41}$$

To evaluate both integrals, we use a symmetric expression. The advantage of combining both integrals is the enormous simplification of the calculation, which becomes very involved whyn considering the integrals separately instead. The result of the two-fold integral then also is obtained in a term-by-term symmetric form[17]. Such an expression can be derived straight forward from integrating one variable at a time.

A double integral for this problem could be

$$\int \mathrm{d}^d x \int \mathrm{d}^d y \frac{x^\mu y^\nu}{(-(\phi + \chi)x^2 - (\phi + \psi)y^2 + 2\phi xy + M^2)^{\frac{3d}{2}-1}} = \frac{\pi^4 \phi}{\Lambda^3} \frac{g^{\mu\nu}}{2} \frac{\Gamma(\epsilon)}{\Gamma(\frac{3d}{2} - 1)} (M^2)^{d-n+1}. \tag{3.42}$$

In these type of integrals we keep a mass regulator M for the gluon propagators. It does not contribute to the $1/\epsilon$ pole, and to us is a consistency feature only. To obtain the pole we assume $n = 3$. We here evaluate in $d = 4$ to obtain the desired term, and thus have the explicit prefactor independent of $d(\epsilon)$. In the Γ factor in the denominator it is convenient to keep the implicit form because this factor cancels symbolically. A complete list of the double integrals in d dimensions used in the computation is listed in appendix A.

In contrast to the other diagrams, in diagram C we a mixing of different powers of the denominator. The different powers arise from the derivatives on the propagators. A double derivative results in two terms: One term with two derivatives acting on the denominator only, and one term with the derivative acting on the numerator created by the first derivative. The second of these terms produces a metric tensor of the Lorentz indices of the derivatives.

Terms with this structure also appear for other diagrams, but they do no contribute as they are $\propto n^T$, where T is a transverse index. In diagram C there is similar behaviour: For the two double derivatives appearing in one contraction, the term producing the double metric contribution does not contribute for the discussed reason. The

[16]It turns out that $\frac{3d}{2} - 1$ does not contribute. This can be understood already from an analysis of the resulting numerator structure of terms resulting in this power of the denominator.

[17]That is, for symmetric terms as an input to the formula.

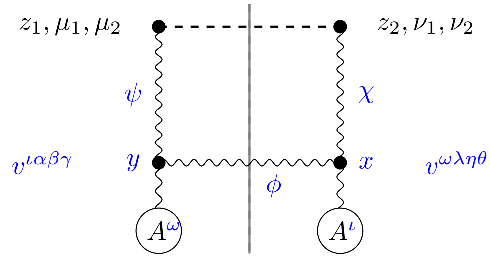


Figure 11: Symmetric display of diagram C for gluon twist*-2. The grey line is the axis of symmetry. The blue (black) colour labels internal variables (external) variables. Colour group indices are omitted.

terms with the two denominator powers $3d/2$, $3d/2+1$ are not zero by contraction and contribute to the diagram.

The power of the denominator determines how many terms of the integration variables need to be collected in the numerator of the loop integral structure to obtain the $1/\epsilon$ pole (e.g. eq. (3.42)). We separately treat both contributing powers of the denominator in independent computations for that reason. In the end their contributions are combined. Both structures individually carry the symmetry of figure 11. For the case of $3d/2$ as the power of the denominator, the combined power in the integration variables must be four to obtain $\Gamma(\epsilon)$.

For the computation carried out in the described way, we are able to obtain the result for diagram C. The symmetry of the diagram is preserved in the result. The individual components also obey relations that satisfy this symmetry; this has been explicitly checked. The explicit result can be found in appendix A. Here we present the parametrized form for diagram C as

$$
\begin{aligned}
\mathbb{G}_C = \int [\mathrm{d}\alpha \mathrm{d}\beta \mathrm{d}\gamma] & \frac{-2\mathrm{i}g^2 C_A f^{A'D'E}\chi\psi}{32\pi^2\epsilon\Lambda} \\
& \Big(g_{\mu\nu} A^{A'}_{\alpha}(z_3\psi' n + z1\overline{\psi}' n) A^{D'}_{\alpha}(z_1\chi' n + z_3\overline{\chi}' n) \left\{\mathcal{F}_1 \partial^2_{\psi+} + \mathcal{F}_2 \partial^2_{\chi+} + \mathcal{F}_3 \partial_{\psi+}\partial_{\chi+}\right\} \\
& + A^{A'}_{\nu}(z_3\psi' n + z1\overline{\psi}' n) A^{D'}_{\mu}(z_1\chi' n + z_3\overline{\chi}' n) \left\{\mathcal{G}_1 \partial^2_{\psi+} + \mathcal{G}_2 \partial^2_{\chi+} + \mathcal{G}_3 \partial_{\psi+}\partial_{\chi+}\right\} \\
& + A^{A'}_{\mu}(z_3\psi' n + z1\overline{\psi}' n) A^{D'}_{\nu}(z_1\chi' n + z_3\overline{\chi}' n) \Big\{\mathcal{H}_1 \partial^2_{\psi+} + \mathcal{H}_2 \partial^2_{\chi+} + \mathcal{H}_3 \partial_{\psi+}\partial_{\chi+}\Big\}\Big)
\end{aligned}
\tag{3.43}
$$

where all appearing labels are rational functions of (ϕ, χ, ψ) to the unique power of (-3). As mentioned in the second point, we respect the symmetry of diagram C at each step in the computation.

Diagram $\mathbb{G}_\mathbf{D}$. The structure of the diagram $\mathbb{G}_D$ is the first contribution we find that includes the four gluon interaction term of eq. (3.13). The correction can be computed either with the effective action in BFM, or by applying the Feynman rule eq. (3.17) consistent with [51]. The diagram matches the expression

$$
\begin{aligned}
\mathbb{G}_D = {} & g^2 f^{abE} \frac{C_A}{2} 4\left(g_{\alpha\mu_2} g_{\nu_2\beta} - g_{\alpha\nu_2} g_{\beta\mu_2}\right) \\
& \times \int \mathrm{d}^d x \partial_{\mu_1}\Delta_(z_1 - x)\partial_{\nu_1}\Delta(z_2 - x) A^a_\alpha(x) A^b_\beta(x) \\
& - (\mu_1 \leftrightarrow \mu_2) - (\nu_1 \leftrightarrow \nu_2) + (\mu_1 \leftrightarrow \mu_2, \nu_1 \leftrightarrow \nu_2)\,,
\end{aligned}
\tag{3.44}
$$

with a convenient short hand notation regarding the field indices. After antisymmetrizing $\mu_{1,2}$ in eq. (3.44) the indices can be restored, $\mu_1 = \mu$, $\mu_2 = +$, and identical for $\nu_{1,2}$. From this expression the integral can be taken, by expanding the fields $A(x)$ for small positions, and selecting terms which contribute to the $1/\epsilon$ term.

Unlike the other diagrams, here both external fields A originate from the same vertex. Since both gluon fields are at the same position, they are equivalent in terms of colour ordering along the Wilson line. In order to match an expression with two colour ordered fields, we use the relation,

$$A_\mu(xn) = -\int_L^x \mathrm{d}w F_{\mu+}(wn). \tag{3.45}$$

and with eq. (3.45) we have our result

$$\mathbb{G}_D = \frac{-g^2 C_A}{4\pi^2\epsilon} \int_0^1 \mathrm{d}u(\overline{u}-u) \int_L^u \mathrm{d}w \left(\mathcal{X}^E_{\mu,\nu}\left(z^u_{21}n, z^w_{21}n\right) - \mathcal{X}^E_{\nu,\mu}\left(z^u_{21}n, z^w_{21}n\right)\right). \tag{3.46}$$

The relative sign in eq. (3.46) reflects the colour ordering along the Wilson line. The short hand notation

$$z^u_{ij} = z_i + u\left(z_j - z_i\right) \tag{3.47}$$

describes the connecting line between z_i and z_j along the parameter u. This combination appears in most of the diagrams to gluon twist*-2.

IV Extraction of non perturbative TMDPDFs

In order to motivate the study of TMDs we summarize here a discussion of [4]. It has been found that the differential cross section of the transverse momentum of produced bosons (on the example of Z boson) in pp collision can be described very well for high transverse momenta of the virtual boson q_T. However, with decreasing q_T the description of the data eventually fails. The given explanation [4] is that, close to the energy threshold at which Z production becomes physical, the process is dominated by QCD radiation – and hence a standard cutoff of the perturbative series is not justified. There exists a transition region in which the momentum scale q_T and the coupling can be coupled such that the data can be described by a rearrangement of the collinear factorization terms. This procedure is called resummation, and applies for moderate q_T. For small q_T this ansatz, too, cannot describe the experimental results [44].

Phenomenology of TMDPDFs is a research field since the differential cross section of transverse momentum dependent short lived bosons has been worked out in terms of TMDPDF and TMDFFs eq. (2.69) eq. (2.101) First studies have been provided by [54].

In this section we provide an overview of the explicit methods of extractions of TMDs, summarizing the TMDPDFs and the TMDFFs but also the RAD in one name. We then discuss the models and the outcome for the TMD functions for the extraction from DY data.

The TMD distributions can be categorized into eight distributions for the TMDPDF [42]. The unpolarized distributions for quarks are the best constrained by the available data. In the current state of the art, the unpolarized distributions for TMDs are extracted only[55, 11, 12, 56], and this is the case here as well. This is why the other distributions are not further discussed here.

All the results of presented in the following (section B, C, D) have been produced in the context of this thesis. They are product of the collaboration with my colleagues P. Zurita, A. Vladimirov and I. Scimemi. The contents the TMDPDF extraction, i.e. the results in section B, are already published in [9] and we call this work ART23. The work about the TMDFF section C and pion TMDPDF section D is very recent and the there contained results are published in this thesis for the first time.

A TMD functions and their parametrization

In this section we discuss the framework for the extraction of TMD distributions, present the model for the extraction ART23 in section B and give an overview of ansätze found in the literature.

A.a TMD distributions

The TMDPDF is a non perturbative function defined in eq. (2.62) and, in principle, independent of collinear PDFs. However, the TMDPDFs can be matched to the collinear PDFs in the limit $b \to 0$. This is achieved by OPE of the TMD operator eq. (3.3). Like a collinearly separated operator can be expressed in terms of local operators, here the OPE reduces the transverse separation. This technique of matching reduces the problem of finding the TMDPDF as a function of x, $\boldsymbol{b}$ and possibly also the flavour of the quark. It does not, however, remove the fact that the TMDPDF is a non perturbative function, like the matrix element of the collinear operator, the collinear PDF. The relation discussed in the following is immensely useful for TMD phenomenology, but it is not necessary. TMDPDFs can be independently both understood and determined as functions, without the requirement of matching them to PDFs.

The forward matrix element of a state with large momentum P and polarization S of the TMD operator turns to the collinear PDF operator in the limit $\boldsymbol{b} \to 0$. For the limit $\boldsymbol{b} = 0$ the expression is exact,

$$\lim_{\boldsymbol{b}\to 0} F_{1,f\leftarrow h}(x,b) = \sum_{f'} \int_x^1 \frac{\mathrm{d}y}{y} \sum_{n=0}^{\infty} \alpha_s^n(\mu_{\mathrm{OPE}}) C_{f\leftarrow f'}^{[n]}\left(\frac{x}{y}, \mathbb{L}_{\mu_{\mathrm{OPE}}}\right) f_{f'\leftarrow h}(y,\mu_{\mathrm{OPE}}), \tag{4.1}$$

with $f_{f'\leftarrow h}$ the unpolarized collinear PDF, and

$$\mathbb{L}_\mu = \ln\left(\frac{\boldsymbol{b}^2\mu^2 e^{2\gamma_E}}{4}\right). \tag{4.2}$$

$\gamma_E \approx 0.577$ is the Euler-Mascheroni constant, and $C_{f\leftarrow f'}^{[n]}$ denote the coefficient functions from the OPE. At LO the coefficient function

$$C_{f\leftarrow f'}^{[0]} = \delta_{ff'}\delta(1-x). \tag{4.3}$$

The remaining functions, up to N^{3}LO, are given in the published work [9]. The parameter μ_{OPE} is the scale of the OPE. It can be chosen freely as it is not connected to the scales of TMD factorization. We choose the scale value

$$\mu_{\mathrm{OPE}} = \frac{2e^{-\gamma_E}}{\boldsymbol{b}} + 2\,\mathrm{GeV}. \tag{4.4}$$

The effect of defining the intrinsic scale as in eq. (4.4) is that eq. (4.2) vanishes in the limit $b \to 0$. The offset of $2\,\mathrm{GeV}$ guarantees to avoid the non perturbative region even for large $\boldsymbol{b}$. This choice, including the offset, is consistent with [12].

The TMDPDF for non vanishing b is given by a modification of eq. (4.1):

$$F_{1,f\leftarrow h}(x,\boldsymbol{b}) = \sum_{f'} \int_x^1 \frac{\mathrm{d}y}{y} \sum_{n=0}^{\infty} \alpha_s^n(\mu_{\mathrm{OPE}}) C_{f\leftarrow f'}^{[n]}\left(\frac{x}{y}, \mathbb{L}_{\mu_{\mathrm{OPE}}}\right) f_{f'\leftarrow h}(y,\mu_{\mathrm{OPE}}) F_f^{\mathrm{NP}}(x,\boldsymbol{b}), \tag{4.5}$$

where the ansatz F_f^{NP} is the non perturbative model to describe TMD factorisation at *moderate*[18] $\boldsymbol{b}$.

The requirements for the NP ansatz are:

- F_f^{NP} must be an even function in $\boldsymbol{b}$, because the TMDPDF must be a symmetric function in $\boldsymbol{b}$, and also the expansion in $\boldsymbol{b}^\mu$ of the TMDPDF yields a composition of even functions (up to a global factor resulting in an antisymmetric term) [42, 50];

- the ansatz must not violate the small-$\boldsymbol{b}$ matching eq. (4.1). Thus it satisfies the boundary condition

$$\lim_{\boldsymbol{b}\to 0} F_f^{\mathrm{NP}}(x,\boldsymbol{b}) = 1. \tag{4.6}$$

Another important constraint on F_f^{NP} is that it must decrease rapidly at larger $\boldsymbol{b}$. The resulting TMDPDF must be an integrable function such that it can describe the differential cross section. This implicitly means that the logarithms $\sim \ln(\boldsymbol{b}^2)$ inside the coefficient function need to be supressed.

The purpose of this ansatz is to contain all the non perturbative information such that eq. (4.5) results to be the TMDPDF. Its parametrization is determined by comparing to TMD relevant experimental data. Finding the best form for the non-perturbative part of the TMDPDF is the key aspect of phenomenological works.

A list of ansätze for F_f^{NP} is given in table 1. In the past, the flavour-independent ansatzes were formulated,

[18] A discussion on what *moderate* precisely stands for in this context is given in the next section.

reference	NP par.	F_f^{NP}	$N_{par.}$
SV19 [12]	κ	$\exp\left(-\frac{(\kappa_1\overline{x}+\kappa_2 x+\kappa_3\overline{x}x)\,\boldsymbol{b}^2}{\sqrt{1+\kappa_4 x^{\kappa_5}\boldsymbol{b}^2}}\right)$	5
PV19 [55]	$g^*,\ \lambda$	$\frac{1}{2\pi}\exp\left(-g_{1a}\frac{\boldsymbol{b}^2}{4}\right)\left(1-\frac{\lambda g_{1a}^2}{1+\lambda g_{1a}}\frac{\boldsymbol{b}^2}{4}\right)$	4
MAP22 [60]	$g^*,\ \lambda$	$\frac{g_1(x)e^{-g_1(x)\frac{\boldsymbol{b}^2}{4}}+\lambda^2 g_{1B}^2(x)\left[1-g_{1B}\frac{\boldsymbol{b}^2}{4}\right]e^{-g_{1B}(x)\frac{\boldsymbol{b}^2}{4}}+\lambda_2^2 g_{1C}(x)e^{-g_{1C}(x)\frac{\boldsymbol{b}^2}{4}}}{g_1(x)+\lambda^2 g_{1B}^2(x)+\lambda_2^2 g_{1C}(x)}$	11

Table 1: Different ansatzes for used in the literature for F_f^{NP} in eq. (4.5). Due to differences in (a.o.) prescriptions and choice of collinear PDF the ansatzes cannot be compared exactly one to one. This table serves as an overview.

i.e., for all flavours, the function is identical. Notably a flavour independent ansatz leads to a flavour dependent TMDPDF nonetheless due to the splitting kernels $C_{f\leftarrow f'}^{[n]}$ and the (flavour dependent) collinear PDF in eq. (4.5).

In full analogy to TMDPDFs the above discussion can be repeated for the TMDFFs. The TMDFFs can be related to their collinear counterparts, and the matching is modified by a NP function D^{NP}, to absorb effects at moderate $\boldsymbol{b}$. It reads

$$D_{1,f\to h}(z,\boldsymbol{b})=\frac{1}{z^2}\int_z^1\frac{\mathrm{d}y}{y}y^2\mathcal{C}_{f\to f'}(y,L_{\mu_{\mathrm{OPE}}},\alpha_s\left(\mu_{\mathrm{OPE}}\right)d_{1,f'\to h}\left(\frac{z}{y},\mu_{\mathrm{OPE}}\right)D_{f\to h}^{\mathrm{NP}}(z,\boldsymbol{b}) \tag{4.7}$$

and has the same boundary condition as the NP TMDPDF,

$$\lim_{\boldsymbol{b}\to 0}D_{\mathrm{NP}}(z,\boldsymbol{b})=1. \tag{4.8}$$

as the TMDPDF. The ansatz in general depends on the flavour of the parton and the state in hadronises into. The TMDFFs already are hadron and flavour dependent via the collinear FFs and the coefficient functions.

The $\mathcal{C}_{f\to f'}$ are coefficient functions from the OPE, and the collinear matrix elements $d_{1,f'\to h}$ are the same distributions as in eq. (2.46)eq. (2.47). In [12] $\mathcal{C}_{f\to f'}$ are explicitly given to NLO. There it is explained as well that the coefficients are obtained by same pattern as the coefficients C for the TMDPDF are obtained. Thus they can be written down by using the form of $C_{f\leftarrow f'}$ with the replacement of the PDF DGLAP kernels $P_{f\leftarrow f'}^{(n)}(x)$ by the FF DGLAP kernels $P_{f\leftarrow f'}^{(n)}(z)$ (which can be found in [57]) and the functions $C_{f\leftarrow f'}^{(n,0)}(x)$ by the functions $\mathcal{C}_{f\leftarrow f'}^{(n,0)}(z)$ (can be found in [58, 59]).

The ansatz for the (proton) TMDPDF proposed in this work is

$$F_f^{\mathrm{NP}}(x,\boldsymbol{b})=\frac{1}{\cosh\left[\left(\lambda_1^f\overline{x}+\lambda_2^f x\right)\boldsymbol{b}\right]}. \tag{4.9}$$

The parameters $\lambda_{1,2}^f$ are explicitly flavour dependent, i.e. we distinguish between $f\in\{u,\overline{u},d,\overline{d},r\}$. For the remaining sea quarks, labelled by r, we do not have further flavour separation in our ansatz. The motivation for this ansatz is to have a parameter λ_1^f to control the behaviour at small x, where $\overline{x}=1-x$ is dominant. The parameter λ_2^f then should capture the information needed to describe the TMDPDF at moderate x. The two parameters can be translated to a constant and a linear term in the argument of the hyperbolic cosine.

The proposed ansatz contains 10 free parameters, which we obtain by matching to data. This formulation allows for a very simple functional form for each flavour. This certainly is an advantage of this framework and implicit confirmation of the necessity of flavour dependence in the ansatz.

A.b Rapidity anomalous dimension

The rapidity anomalous dimension (RAD), or Collins-Soper kernel (CS), is a NP function that describes the rapidity evolution of the TMDPDF eq. (2.110) (and also of the TMDFF). The RAD implicitly appears in physical observables (differential cross section) via the evolution from the initial scale to the physical scale (Q^2), see eq. (2.110).

In small-$\boldsymbol{b}$ the RAD can be expressed as a perturbative series, plus a NP part. It reads

$$\mathcal{D}(\boldsymbol{b},\mu) = \mathcal{D}_{\text{small-}\boldsymbol{b}}(\boldsymbol{b}^*,\mu^*) + \int_{\mu^*}^{\mu} \frac{\mathrm{d}\mu'}{\mu'}\Gamma_{\text{cusp}}(\mu') + \mathcal{D}_{\text{NP}}(\boldsymbol{b}). \tag{4.10}$$

The coefficients of Γ_{cusp} are obtained by higher order computation of the cross section and the small-$\boldsymbol{b}$ expansion of the RAD can be found in [49]. The intrinsic parameters b^* and μ^* are defined as

$$\boldsymbol{b}^*(\boldsymbol{b}) = \frac{\boldsymbol{b}}{\sqrt{\left(1+\boldsymbol{b}^2/B_{\text{NP}}^2\right)}}, \qquad \mu^*(\boldsymbol{b}) = \frac{2e^{-\gamma_E}}{\boldsymbol{b}^*(\boldsymbol{b})}, \tag{4.11}$$

and depend on the NP parameter B_{NP}. The definition has the effect that the logarithms in the perturbative series for the small-$\boldsymbol{b}$ RAD function vanish. The RAD depends on the transverse distance $\boldsymbol{b}$, which in eq. (4.10) is separated into two (overlapping) regions: the region where $\boldsymbol{b}$ is perturbative, i.e. the description of the RAD is achieved via the perturbative series, and the region of moderate $\boldsymbol{b}$, at which non perturbative physics is dominant. These two regions are separated (and thus defined) by B_{NP} (also: $\boldsymbol{b}_{\text{max}}$). This scale should be *roughly* located somewhere below $\Lambda_{\text{QCD}}^{-1} \approx 2\,\text{GeV}^{-1}$, when length scales are reached where perturbative physics is not applicable. We determine B_{NP} as a NP parameter in our extraction. Other works instead fix the parameter to a value, which is in the expected range [11, 8].

The RAD eq. (4.10) does not depend on the intrinsic scale μ^* in principle, but artificially so due to truncation of the perturbative series. The b^* function has linear behaviour for small values of b, and asymptotically turns to a constant for large b, and the expansion eq. (4.10) stays valid. Thus the NP part needs to compensate the truncation of the series and the effects of moderately large b. For the NP part of the RAD in previous studies [12, 56, 11, 8] typically two parameters are proven to be sufficient to obtain a good fit.

With a second parameter c_0 the ansatz for the NP part of the RAD in many works in the literature[19] is

$$\mathcal{D}_{\text{NP}}(b) = c_0 \boldsymbol{b}\boldsymbol{b}^*(b), \tag{4.12}$$

with obtained values shown in table 2.

Notably in table 2 the difference between the fits with NNLO and N^{3}LO is much smaller than the difference of including SIDIS data to the fit. This is observed for the combined fit of DY and SIDIS. The SIDIS data ranges to smaller resolution scales Q, to which the DY fit could not be tuned. The soft scale that separates perturbative $\boldsymbol{b}$ from non perturbative ones appears to be strongly correlated to the boundaries of the tested resolution scale. In the extraction ART23, we modified eq. (4.12) by adding a term $\sim bb^* \ln(b^*/B_{NP})$. We find that its introduction

[19] The ansatz eq. (4.12) is used in [56, 12]. The works in [11, 8] instead use $\mathcal{D}_{\text{NP}} = c_0\boldsymbol{b}^2$.

Extraction	$B_{NP}(\boldsymbol{b}_{max})$	c_0
SV19 DY [56]	1.7 ± 0.30	0.297 ± 0.006
SV19 DY+SIDIS (NNLO) [12]	1.93 ± 0.17	0.0391 ± 0.0063
SV19 DY+SIDIS (N^{3}LO) [12]	1.93 ± 0.22	0.0472 ± 0.0105
PV17 [11]	$2e^{-\gamma_E} \approx 1.123$	$0.5 \times (0.13 \pm 0.01)$
MAP22 [8]	1.5	$0.5 \times (0.248 \pm 0.008)$

Table 2: NP parameters obtained for the RAD as in eq. (4.12). The light grey entries are for a different prescription of the NP ansatz for RAD. Additionally, in these cases the $B_{\text{NP}}(\boldsymbol{b}_{\text{max}})$ is a fixed instead of a free parameter. The parameters can still be compared.

ad-hoc improves quality of the fit significantly. The ansatz for the RAD in ART23 is

$$\mathcal{D}_{\text{NP}}(b) = c_0 b b^*(b) + c_1 \ln\left(\frac{b^*}{B_{\text{NP}}}\right), \tag{4.13}$$

which means we introduce a total of 3 parameters to describe the RAD at small- and moderate $\boldsymbol{b}$ in our extraction.

In summary of the previous section we have the model for TMDPDF described by eq. (4.5)eq. (4.9) and the model for the RAD by eq. (4.10)eq. (4.13). In these functions we find a total of 13 free parameters to be determined by a fit to the experimental data;

$$\overrightarrow{\lambda} = \left\{B_{\text{NP}}, c_0, c_1, \lambda_1^u, \lambda_2^u, \lambda_1^{\overline{u}}, \lambda_2^{\overline{u}}, \lambda_1^d, \lambda_2^d, \lambda_1^{\overline{d}}.\lambda_2^{\overline{d}}, \lambda_1^r, \lambda_2^r\right\} \tag{4.14}$$

B Extraction of the proton TMDPDF from DY data

B.a Discussion of the setup

The obtained TMDPDF depends on various additional parameters which we specify in the following. First, for the χ^2 test function eq. (2.41), the correlation matrix needs to be defined. In this work it is taken to be

$$V_{ij} = \delta_{ij}\Delta^2_{i,\text{uncorr}} + \sum_c \Delta^c_{i,\text{corr.}}\Delta^c_{j,\text{corr.}}, \tag{4.15}$$

where $\Delta_{i,\text{uncorr}}$ is the sum of the squares of the uncorrelated uncertainties for datapoint i, and $\Delta^c_{i/j,\text{corr.}}$ is the c-th correlated uncertainty between the datapoints i and j. We further present our χ^2 result in a decomposed from,

$$\chi^2 = \chi^2_D + \chi^2_\lambda, \tag{4.16}$$

where χ^2_D represents the quality of the fit in terms of reproducing the shape of the data shape, whereas χ^2_λ represents the contribution associated to normalization discrepancies[20]. The definitions of the decomposed contributions are

$$\chi^2_D = \sum_i \frac{(m_i - t_i)^2}{\Delta^2_{i,\text{uncorr}}}, \qquad \chi^2_\lambda = \sum_c \lambda_c^2. \tag{4.17}$$

[20]We acknowledge that this notation is ambiguous, since the index λ could be identified with the parameters of the model from the previous section. This is not the case. The λ is associated with the mentioned nuisance parameters. We stick to this notation to be consistent with the literature [61].

The second sum is over all correlated errors with index c. The λ_c are the "nuisance" parameters, which can be computed for a theory prediction [61]. They allow for a decomposition of the theory prediction

$$t_i = \bar{t}_i - d_i, \quad d_i = \sum_c \lambda_c \Delta_{\text{uncorr}}{}^c, \tag{4.18}$$

where the sum is over all uncorrelated In the literature [61], it is precisely explained how the parameters λ_c can be computed. Essentially, it is an optimised correlated shift (to the theory) of all data to describe the experimental values. This decomposition is very useful to get an understanding for the decomposition of the χ^2 value. We also use the shift to display the theory prediction with the (estimated) shift in the experimental data being compensated. This is why in all data plots we show the value from the experiment and plot it against the shifted theory value $\bar{t}_i$.

Next, the expressions and implicitly the perturbative accuracies used are listed. In our prediction we rely on the $\text{N}^4\text{LO}(\sim \alpha_s^4)$ expression for

- the hard coefficient function C_V [13],
- the light-like quark anomalous dimension γ_V [13] and
- for the RAD $\mathcal{D}_{\text{small-}b}$ [62, 63],

and the $\text{N}^3\text{LO}(\sim \alpha_s^3)$ expression for the matching coefficient functions $C_{f\leftarrow f'}$ [64, 65].

For the data in this context, we speak of experiments which provide their results accessible via the expressions for the differential cross section eq. (2.99) in DY (and eq. (2.69) in SIDIS for section C). It is worth to mention that integrated cross sections from these experiments are already used for the determinations of the collinear PDFs. In this way the experiments as a whole quasi double contribute to the extraction of TMDPDF.

The collinear PDFs themselves ideally already describe the q_T-integrated cross section eq. (2.106) - we assume - perfectly and there is no ambiguity. In reality, different collaborations use different parametrizations, and eventually their results are in good agreement with each other [66]. Therefore the (a priori) arbitrary choice for the collinear PDF in eq. (4.5) *should* be irrelevant. Nonetheless in [10] it was found that

- the inclusion of the uncertainty of collinear PDFs in the fit is necessary to improve the reliability of the uncertainty predicted for the TMDPDFs;
- different collinear PDFs for the matching eq. (4.5) result in quite different shapes of the test function eq. (2.41) as a distribution of the replica;
- a different PDF choice leads to a significant difference in the extracted NP part F_f^{NP} and also the TMDPDF as a function itself. This is stressed because a difference in the parameters could originate in and also be compensated by the collinear PDF.

The results [10] are obtained with a flavour dependent ansatz for the NP b profile, as it is done in this work. Therefore the above enumerated findings are especially relevant for this extraction.

The collinear PDF used to obtain the published result ART23 is `MSHT20PDF` [14]. The choice is reasoned with the good distribution for the χ^2 test function as a variable of the PDFs Monte Carlo replicas[21]. Compared to the other PDF sets `NNPDF3.1` [67], `HERAPDF2.0` [68] and `CT18` [69], `MSHT20PDF` shows the best properties in the following criteria:

- the distribution is the most compactly centred as expected from a χ^2 distribution;

[21]The method to determine the distribution of theory predictions is illustrated in appendix B.A.

- both replica distributions for either varying the data or the collinear input PDF yield the same average value, which is a good feature.

The above are positive arguments to support the choice of `MSHT20PDF`. Moreover, the other PDFs are discarded for the following reason(s):

`NNPDF3.1`. There is a stronger difference between the resulting χ^2 distribution of the replica from data variation to the replica from PDF set variation. The latter contributes with large outliers.

`HERAPDF2.0`. This PDF was not used because it was obtained with HERA (DIS) data only[22] [12].

`CT18`. For this PDF set the χ^2 distribution is the most spread out.

Let us note here that the obtained χ^2 values are the best for `HERAPDF2.0` and `CT18` in [10]. Typically the comparison of fit results is broken down to this number. The average/best-χ^2 value standalone is not the criterium by which the collinear PDF is selected.

To conclude the discussion of the framework, we discuss the data used for the TMDPDF extraction and the usage of libraries. A detailed overview of the data sets and their kinematic ranges in resolution scale Q and rapidity y is given in [9] and we do not repeat it here. To summarize the discussion we list the contributing experimental setups. The experimental data for DY lepton-pair production for this fit are measurements of proton-proton collisions taken at the LHC (ATLAS, CMS, LHCb) and RHIC (PHENIX, STAR), of proton-antiproton collisions at Tevatron (CDF, D0) and proton-nucleus collisions at Fermilab (E228, E606, E772).

The kinematic cuts imposed on all the data from the listed experiments are

$$\frac{\langle q_T\rangle}{\langle Q\rangle} \equiv \delta < 0.25 \quad \text{and} \quad \left(\delta^2 < 2\sigma \quad \text{or} \quad \langle q_T\rangle < 10\,\text{GeV}\right), \tag{4.19}$$

where the bin-average values for q_T and Q are used. They are part of the experimental measurement and reported with the datapoint. σ is the relative uncorrelated uncertainty. The second condition removes datapoints which have a large δ and, at the same time a small, uncorrelated uncertainty. It prevents the fit to try to describe data which are on the edge of the TMD region but contribute strongly due to their small uncertainty eq. (4.15). With these cuts, there are a total of 627 datapoints in the fit.

The gradient descent method used is implemented in the `migrad` function of the `minuit` package [71]. The numerical computation of the cross section, i.e. all the involved calculation from the distributions to the evolution is implemented in the `artemide`[72] package. To name only too aspects which contribute to this, we did not discuss the effects of detector design related cuts in the final state phase space for the leptons for DY or SIDIS. Also the datapoints provided by experiment can be integrated in one of the variables, and to compare the differential cross section needs to be integrated over the same region as in the experiment[23]. All of this is handled in the package. Implicitly the package uses LHAPDF [73] for the values of the PDFs α_s. The storage and replica generation of data and among others the computation of the χ^2 function is taken care of by the package `artemide-DataProcessor`.[74].

B.b Results

In the following we present the results of this work and discuss what can be learned from them. In figure 12 we present the obtained distribution of χ^2 values for fitting randomly generated replica of the data and using a random replica of the collinear PDF. The distribution serves as result to everything that follows. I.e., we use 1000 of the generated fits as our ensemble, from which we compute the

[22] The SV19 NNLO extraction with this set produced an outstandingly low χ^2 with values of 0.97 (DY) and 0.76 (SIDIS) [12]. It would be intresting to compare the extraction of ART23 to a result using this PDF set, given the good quality of the fit. The ansatz to use this PDF set to then describe DY data is pursued in MCEG studies [70] as well .

[23] For a more detailed discussion of the effects and the control of fiducial cuts and bin averaging we refer to [12, 9].

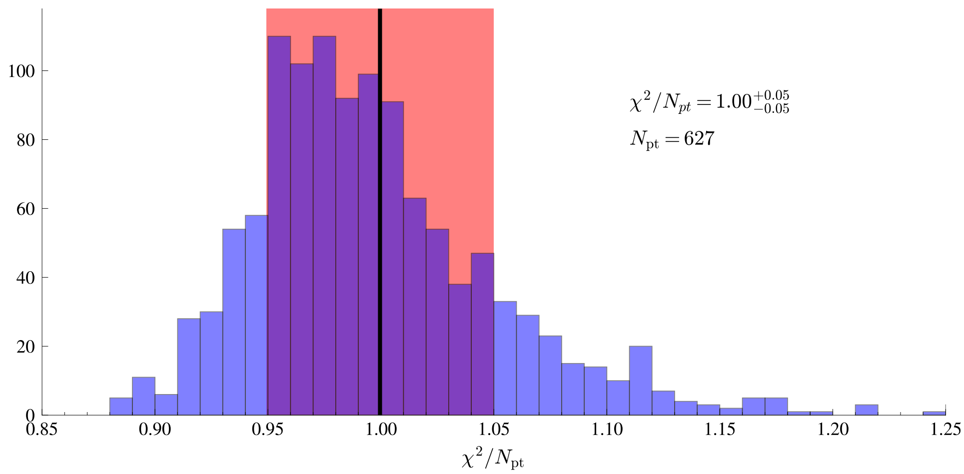


Figure 12: The distribution of the χ^2 over all generated replicas, without cuts or restrictions. The black line indicates the average value, and the red band the 68%CI.

- obtained parameter values,
- theory prediction,
- TMDPDF and RAD,

and the corresponding uncertainties. Thus $\sim$ all fits contribute, not only those which are in the 68% confidence interval (CI) around the average χ^2 of the distribution in figure 12.

The parameters obtained in the extraction for the RAD eq. (4.13) are

$$B_{\rm NP} = 1.56^{+0.13}_{-0.09}\text{GeV}, \quad c_0 = 3.69^{+0.65}_{-0.61} \cdot 10^{-2}, \quad c_1 = 5.82^{+0.64}_{-0.88} \cdot 10^{-2} \tag{4.20}$$

and the corresponding plot for the RAD is given in figure 13. The RAD agrees well with the extraction from MAP22 [8]. The additional logarithmic term helps increasing sensitivity to the region $\boldsymbol{b} \sim 1\,\text{GeV}^{-1} - 2\,\text{GeV}^{-1}$.

Due to the NP ansatz the RAD will grow uncontrolled for large $\boldsymbol{b}$. In our ansatz the behaviour at large $\boldsymbol{b}$ is linear, in other models [8] it is quadratic. It should be stressed that the region $\boldsymbol{b} \gtrsim 4\,\text{GeV}^{-1}$ cannot be determined from data at the current stage. This region effectively does not contribute to the numerical integral for the cross section because the functions initial scale value is so suppressed. Therefore the behaviour for the RAD at these large $\boldsymbol{b}$ values in all the models is a relict of the framework but physically not probed by the fits. The extraction of DY at small q_T is most sensitive to the region $b < 3\,\text{GeV}^{-1}$. The behaviour at larger $\boldsymbol{b}$ does not significantly affect the estimated cross section. Hence, with the NP ansatz we can extend the region in $\boldsymbol{b}$ space in which we can extract the TMDPDF with some confidence. For growing $\boldsymbol{b}$ we reach a point at which we are not able to make a reliable prediction. The TMDPDF and also the RAD cannot be determined for the region beyond.

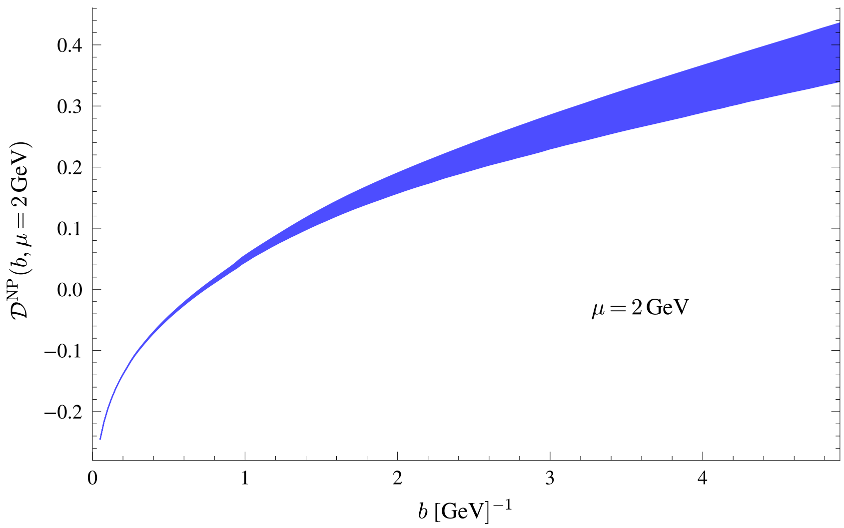


Figure 13: The RAD extracted at $\mu = 2.0\,\mathrm{GeV}$

The parameters the for the NP TMDPDF ansatz eq. (4.9) we obtain are

$$\begin{aligned} \lambda_1^u &= 0.87^{+0.10}_{-0.10}, & \lambda_2^u &= 0.91^{+0.33}_{-0.29}, \\ \lambda_1^d &= 0.99^{+0.09}_{-0.12}, & \lambda_2^d &= 6.06^{+1.36}_{-1.34}, \\ \lambda_1^{\bar{u}} &= 0.35^{+0.23}_{-0.22}, & \lambda_2^{\bar{u}} &= 46.6^{+7.9}_{-8.1}, \\ \lambda_1^{\bar{d}} &= 0.12^{+0.13}_{-0.11}, & \lambda_2^{\bar{d}} &= 1.53^{+0.54}_{-0.17}, \\ \lambda_1^r &= 1.32^{+0.23}_{-0.24}, & \lambda_2^r &= 0.46^{+0.13}_{-0.45}, \end{aligned} \tag{4.21}$$

and we also provide the obtained correlation matrix in table 3.

The extracted TMDPDFs as a function of $\boldsymbol{b}$ can be seen in figure 36 , and additional figures is given in appendix B. A plot comparing our theory prediction with the most precise data available from the ATLAS collaboration is given in figure 15. In general we find that the predictions obtained match the data very well. We observe that the data of the $W-$boson does not deliver a strong constraint on the flavour separation. This is because the experimental errors are very large and the fit is merely influenced hereby, see figure 34. A collection of all data plots is given in appendix B.

B.c Discussion

The extraction of the relevant NP parameters for the RAD and TMDPDF can be repeated with (minor) modifications, such that their impact on the final outcome can be discussed. We repeat the extraction with the following one-at-a-time modifications (to study the effect of):

- `NNPDF3.1` as an input collinear PDF (impact of the choice of the PDF)
- fixing the PDF to the central value (impact of the PDFs uncertainty onto the extraction)
- removal of the $\sim \log(b)$ term in the NP RAD (impact and justification)
- removal of the W-boson production data (impact of the flavour mixing process on a flavour dependent model)

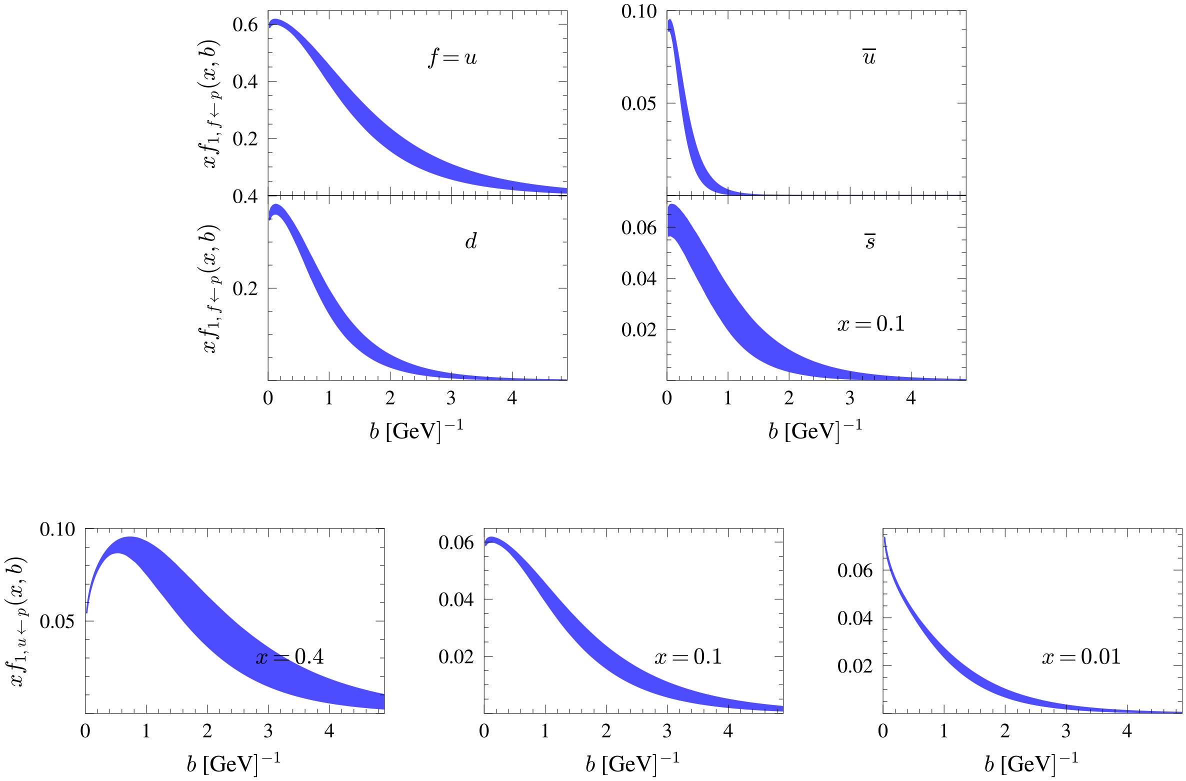


Figure 14: The TMDPDF obtained with uncertainty band obtained at the optimal scale. In the upper panels are different flavours for fixed $x = 0.1$. In the lower panels is the TMDPDF for u flavour in different x bins. Note that the optimal scale is $\boldsymbol{b}$ dependent.

	B_{NP}	c_0	c_1	λ_1^u	λ_2^u	λ_1^d	λ_2^d	$\lambda_1^{\overline{u}}$	$\lambda_2^{\overline{u}}$	$\lambda_1^{\overline{d}}$	$\lambda_2^{\overline{d}}$	λ_1^r	λ_2^r
B_{NP}	1	−0.322	−0.054	0.087	0.133	0.071	0.127	−0.053	0.226	−0.046	−0.127	0.234	0.046
c_0	−0.322	1	−0.209	−0.673	−0.446	−0.445	0.262	−0.065	−0.264	−0.111	0.313	−0.636	0.026
c_1	−0.054	−0.209	1	0.470	0.655	0.690	0.354	0.287	0.487	−0.196	−0.582	0.545	−0.045
λ_1^u	0.087	−0.673	0.470	1	0.500	0.357	−0.228	0.078	0.298	0.010	−0.608	0.542	−0.064
λ_2^u	0.133	−0.446	0.655	0.500	1	0.578	−0.023	0.131	0.403	−0.365	−0.740	0.502	−0.062
λ_1^d	0.071	−0.445	0.690	0.357	0.578	1	−0.084	0.135	0.302	−0.190	−0.322	0.459	−0.022
λ_2^d	0.127	0.262	0.354	−0.228	−0.023	−0.084	1.000	0.379	0.331	0.086	−0.073	0.112	0.018
$\lambda_1^{\overline{u}}$	−0.053	−0.065	0.287	0.078	0.131	0.135	0.379	1	−0.021	0.044	−0.133	−0.007	0.109
$\lambda_2^{\overline{u}}$	0.226	−0.264	0.487	0.298	0.403	0.302	0.331	−0.021	1	−0.083	−0.316	0.475	−0.110
$\lambda_1^{\overline{d}}$	−0.046	−0.111	−0.196	0.010	−0.365	−0.190	0.086	0.044	−0.083	1	−0.136	0.043	0.057
$\lambda_2^{\overline{d}}$	−0.127	0.313	−0.582	−0.608	−0.740	−0.322	−0.073	−0.133	−0.316	−0.136	1	−0.433	0.022
λ_1^r	0.234	−0.636	0.545	0.542	0.502	0.459	0.112	−0.007	0.475	0.043	−0.433	1	−0.003
λ_2^r	0.046	0.026	−0.045	−0.064	−0.062	−0.022	0.018	0.109	−0.110	0.057	0.022	−0.003	1

Table 3: Correlation between NP parameters eq. (4.14) from the extraction ART23. The numbers are rounded to four digits to display them here.

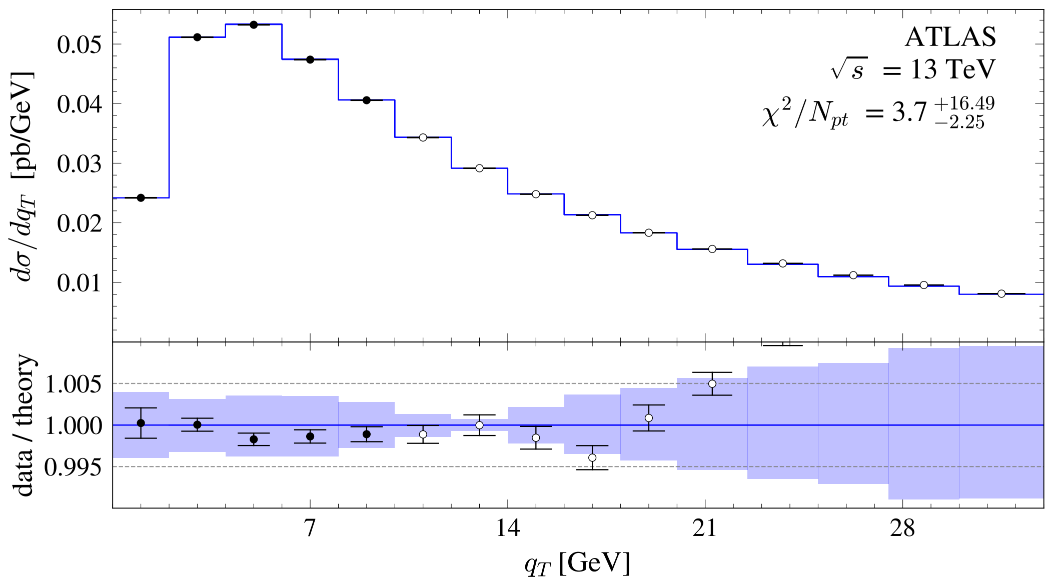


Figure 15: Theory prediction versus experiment for the ATLAS [75] experiment, here Z boson production at $\sqrt{s} = 13\,\text{TeV}$. The blue band is the 68%CI. The filled (empty) datapoints were (not) included in the fit. With increasing q_T it is (at some point) expected that the data cannot be described. The displayed χ^2 only regards to the data used in the fit.

In the fit of ART23 the extracted result with `MSHT20PDF` was cross checked versus the extraction with the `NNPDF3.1` , and the comparison between modified extractions are depicted in figures 37-39. Thus here we compare two collinear PDF sets in the same framework and for the same models[24]. The findings agree with [10] in two key observations:

- the χ^2 distribution is more spread for `NNPDF3.1` than for `MSHT20PDF` , and it has a larger average value, see figure 37;
- the parameters and the corresponding error we obtain are very different, see figure 16figure 20.

Let us additionally make the following comments. It can naturally be argued that for any collinear PDF there will be an optimal parametrization and hence also an optimal fit quality in terms of the criteria discussed above. The model (NP RAD and TMDPDFs) was developed with the `MSHT20PDF` , which can reflect a bias. This, however, starts an endless discussion effects on the extraction of mostly minor impact and (most likely) of minor physical interest as well. What we state is that we find a very good quality of data prescription with `MSHT20PDF` ; for `NNPDF3.1` we confirm already reported issues.

The RAD governs the rapidity scale evolution of TMD distributions. It is the same function for TMDPDF and TMDFFs, which underlines that it should not depend on the distribution itself. Hence, the RAD also should not depend on the input PDFs (and FFs). Formally the correlation is cured by the ζ-prescription [49] in which the TMDPDF and RAD are viewed as independent functions. In the result we see that the parameters of the TMDPDF show an acceptable correlation to those of the RAD. It cannot be avoided that the parameters all depend on each other in a numerical implementation for the fit. From here we can already see that the input PDF

[24]If we instead try to compare the results of different generations of fits or between collaborations, it is very unclear to pin down the exact reason why certain aspects work better/worse in one result than in another.

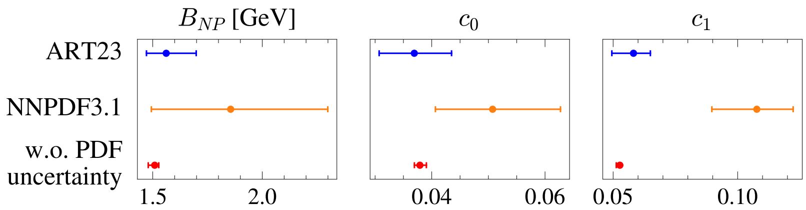


Figure 16: Extracted NP parameters for the RAD eq. (4.13) for modified extractions.

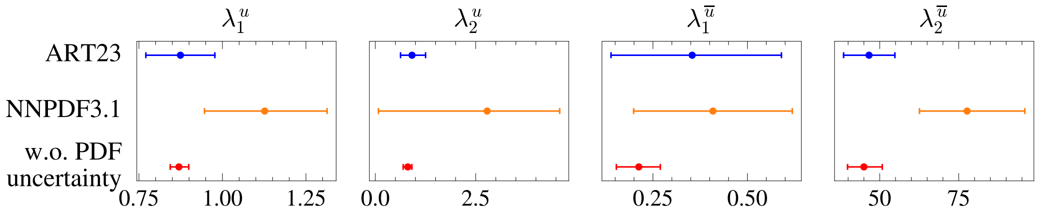


Figure 17: Extracted NP parameters for the TMDPDF eq. (4.9) for modified extractions.

eventually affects all parameters, including those from the RAD. What we observe in our fit is that the extracted RADs for the two input PDFs we compare are compatible. In the region where we want most sensitivity to NP physics($\sim 1 - 3\,\mathrm{GeV}^{-1}$), however, the two extractions do not agree within their respective uncertainties, which arises in the current implementations.

The extraction with `MSHT20PDF` was repeated with its central set for all replicas. The Monte Carlo replicas then only vary in the sampled experimental data. I.e., we repeat the fit without taking into account the PDFs uncertainties.

The theory prediction of the cross section, for most experiments, does not significantly change. At the same time, the predicted uncertainty shrinks drastically. Hence, we find that the uncertainty of our extraction mainly originates from the collinear PDF.

On the other hand, for the extracted parameters the average value can slightly vary from the main extraction, and the error bars for most parameters are reduced. The same holds for the obtained TMDPDFs: the average value can indeed shift, and the standard deviation is reduced. For most flavours, the modified result is included in the estimated error of the main extraction. The main statement that we deduce from this is that the variation from the PDF should be included in a state of the art TMD extraction to obtain a more trustworthy, reliable, and sensible uncertainty estimate.

It should be stressed that from this current implementation we can only reproduce the impact of the statistical fluctuation of one specific PDF set. The systematic bias introduced by choosing an available PDF set is not included in the error that we can relate to our extraction. It cannot quantitavely be determined but estimated by comparisons with extractions using other PDF sets. It can be seen from the earlier discussion about the choice of the PDF that a better understanding of this effect would be helpful. At this point it is unclear if the effect has some meaningful origin or merely is an artefact of the framework.

In figure 34 we show the theory prediction for W-boson production measured at Tevatron. We cross checked the impact of DY data with W-boson production on the extraction, by recomputing without taking these data into account in the fit. The central fit which leads to the initial parameters of the gradient descent method in our implementation does include all data, also the W-boson. Yet, due to variation of collinear PDFs, and, in this case perhaps more important, the variation of the W-boson data themselves, the fit should be influenced by the data strongly in the statistical fit. The fit without the W-boson data shows the same parameter distribution with practically no difference, see figure 16figure 20. What we find is that they have practically no influence. This can be seen from the data plot. Given that the data are already described inside the large experimental uncertainties, they do not provide a strong constraint on the fit.

From this discussion there arise serval points which should further be investigated. The very first point is the dependence on the collinear PDFs. The fit can be repeated for several PDF results of interest. One of these is be `HERAPDF2.0` , as mentioned already, since it provides an outstanding result for the SV19 configuration, and because its determination relies on experiments of which none are used in the TMD extraction. Apart from computational cost, these studies are also cheap in manpower.

Instead of varying between different existing input PDFs, the framework can be challenged itself. From a technical point of view it is very convenient to use the matching relation in eq. (4.5), but it also brings its problems. We can picture this as a combined model for the TMDPDF, of which the part that must describe the cross section in the collinear regime is already constrained fixated. As we discuss in the next section, an overall better result is accomplished by simultaneous fit. Hence it is a quite obvious idea to simply release the parameter for the PDF, or formulate a combined model overall. Then the results of such a combined fit of TMDPDFs and collinear PDFs can be explored.

It is obvious that, from a purely technical point of view, a combined fit complicates the problem. It is much simpler to determine one part of the parameter first, and control the remaining part at a later step. This is coupled to the technical question of how to determine a large amount of parameters, or in general an unknown function. With the in section II.A.c mentioned example of the `MSHT20PDF` extraction of PDFs, there are 52 parameters to be constrained for the collinear PDF alone. Together with the proposed ansatz in this work, 65 parameters needed to be determined simultaneously.

We estimate the creation of the full statistics came at a cost of ∼350 000 core hours. The computational time depends on the precision of the numerical computation for the cross-section. Mostly however the time for the fit to converge is determined by the amount of free parameters in the gradient-descend method. It is a purely technical, but a very important question, which methods are best suited to find the parameters for a model like in this study. The modern perturbative accuracy seems to allow require effects like flavour dependence in TMDPDFs. This suggests that we are at a precision where we are sensitive to these effects, needing a more involved but meaningful parametrization. E.g., from table 1 it can be seen that the number of parameters increased over time (the names of the extractions indicate the year of their publication). However this faces the technical problem that, typically, computational time is neither cheap nor unlimitedly accessible. Even if the first point is not an issue when a large amount of computational ressources is available, the time consumed grows exponentially with the number of parameters, reducing the effectiveness of such a method dramatically. These problems slow down the time for the fit in general. Not only does it increase the time until the statistics have been computed, it also slows down and complicates the process of finding of a good parametrization in general. For the simultaneous phenomenology of TMDPDFs and collinear PDFs togegher (or similar studies) it will be an essential issue to find a good way of solving this minimization problem efficiently.

There are algorithms that perform better than the gradient descent method for this large amount of parameters. One of the most discussed alternative is the solving via neural networks. To advance towards this direction can be expected to be very difficult mostly in aspects of technical implementation. Yet it is necessary to put the available technology to the test, in order to push the boundaries on what can be understood from the available data in our field.

The flavour dependence should ideally be improved by the W-boson data. Certainly the large errors are the main effect of why the data provides such a poor constraint on the fit. Nevertheless more can be done with the available data. It can be verified that by varying initial conditions for the fit the final result does not change.

Another constraint to the flavour dependence is SIDIS, as the hadronisation strongly depends on the flavour of the quark to form a hadron.

C Extraction of TMDFFs from data

To extract TMD fragmentation functions (FF), we study the process of SIDIS. Alternatively, they could be studied in the process of semi-inclusive annihilation (SIA), which purely is a fragmentation process from quarks into hadrons. It would be a demonstration of the success of the framework and the reliability of the extracted functions to determine (TMD)PDFs purely from DY, and (TMD)FFs solely from SIA and then verify that SIDIS, which requires both ingredients, can be described in the TMD regime from these input – without additional modification.

Unfortunately, the data in SIA fragmenting into two unpolarized related hadrons are not available and unpolarized TMDFFs related to TMD factorisation cannot be extracted for the time being[76]. In the past it has been a standard procedure to fit TMDPDFs from DY, and determine TMDFFs in a joint ("global") fit [12, 55]. This is how there is no independent fit of TMDFFs so far. The combined fit of SIDIS allows for the determines the TMDFFs, and it also provides a constraint on the TMDPDFs.

In what follows we show a combined (DY and SIDIS) fit, in which we simultaneously extract TMDPDFs and TMDFFs from DY and SIDIS data. The applied framework coincides with what has been discussed in section A; the ansatz for the TMDPDF is the model for ART23, with an additional constraint which we discuss in section C.b.

C.a Data review

We discuss here the SIDIS data used in a combined fit. The DY data are discussed earlier and can be found in [9]. For the fit we consider data from the HERMES [18] and COMPASS [20] collaborations. As in [12] we use the datasets from HERMES which are binned in the variables x, z and p_T, which have finer p_T binning. The data are split into bins for the ranges $0.023 < x < 0.6$ in six bins, $0.1 < z < 1.1$ in 8 bins and $0\,\text{GeV} < p_T < 1.2\,\text{GeV}$ in 7 bins. The data from COMPASS are given in four fold differential binning in the variables Q, x, z and p_T^2. They are provided in the kinematic ranges of $0.003 < x < 0.4$ in 8 bins, $1.0\,\text{GeV}^2 < Q^2 < 81\,\text{GeV}^2$ in 5 bins, $0.2 < z < 0.8$ in 4 bins and $0.02\,\text{GeV}^2 < p_T^2 < 3.0, \text{GeV}^2$ in 30 bins. The COMPASS data are given with subtraction of the vector-boson channel. Thus for COMPASS and HERMES we use this kind of data, for consistency. The multiplicities in the HERMES data distinguish between charged pions and kaons, while the COMPASS collaboration provides the data as the sum of pions, kaons and (anti-)protons, which they summarize as $h^{\pm}$. Thus, the theory computation has to take into account all contributions, i.e. sum the cross sections for the individual hadron productions. Because the fraction of (anti-)protons in the COMPASS measurement is negligible, we do not take them into account and approximate $\pi^{\pm} + K^{\pm} = h^{\pm}$.

The data are restricted to fulfil the kinematic constraints to be considered in the fit. They are

$$\langle Q \rangle > 2, \quad 0.2 < z < 0.8, \tag{4.22}$$

and additionally the same constraint as for the DY data eq. (4.19) [25]. The first condition should remove those points where it becomes doubtful that perturbative physics apply. For resolution scales $Q \sim M$ comparable to the mass of the proton, the perturbative factorisation approach certainly will break down. Nevertheless we have to retain some data for the analysis and we thus allow for a rather low cut on $\langle Q \rangle$. The cut on z only affects HERMES data and restricts them to be in the same range as the COMPASS data.

We also disregard datapoints in which the kinematics in the extreme bin region lead to a violation of the TMD constraints, even though the average values fulfil it. This is done by defining an approximate $\zeta_{\max}$ with the

[25] Note that for SIDIS q_T is defined in eq. (2.58).

Experiment	channel	$N_{\rm pt}$
HERMES	$p \to \pi^+$	24
	$p \to \pi^-$	24
	$p \to K^+$	24
	$p \to K^-$	24
	$d \to \pi^+$	24
	$d \to \pi^-$	24
	$d \to K^+$	24
	$d \to K^-$	24
COMPASS	$d \to h^+$	195
	$d \to h^-$	195
Total	SIDIS	**582**

Table 4: SIDIS measurements taken into account in the fit. They are listed in the separate channels.

definition eq. (2.52) and the maximum x value, as an approximation. Then $\zeta_{\max} < 1$ is required, and the maximal $q_{T,\max}$ is approximated with $\zeta_{\max}$, the maximal value x and minimal value z in the respective bin. The condition is then that the maximal value has to satisfy

$$\frac{\langle q_{T,\max}\rangle}{\langle Q\rangle} < 0.5, \tag{4.23}$$

or else the entire datapoint is disregarded.

We list the data in table 4 with reported number of points included in the fit from the different analyses. Both experiments provide their data as multiplicities, i.e. ratios of the SIDIS cross section normalized to the corresponding DIS cross section. The DIS cross section is computed with the same PDF as the SIDIS cross section (i.e. MSHT20 [14]). To compute the DIS cross section, we compared the results of three codes, `apfel` [77], `apfel++` mcapfel++ and the code, used for the extraction of MSHT20 and MMHT14, which is available on `xfitter` [78]. The DIS cross sections computed with the different codes are in agreement. We choose to use the code from the MSHT collaboration to predict the DIS cross section, for consistency.

C.b Modifications on the TMDPDFs

Naively the extraction for the TMDFFs could be done using the results from ART23 [9]. That implementation would only determine the parameters of the TMDFF eq. (4.25). This would also be a desirable approach mainly for technical reasons. I.e., if the TMDPDF and the RAD are fixed, the amount of parameters for the fit is much smaller compared to the sum of all NP parameters from all TMD models involved. In our case we would only need to determine four parameters from the fit, instead of 17. This would be a drastic difference in terms of runtime and in the complexity of the parameter space, for a single fit configuration. By a configuration here we speak of the including exact model, included datapoints, collinear input functions (the proton PDF and FFs for two pions and kaons); a different configuration would use a different model, other data or other input functions.

This could not be set into practice for two reasons.

1. The TMDPDF obtained by DY fits section B as demonstrated describes the data very well. Even though in ART23 the probed resolution ranges from $4\,\mathrm{GeV}^2$ to $10^3\,\mathrm{GeV}^2$, it appears not to be sufficient to describe data ranging down to $Q^2 = 3\,\mathrm{GeV}^2$[26]. The cross section to SIDIS processes tests larger $\boldsymbol{b}$ regions of the distributions, which are not as constrained by fits to DY data only. This is why typically the TMDFFs are

[26]Due to the binning the lowest Q for the COMPASS data appears in the bin $3\,\mathrm{GeV}^2 < Q^2 < 7\,\mathrm{GeV}^2$.

obtained in a combined fit [8, 12]; also in this work we find that a satisfactory result in terms of fit quality (χ^2) can only be achieved by a combined fit of DY and SIDIS data, and releasing the TMDPDF and RAD parameters. This decreases the quality of the fit to DY, but dramatically improves the quality for the SIDIS description.

2. We observe a flaw in the flavour dependent ansatz in the TMDPDF of ART23. The flavour separation between q and $\bar{q}$ in the region of small x presents a problem, even though the DY description with this model works very well. We fix this issue by imposing that the flavour separation only holds for large x, of which the behaviour should be captured inside λ_2^f in eq. (4.9). For small x we remove the flavour separation in our model. We implement this without changing the parametrization, but demand

$$\lambda_1^u = \lambda_1^{\bar{u}} \quad \lambda_1^d = \lambda_1^{\bar{d}} \tag{4.24}$$

for the combined fit to DY and SIDIS data. The physical problem can be explained and motivated by the idea that PDFs at small momentum fractions x are dominated by sea quarks. The sea quarks are emerge as colour dipoles from gluons, and hence are created in pairs, and thus have symmetric properties with respect to their (anti)quark pair partner.

C.c NP model and fit setup summary

As mentioned in the previous section, a satisfactory extraction with a justified amount of parameters and simultaneously a good χ^2 for the fit quality could not be constructed with the results from the TMDPDF result from section B. With changes mentioned above, we are able to use a very simple parametrization for the NP part of the TMDFFs,

$$D^{\mathrm{NP}}_{f\to h}(z, \boldsymbol{b}) = \frac{1}{\cosh\left(\eta_0^h \boldsymbol{b}/z\right)}\left(1 + \eta_1^h \boldsymbol{b}^2/z^2\right). \tag{4.25}$$

The parameters $\eta_{0,1}^h$ depend on the hadron, which can be a pion or a kaon in the current work. This means there are four parameters in the following extraction, two for the pion FF and two for the kaon FF.

Further separation as flavour has been studied. The separation into u, $\bar{d}/\bar{s}$ and remaining flavours r for pions/kaons, respectively has been checked with several fits to the central data. It can be observed that the parameters do not converge to the same point. However, because the fit already produces a good χ^2 with the flavour-independent ansatz eq. (4.25), and for the technical reason to keep the number of additional free parameters as large as necessary but as small as possible, we decide not to enforce flavour separation for the TMDFF ansatz.

For the extractions we use the parametrizations for the NP functions RAD as in eq. (4.13), the TMDPDFs as in eq. (4.9) with the modification in section C.b, and the TMDPFFs as in eq. (4.25). We use 15 parameters in total in the NP model to describe DY and SIDIS in the TMD regime. The RAD parameter B_{NP} is fixed to $1.5\,\mathrm{GeV}^{-1}$ because we see that it can be compensated by the remaining parameters inside the NP ansatz. It has been demonstrated already previously by another group that this is possible [11, 8]. Our findings confirm their result.

For the TMDFF we obtain the η_0^h by the central fit, but fix them for the statistical replica fits due to of strong correlation with the parameter η_1^h, for both pions and kaons respectively. In this setup we then have 12 parameters to obtain by simultaneous variation of the experimental data, the PDFs and FFs replicas. The collinear PDF used in the current work is `MSHT20NNLO` [14]. This set was used as well in ART23. The collinear FFs used are DSS22Pion for the pion FF [15] (extracted at approximate NNLO), and the DSS17 Kaon result [16] for the kaon FF (at NLO precision).

C.d Results

With the modified result of ART23 and the ansatz discussed in the previous section, we find that we can describe the SIDIS data very well. With this setup, we obtain a $\chi^2/N_{\text{pt}} = 0.93$ for the SIDIS data[27]. At the same time, the DY data still produce a nice result with $\chi^2/N_{\text{pt}} = 1.21$[28]. Here we should clarify that, the combined fit with SIDIS is expected to worsen the result of the test function for DY, simply because we add constraints. We will discuss this in more detail later on, but stress that this number should not blindly be compared to the ART23 result. All χ^2 values are listed in table 5 bin-wise for DY and channel-wise for SIDIS for the individual experiments used in the fit.

With the restrictions as discussed we obtain the parameters

$$B_{\text{NP}} = 1.5 \quad \text{GeV} \quad \text{fixed}, \quad c_0 = 7.40^{+0.76}_{-0.60} \cdot 10^{-2}, \quad c_1 = 2.80^{+0.67}_{-0.59} \cdot 10^{-2} \tag{4.26}$$

for the NP part of the RAD eq. (4.13). The parameters for the TMDPDF in the combined DY and SIDIS extraction are,

$$\begin{aligned}
\lambda_1^u &= 0.44^{+0.08}_{-0.08}, & \lambda_2^u &= 0.24^{+0.15}_{-0.16}, \\
\lambda_1^d &= 0.42^{+0.08}_{-0.08}, & \lambda_2^d &= 0.64^{+0.36}_{-0.55}, \\
& & \lambda_2^{\bar{u}} &= 21.8^{+7.3}_{-7.4}, \\
& & \lambda_2^{\bar{d}} &= 0.12^{+0.17}_{-0.12}, \\
\lambda_1^r &= 0.51^{+0.12}_{-0.17}, & \lambda_2^r &= 0.0008^{+3.9789}_{-0.0006},
\end{aligned} \tag{4.27}$$

and we do not display the duplicated parameter values. For λ_2^r we print the median instead of the mean value, because the mean value is actually outside of the 68%CI with regular definitions. The parameter tends to very small values for most replicas, yet has some outliers for much larger values, such that the parameter interval with 68% of replica inside spans over four orders of magnitude. Clearly, there is a problem with this parameter. One can fix the parameter value, but the fit quality decreases as a consequence. Alternatively the parameter can be restrained to a certain domain and all outliers can be thus removed; yet we do not want to bias the result without further investigation which explain the behaviour or motivate such cuts. This is why for the moment we keep the parameter free, but acknowledge the issue on its behaviour. The parameters for the FFs for the pion and kaon in eq. (4.25) are

$$\begin{aligned}
\eta_0^\pi &= 0.47, & \eta_1^\pi &= 0.16^{+0.01}_{-0.01}, \\
\eta_0^K &= 1.13, & \eta_1^K &= 2.33^{+0.26}_{-0.24}.
\end{aligned} \tag{4.28}$$

Because we observe a strong correlation between $\eta^h_{0,1}$ for both hadrons, we currently fix one of the parameters to its central fit value. This is why no uncertainty is provided on these parameters.

C.e Discussion

In this extraction the formalism of TMD factorization has been put to the test. We observe that the constraint from DY data is not enough to determine the TMDPDFs such that also the SIDIS data can be described well. In

[27] This value regards the average of all replicas. The central fit yields a lower χ^2 value, but we regard it as less reliable.
[28] See 27

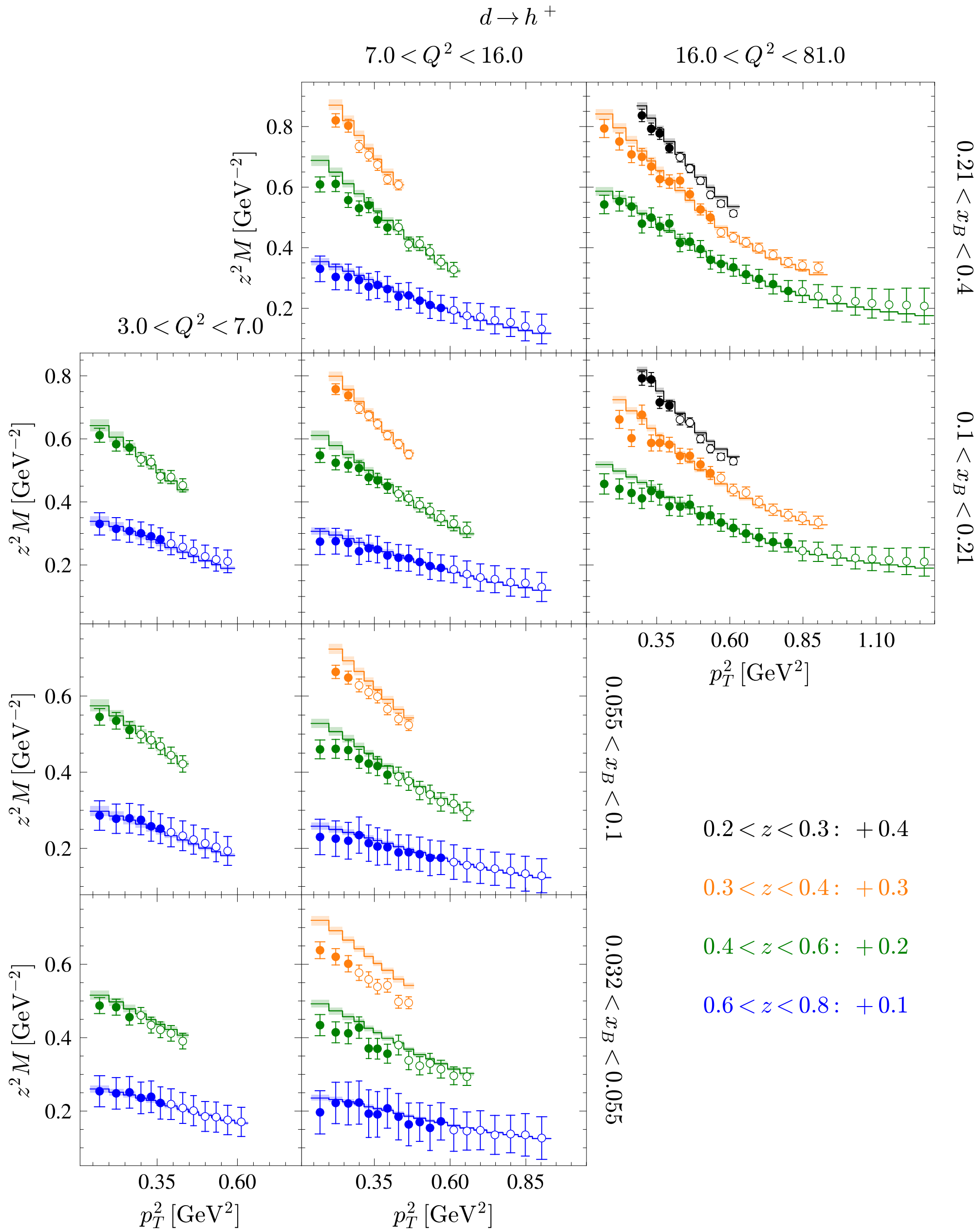


Figure 18: Theory prediction for charged (+) hadron production in SIDIS for the COMPASS experiment. The data are given as multiplicities in bins of the variables Q, x, z and the p_T profile is plotted. In the experiment, pions, kaons and (anti)protons are observed. For the theory prediction protons are neglected. To show multiple kinematic regions of z we use an offset as indicated in the figure.

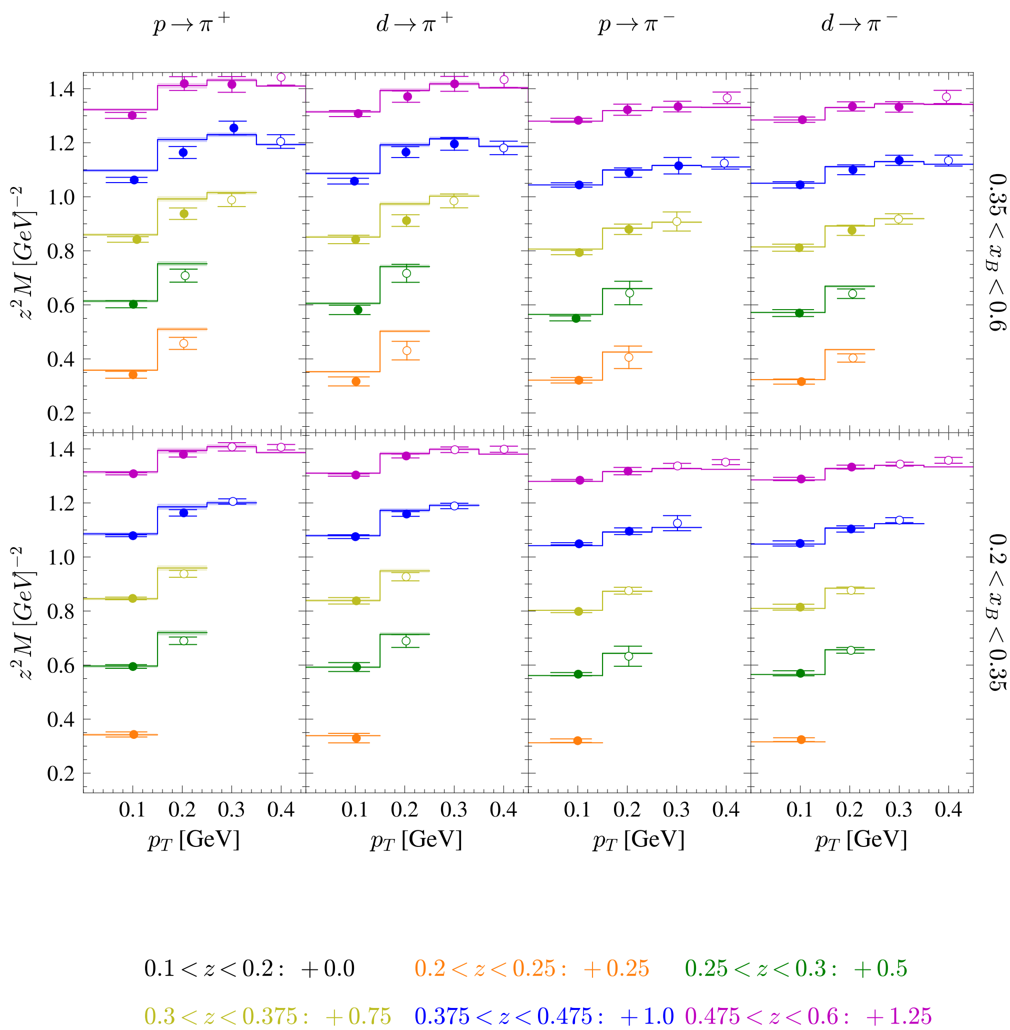


Figure 19: Theory prediction for charged pion production in SIDIS for the HERMES experiment. The data are given as multiplicities in bins of the variables x, z and the p_T profile is plotted. To show multiple kinematic regions of z we use an offset as indicated in the figure.

Experiment	N_{pt}	χ^2_D/N_{pt}	$\chi^2_\lambda/N_{\text{pt}}$	χ^2/N_{pt}	χ^2/N_{pt}	$\langle d/\sigma \rangle$
CDF (run1)	33	0.66	0.15	0.81	$0.84^{+0.06}_{-0.07}$	8.7%
CDF (run2)	45	2.71	0.01	2.72	$2.81^{+0.44}_{-0.43}$	3.5%
D0 (run1)	16	0.87	0.00	0.87	$0.87^{+0.08}_{-0.08}$	0.2%
D0 (run2)	9	3.19	0.00	3.19	$3.11^{+0.40}_{-0.39}$	0.0%
D0 (run2: μ)	4	2.87	0.61	3.48	$3.13^{+0.39}_{-0.35}$	1.9%
ATLAS 7 TeV $0.0 < y < 1.0$	5	4.01	0.01	4.02	$4.85^{+2.47}_{-2.28}$	-0.1%
ATLAS 7 TeV $2.0 < y < 2.0$	5	4.84	0.02	4.85	$5.92^{+2.43}_{-2.20}$	-0.1%
ATLAS 7 TeV $2.0 < y < 2.4$	8	0.65	0.00	0.65	$0.98^{+0.38}_{-0.43}$	0.0%
ATLAS 8 TeV $0.0 < y < 0.4$	5	4.03	0.68	4.71	$5.63^{+1.96}_{-1.87}$	3.2%
ATLAS 8 TeV $0.4 < y < 0.8$	5	1.70	0.24	1.94	$2.63^{+0.89}_{-0.69}$	3.2%
ATLAS 8 TeV $0.8 < y < 1.2$	5	0.37	0.22	0.59	$1.06^{+0.56}_{-0.66}$	3.0%
ATLAS 8 TeV $1.2 < y < 1.6$	5	1.65	0.47	2.12	$3.10^{+1.22}_{-1.12}$	3.2%
ATLAS 8 TeV $1.6 < y < 2.0$	5	0.12	0.29	0.41	$1.22^{+0.82}_{-0.68}$	3.5%
ATLAS 8 TeV $2.0 < y < 2.4$	5	1.00	0.68	1.68	$2.12^{+0.80}_{-0.78}$	3.9%
ATLAS 8 TeV $46 < Q/\text{GeV} < 66$	5	1.44	0.10	1.54	$1.55^{+0.49}_{-0.42}$	1.6%
ATLAS 8 TeV $116 < Q/\text{GeV} < 150$	9	0.92	0.06	0.98	$1.10^{+0.18}_{-0.19}$	2.3%
ATLAS 13 TeV (norm.)	5	0.19	0.00	0.19	$4.09^{+3.08}_{-3.32}$	0.0%
CMS 7 TeV	8	1.23	0.00	1.23	$1.24^{+0.03}_{-0.03}$	0.0%
CMS 8 TeV	8	0.78	0.00	0.78	$0.79^{+0.02}_{-0.02}$	0.0%
CMS 13 TeV $0.0 < y < 0.4$	10	1.46	0.18	1.63	$1.86^{+0.34}_{-0.33}$	3.3%
CMS 13 TeV $0.4 < y < 0.8$	12	0.70	0.17	0.87	$0.99^{+0.16}_{-0.12}$	3.6%
CMS 13 TeV $0.8 < y < 1.2$	13	0.39	0.13	0.52	$0.61^{+0.16}_{-0.10}$	3.3%
CMS 13 TeV $1.2 < y < 1.6$	14	0.36	0.15	0.51	$0.54^{+0.10}_{-0.08}$	3.7%
CMS 13 TeV $1.6 < y < 2.4$	15	0.57	0.10	0.67	$0.69^{+0.10}_{-0.09}$	3.0%
CMS 13 TeV $50 < Q/\text{GeV} < 76$	6	7.67	12.33	20.00	$19.97^{+2.94}_{-2.65}$	10.3%
CMS 13 TeV $106 < Q/\text{GeV} < 170$	10	0.78	0.35	1.13	$1.45^{+0.37}_{-0.33}$	2.2%
CMS 13 TeV $170 < Q/\text{GeV} < 150$	11	0.45	0.05	0.50	$0.58^{+0.12}_{-0.09}$	0.9%
CMS 13 TeV $350 < Q/\text{GeV} < 1000$	12	0.67	0.03	0.70	$0.82^{+0.23}_{-0.24}$	-0.8%
LHCb 7 TeV	10	1.18	0.61	1.79	$1.87^{+0.35}_{-0.31}$	5.3%
LHCb 8 TeV	9	0.94	0.74	1.68	$1.89^{+0.64}_{-0.60}$	4.2%
LHCb 13 TeV	49	1.05	0.11	1.17	$1.38^{+0.18}_{-0.14}$	4.7%
PHENIX	3	0.27	0.00	0.27	$0.32^{+0.04}_{-0.05}$	0.6%
STAR	11	1.15	0.17	1.32	$1.42^{+0.10}_{-0.10}$	11.8%

Table 5: Part 1. The resulting list of χ^2/N_{pt} for all DY and SIDIS datasets. The number of datapoints corresponds to those which are used in the fits. For the central replica we show the decomposition into χ^2_D and chi^2_λ representing the contributions to the χ^2 from disagreement with shape and normalisation of the data, respectively. For the replica average we also provide the 68% CI.

Experiment	$N_{\rm pt}$	$\chi^2_D/N_{\rm pt}$	$\chi^2_\lambda/N_{\rm pt}$	$\chi^2/N_{\rm pt}$	$\chi^2/N_{\rm pt}$	$\langle d/\sigma\rangle$
E228 200 GeV	43	0.35	0.03	0.38	$0.49^{+0.10}_{-0.08}$	30.3%
E228 300 GeV	53	0.48	0.04	0.51	$0.66^{+0.12}_{-0.08}$	33.9%
E228 400 GeV	79	0.67	0.03	0.69	$1.10^{+0.48}_{-0.33}$	35.5%
E772	35	1.28	0.11	1.39	$1.52^{+0.17}_{-0.18}$	19.4%
E605	53	0.29	0.13	0.42	$0.63^{+0.21}_{-0.21}$	39.5%
D0 (W-Boson)	7	3.07	0.00	3.07	$3.14^{+0.17}_{-0.19}$	0.0%
CDF (W-Boson)	6	0.39	0.00	0.39	$0.40^{+0.02}_{-0.02}$	0.0%
DY total	**627**				$1.21^{+0.10}_{-0.09}$	
Hermes $p \to \pi^+$	24	2.84	0.08	2.92	$2.75^{+0.53}_{-0.55}$	2.8%
Hermes $p \to \pi^-$	24	2.18	0.38	2.56	$2.33^{+0.35}_{-0.31}$	6.3%
Hermes $d \to \pi^+$	24	1.51	0.14	1.65	$2.25^{+0.26}_{-0.23}$	4.1%
Hermes $d \to \pi^-$	24	1.58	0.25	1.83	$0.99^{+0.22}_{-0.24}$	5.5%
Hermes pions total	**96**			**2.08**		
Hermes $p \to K^+$	24	0.80	0.00	0.81	$0.98^{+0.17}_{-0.20}$	0.7%
Hermes $p \to K^-$	24	0.73	0.01	0.74	$1.07^{+0.35}_{-0.36}$	1.2%
Hermes $d \to K^+$	24	0.94	0.02	0.96	$1.25^{+0.34}_{-0.43}$	1.6%
Hermes $d \to K^-$	24	1.53	0.08	1.61	$3.28^{+1.43}_{-1.39}$	3.1%
Hermes kaons total	**96**			**1.65**		
Compass $d \to h^+$	195	0.53	0.00	0.53	$0.66^{+0.11}_{-0.10}$	1.7%
Compass $d \to h^-$	195	0.39	0.00	0.39	$0.45^{+0.06}_{-0.06}$	-1.2%
Compass total	**390**			**0.55**		
SIDIS total	**582**				$0.93^{+0.10}_{-0.10}$	

Table 6: Part two of table 5.

Experiment	$\sqrt{s}$ [GeV]	Q [GeV] / x_F	N_{pt}	corr. err.	ref.
E537 ($Q-diff$)	15.3	$4 < Q < 9$ in 10 bins	147	corr. err.	[79]
E537 (x_F-diff)	15.3	$-0.1 < x_F < 1.0$ in 11 bins	165	corr. err.	[79]
E615 ($Q-diff$)	21.7	$4.05 < Q < 13.05$ in 10 bins	155	corr. err.	[80]
E615 (x_F-diff)	21.7	$0.0 < x_F < 1.0$ in 10 bins	159	corr. err.	[80]

Table 7: Overview of the available data. The correlated error is the estimated shifted uncertainty from the extraction with the central replica.

the combined fit of DY and SIDIS data we have a total of 1209 datapoints and a very good fit quality.

About two third of the SIDIS data we used in the fit are from the COMPASS collaboration. We observe that the description of these data is accomplished also with the ART23 results without modification. The greatest challenge is given in reproducing the HERMES data, especially the pion production channel. Their description has been greatly improved by the discussed update on the proton TMDPDF. However, it still is the bottleneck of the SIDIS description in our χ^2 decomposition table 5. Our model follows the data from the COMPASS collaboration qualitatively very well, and reproduces it very well within the experimental uncertainties. Because the COMPASS results contribute about twice as many datapoints, the total value for SIDIS is still very good.

The DY data are described very well with a $\chi^2 = 1.21^{27}(\mathrm{DY})$ in the combined fit. Preliminary results of a fit with this TMDPDF model to DY data only shows a $\chi^2 = 1.13^{27}(\mathrm{DY})$. This means that the quality of the DY fit decreases in direct comparison to ART23 because of two reasons. The first reason is that the model here has the same functional from but less (free) parameters. In ART23 there are 13 parameters, here there are 10 free parameters plus the B_{NP} a fixed parameter. As the second reason we list that the SIDIS data modifies the TMDPDF for a better combined description at the cost of a (slightly) worse description of DY.

It will be interesting to see the effect of modifications of exterior constraints, namely the choice of a different collinear PDF (tested in ART23), the choice of a different collinear FF, minor modifications of the NP functions, and the variation of the cuts in data. Once it is possible to manage a combined fit on PDFs and TMDPDFs, it will also be interesting to try to implement this for (TMD)FFs.

A most interesting advance would be to extract the NP functions at higher precision in the TMD factorization. Currently, the factorization is implemented at leading power. Thus the relations are approximates for $q_T \ll Q$; we expect corrections to our formalism $\sim q^2/Q^2$. Also, NLP power corrections should provide the access to gluon TMD distributions, which is not constrained at the leading power in TMD factorization.

D Extraction of TMDPDFs for pions

The extraction of pion TMDPDFs follows the framework of described in section A. Experiments provide us with cross sections of DY processes induced by pion-nucleus scattering [80, 79]. The nucleus is described as sum of protons and neutrons, both of which then can be described by the proton[29] TMDPDFs in a previous study[30]. Thus the results for the proton TMDPDFs can be applied and the unpolarized pion TMDPDF is extracted, which enters once in the factorised form of the cross section. This kind of study has been successfully performed already [81, 82]. Hence, with the results from ART23 section B.b we expect to obtain pion TMDPDFs on the basis of the constrained proton TMDPDF and RAD. In contrast to the fit to SIDIS data, the proton TMDPDF and the evolution (RAD) are fixed. The data is not considered to contribute in a combined fit.

In our framework, we observe complications with the direct usage of ART23. Even for a very involved model for the pions TMDPDF, a fit result of $\chi^2 \approx 1.7$ cannot be improved. We achieve a good fit result by modifying the proton TMDPDF, discussed in section C.b.

The challenges of determining the pion TMDPDF are given by the scarce amount of data available. This may be overcome by an upcoming publication of a measurement of pion-induced DY process by the COMPASS collaboration [83]. In this work we are able to describe both datasets used in other pion TMDPDF fits[81, 82]. Large systematic normalization difference between data and theory are reported in [81], and we confirm this observation.

D.a Formalism

Here we briefly summarize the necessary parameters from theory for this extraction.

[29] The neutron PDFs in general are related to those from the proton by isospin symmetry $u \leftrightarrow d$.

[30] Also in that extraction there are heavier nuclei involved, modeled in the same way

The data is given in the variable

$$x_F = \frac{2p_L}{\sqrt{s}}, \tag{4.29}$$

the Feynman-x, centre of mass energy s, with the longitudinal (transverse) component of the virtual photon p_L (p_T) in E615 [80] (the definition in E537 [79] somewhat alters). From [81] we adopt [31] relation to the Bjorken x [81]

$$x_{1,2} = \frac{\pm x_F + \sqrt{x_F + 4\tau}}{2}, \qquad \tau = x_1 x_2 = \frac{Q^2 + q_T^2}{s}, \tag{4.30}$$

and the representation of the differential DY cross section, which can be expressed in x_F instead of the rapidity y. The three fold differential cross section is expressed as a convolution of the TMDPDFs and the Bessel function of the first kind,

$$\begin{aligned}\frac{\mathrm{d}^3\sigma}{\mathrm{d}Q^2\mathrm{d}x_F\mathrm{d}q_T^2} = \frac{2\pi\alpha^2(Q)}{9Q^2 s\sqrt{x_F^2+4\tau}} &\sum_f Q_f^2 \left|C_V\left(Q^2,\mu_F^2\right)\right|^2 \\ &\times \int_0^\infty \mathrm{d}\boldsymbol{b}\boldsymbol{b} J_0(\boldsymbol{b}q_T) F_{1,f\leftarrow N_1}(x_1,\boldsymbol{b};\mu_F,\zeta_1) F_{1,f\leftarrow N_2}(x_2,\boldsymbol{b};\mu_F,\zeta_2).\end{aligned} \tag{4.31}$$

The symbols are identical to those discussed in the earlier sections.

We use the proton TMDPDF form of section A. The parameter values are the results of a central fit to the DY fit only. The reason for this is that at the time of this extraction the data for the combined fit was not available. The parameters for the proton in this extraction are thus fixed to the values

$$B_{\mathrm{NP}} = 1.5\mathrm{GeV}, \quad c_0 = 5.61\cdot 10^{-2}, \quad c_1 = 3.86\cdot 10^{-2} \tag{4.32}$$

for the RAD and

$$\begin{aligned} \lambda_1^u &= 0.57, \quad \lambda_2^u = 0.054, \\ \lambda_1^d &= 0.57, \quad \lambda_2^d = 6.64, \\ &\qquad\quad\ \ \lambda_2^{\bar{u}} = 20.1, \\ &\qquad\quad\ \ \lambda_2^{\bar{d}} = 0.537, \\ \lambda_1^r &= 1.07, \quad \lambda_2^r = 2.39. \end{aligned} \tag{4.33}$$

The perturbative orders are exactly the same as used in the proton TMDPDF extraction clarified in section B.a. In this work we use the exact solution of the ζ-prescription [49].For small-$\boldsymbol{b}$ we match the TMDPDF to the collinear pion PDF `JAM21PionPDFnlo` [84] from the JAM collaboration.

We use a flavour dependent model for the proton TMDPDF, and in combination with this, we expect a flavour separating ansatz to work for the pion TMDPDF, as well. Hence we formulate the flavour dependent ansatz[32] for

[31] It is the same framework (`artemide`). Thus these relations are already implemented and we use them implicitely.

[32] The model could be modified with a term responsible for the large x behaviour. In our extraction we are not sensible to this and the several tries all lead to this term to have coefficient 0.

the NP part of the TMDPDF

$$F^{\mathrm{NP}}_{f\leftarrow\pi}(x,\boldsymbol{b}) = \frac{1}{\cosh\left[\lambda^f_\pi(1-x)\boldsymbol{b}\right]}, \tag{4.34}$$

where we distinguish between the flavours $f = u/\overline{u}, d\overline{d}, r(\text{sea})$. The subscript π should clearly distinguish the parameters here from the ones for the proton. In total we have thus 3 NP parameters to be determined by fit. Already the parameter for the $d\overline{d}$ quark has a much greater uncertainty than the parameter for $u\overline{u}$ flavour. Even though we obtain a distribution with "good[33]" properties for the third parameter, the cross section itself is not affected by the parameter. One can even consider $F^{\mathrm{NP}}_{r\leftarrow\pi} = 0$ and both other parameters can compensate this by minimal adjustment. Another possibility is to use the same parameter value for r as for $d\overline{d}$ flavour.

The parameter λ^r_π does not improve the outcome for the χ^2 function significantly. Yet, flavour separation for sea quarks is kept in the presented work for two reasons.

1. It seems intuitive: the other distinguished flavours are valence quarks of the pion, or their antiquark. The remaining flavours are effects of QCD vacuum, and the same strategy is proved successful for the proton.
2. A significant separation of the parameters is the preferred outcome if all parameters are unconstrained. This is reflected by the values shown later.

D.b Data review and selection

The data sets available for pion induced DY processes are listed in table 7. The data was already stored in the `artemide-DataProcessor` [81] library and could be used directly. The data sets E615 [80] and E537 [79] both are taken at the Fermilab, and have a very similar setup. In the experiment, a $\pi^- - \overline{p}$ beam collides with a tungsten target. In [80] it is explained that the beam hitting the target is generated by a first beam of protons colliding with a Beryllium target.

The experimental cross sections for both E537 and E615 are available in differential form to either the Feynman-x variable x_F or the invariant mass of the produced lepton pair Q. We adopt the argumentation made in [81], stating that the variation in Q (ideally) is fixed by the earlier extraction already. As we should not include the same data twice in the fit, we chose to compare to the cross section differential in x_F provided by E615. We also adopt the reasoning [81] that the large x_F reasoning is not described with the TMD factorisation presented, as it requires resummation.

We find that after a fit only to the data of E615, we describe the data of E537. This means that an inclusion of the mentioned data will have a significant influence on our extracted TMDPDF. For this dataset the normalization issue is a priori minor. Yet we discard this data, as they show unstable behaviour. There are very large uncertainties, and the datapoints clearly show fluctuations. The data from E615 are more reliable, thus, we only fit to them.

The data selection for the fit requires the datapoint to have a ratio of $\delta = q_T/Q < 0.25$. This choice respects a balance between including available data and satisfying the requirement for TMD factorisation, i.e. $\delta \ll 1$. This also means that the proton TMDPDF and RAD were extracted with the same cut parameter; we now extract the pion TMDPDFs for the same value of δ.

The result of the selection is that we fit 72 datapoints from the experiment E615, in the kinematic range of $0.0 < x_F < 0.8$.

[33]In some extractions, some parameters do not show a Gaussian curve for their distribution. They possibly are largely spread, or even have several accumulation points, all of which would be signs of poorly constrained parameters. Here, we do not see any of these criteria.

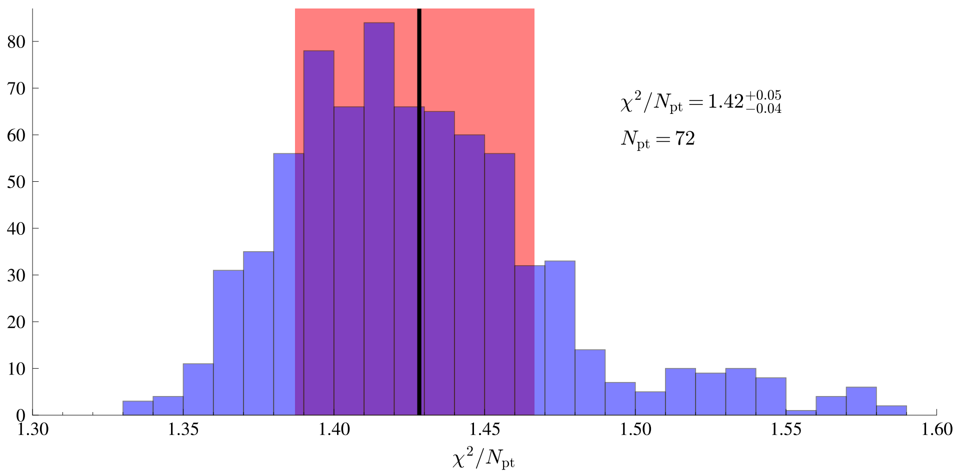


Figure 20: The distribution of the χ^2 over all generated replica fits, without cuts or restrictions. The individual fits each are to a different replica of experimental data, and have a random collinear pion TMDPDF as input. The black line indicates the average value, and the red band the 68%CI.

D.c Results

The results for the NP parameters of eq. (4.34) are

$$\lambda_\pi^{u/\overline{u}} = 0.601^{+0.006}_{-0.004}, \qquad \lambda_\pi^{d/\overline{d}} = 13.14^{+1.88}_{-1.86}, \qquad \lambda_\pi^{r} = 31.73^{+3.26}_{-2.06}. \tag{4.35}$$

The parameter for $f = u/\overline{u}$ is constrained very well, and it is the parameter that matters most in this model. Computation of the cross section with various settings clearly indicates that the $u/\overline{u}$-flavour is the one contributing to the cross section by the largest amount. From the fit also both other parameters for the d type quarks and the remaining quarks appear reasonably constrained and separated from this fit. The parameters obtained for instead constraining $\lambda^{\overline{d}/d} = \lambda^r$ are indeed the same as for the case where the remaining quarks have a free parameter, as in this case

$$\lambda_\pi^{u/\overline{u}} = 0.600^{+0.007}_{-0.007}, \qquad \lambda_\pi^{d/\overline{d}/r} = 13.15^{+1.57}_{-1.45} \tag{4.36}$$

Also the obtained χ^2 distribution is practically identical. This indicates that the separation for rest quarks is mostly irrelevant for the cross section, and for the description of the data. At low energies $Q < 10\,\mathrm{GeV}$ the contribution of the sea quarks, especially those with higher mass, can be neglected. In the collision between pions ($u\overline{d}$) and protons (uud) or neutrons (udd) at small energy suggests that the valence quarks will be most relevant for process.

However, this does not yet explain the big difference in the behaviour of quarks of u and d flavour. We can argue that for the quark-photon-vertex the $u\overline{u}$ flavour is preferred due to its higher EM-charge. This should be a reason due to which the $u\overline{u}$ flavour contributes the most to the cross section and thus we can extract it the most reliable from data.

The comparison between experimental datapoints and the provided theory prediction for the x_F differential cross section is given in figure 21 for E615 and figure 22 for E537. The plots for Q differential cross section are given in appendix D.

D.d Discussion

We confirm a large part of the difference between our theory prediction and the E615 experiment can be associated with a systematic error in agreement with the finding in [81]. In table table 8 it can be seen in numbers that the large χ^2 for small x_F values is a consequence of this systematic error. In figure 21 it is can be seen that the theory prediction (without shift in grey) systematically is below the data - by a factor of ~ 3 for small x_F bins. For moderate $0.3 < x_F < 0.6$ there still is a factor of 2 difference. The situation improves for growing x_F values. At the largest value of x_F that is included in this extraction, the data can be qualitatively described the best without a shift. For the fit quality we can observe an optimum in the agreement for the regions $0.5 < x_F < 0.6$ and $0.6 < x_F < 0.7$ with χ^2 values of 0.59 and 0.76, respectively. For the last included bin $0.7 < x_F < 0.9$ we can still describe the data well with $\chi^2 = 1.01$, however for the remaining bins with values of $x_F > 0.8$ we obtain a much worse χ^2 value. This confirms that the large x_F region cannot be described with the framework. The bin of highest x_F value cannot even be described qualitatively figure 21.

Figure 22 shows the comparison to the E537 experiment. We stress that this data was not used in the fit at all. From the data plot it can be seen that the qualitative data description is given by our prediction. One should note that the prediction is actually the grey band, which clearly has its only characteristic maximum at $q_T \sim 1\,\text{GeV}$. Because of the oscillation from the data the blue band is deformed such that it even produces several maxima. This underlines the problem of this data set. To use it in a fit, one should restrict the data to the region between $0.1 < x_F < 0.6$ where these oscillations are not (as) pronounced. The problem of these oscillations is also present the Q-differential data (see figure 45 in the appendix D). We can achieve a very good result in terms of χ^2/N_{pt} nevertheless. This actually manifests the argument of neglecting this data. Since the experimental errors are large, the description in terms of the test function is very good, even though we only qualitatively describe the data. On the other hand, for E537 the normalisation is not a significant problem, as it is for E615.

In [81] the kinematic constraint is loosened to $\delta = 0.3$[34]. With this softer constraint, there are 80 datapoints inside the fit in total over all x_F binned data included. The resulting data description we observe then is of lower quality with $\chi^2 \approx 1.55$[35]. For the larger cut in δ we observe that a modification of the RAD becomes helpful in this case. I.e., adding an additional term to the RAD proves to have significantly more impact on the fit quality than adding parameters to the model for the pion. This was tested explicitly with a modification of the RAD eq. (4.13) to

$$\mathcal{D}_{\text{NP}}(\boldsymbol{b}) = c_0 \boldsymbol{b}\boldsymbol{b}^*(\boldsymbol{b}) + c_1 \ln\left(\frac{\boldsymbol{b}^*}{B_{\text{NP}}}\right) + c_2 \boldsymbol{b}^4,$$

where the introduced term $\propto c_2$ should add flexibility to the larger $\boldsymbol{b}$ regime and the other parameters are constrained by the proton TMDPDF fit. Here we require c_2 to be a small modification and not intefering with the small $\boldsymbol{b}$ region $\sim\ 1.5\,\text{GeV}$ and smaller, such that the previous extraction stays valid. The modification does not provide a benefit for the cuts as used in the fit ($\delta < 0.25$). We understand this as an effect of tuning to data of (relatively) larger transverse momentum, which was not included in the previous fit to determine the proton TMDPDF and the RAD. On the other hand, this suggests that the cut should not be changed inconsistently with the used cut for proton TMDPDF and RAD determination.

Also in the determination of the pion TMDPDF we use the proton TMDPDF from section C.b and not ART23 directly. With the result of ART23, a fit yields a significantly worse result. To put this effect into numbers, without the update on the model, the fit saturates at a $\chi^2 \sim 1.7$, i.e. adding parameters and using a more involved functional form for the pion TMDPDF does not lead to a better fit quality. Even a fit with released proton TMDPDF and

[34] In the article the precise value is not given, but it is mentioned clearly that not a precise cut $\delta < 0.25$ is used. From the publicly available [74] script file, the value $\delta < 0.3$ can be deduced; this cut also matches the obtained 80 datapoints from set E615.

[35] This number is obtained allowing for recalibration of the NP parameters.

Experiment	x_F / Q	$N_{\rm pt}$	$\chi^2_D/N_{\rm pt}$	$\chi^2_\lambda/N_{\rm pt}$	$\chi^2/N_{\rm pt}$	$\chi^2/N_{\rm pt}$	$\langle d/\sigma\rangle$
E615–dx_F	$0.0 < x_F < 0.1$	9	0.29	1.36	1.66	$1.66^{+0.02}_{-0.02}$	56.1%
	$0.1 < x_F < 0.2$	9	0.41	1.22	1.62	$1.62^{+0.03}_{-0.03}$	53.0%
	$0.2 < x_F < 0.3$	9	0.69	1.21	1.90	$1.89^{+0.04}_{-0.03}$	52.7%
	$0.3 < x_F < 0.4$	9	0.89	0.84	1.74	$1.74^{+0.03}_{-0.03}$	44.1%
	$0.4 < x_F < 0.5$	9	0.91	0.64	1.55	$1.55^{+0.03}_{-0.03}$	38.4%
	$0.5 < x_F < 0.6$	9	0.29	0.32	0.60	$0.59^{+0.03}_{-0.03}$	27.1%
	$0.6 < x_F < 0.7$	9	0.59	0.17	0.76	$0.76^{+0.01}_{-0.01}$	19.8%
	$0.7 < x_F < 0.8$	9	0.93	0.05	0.98	$1.01^{+0.04}_{-0.04}$	10.6%
	$0.8 < x_F < 0.9$	9	8.12	0.09	8.21	$8.15^{+0.31}_{-0.29}$	14.5%
	$0.9 < x_F < 1.0$	9	17.29	0.45	17.74	$18.12^{+0.94}_{-1.09}$	32.2%
Total fit (E615)	$0.0 < x_F < 0.8$	**72**				$\mathbf{1.42^{+0.05}_{-0.04}}$	
E615–dQ	$4.05 < Q < 13.05$	77	2.65	0.10	2.75	$2.44^{+0.10}_{-0.08}$	44.6%
E537–dx_F	$-0.1 < x_F < 1.0$	110	0.47	0.12	0.59	$0.59^{+0.01}_{-0.01}$	23.9%
E537–dQ	$4.0 < Q < 9.0$	60	0.75	0.13	0.88	$0.88^{+0.01}_{-0.01}$	15.9%

Table 8: Resulting χ^2 distribution for the datasets E537 and E615. For the central replica we show the decomposition into χ^2_D and chi^2_λ representing the contributions to the χ^2 from disagreement with shape and normalisation of the data, respectively. The last column is the average shift also for the central fit. For the replica average we also provide the 68% CI. The datasets from E615 in $x_F < 0.7$ bins are included in the fit. The dataset E537 here is over the whole x_F range. All datasets which are not used for the fit are shown in grey and they do not contribute to the fit χ^2.

RAD parameters to the 72 datapoints does not resolve the issue. As in the case of SIDIS, the description was improved *ad-hoc* with the constrained flavour dependent ansatz. This underlines the importance of this adaption of the ART23 result.

We think that it is not enough to give a reliable error estimation with variation of only data and pion PDF. The even though we produce a very good fit result, the uncertainties in [82] appear to be more realistic. The procedure must be repeated for a variation of also proton TMDPDF.

Several typical impact studies can be performed with the given result. The input functions, which applies to the proton TMDPDF and to the collinear PDF for the pion. During the development of the ansatz we already compared to the previous version `JAM18PionPDFnlo` [85], where we did not observe a strong effect, yet a quantitative comparison is standard.

Because in our framework the cost of computing the cross section for the pion induced DY process is negligible in comparison to the cost of computing the theory prediction in the proton TMDPDF extraction, one could think of a simultaneous fit. As demonstrated, the model for the pion TMDPDF does not require many additional parameters. Probably the most important next milestone in the extraction of the pion TMDPDF is the release of the data measured by the COMPASS collaboration; this might also drag more attention to their determination.

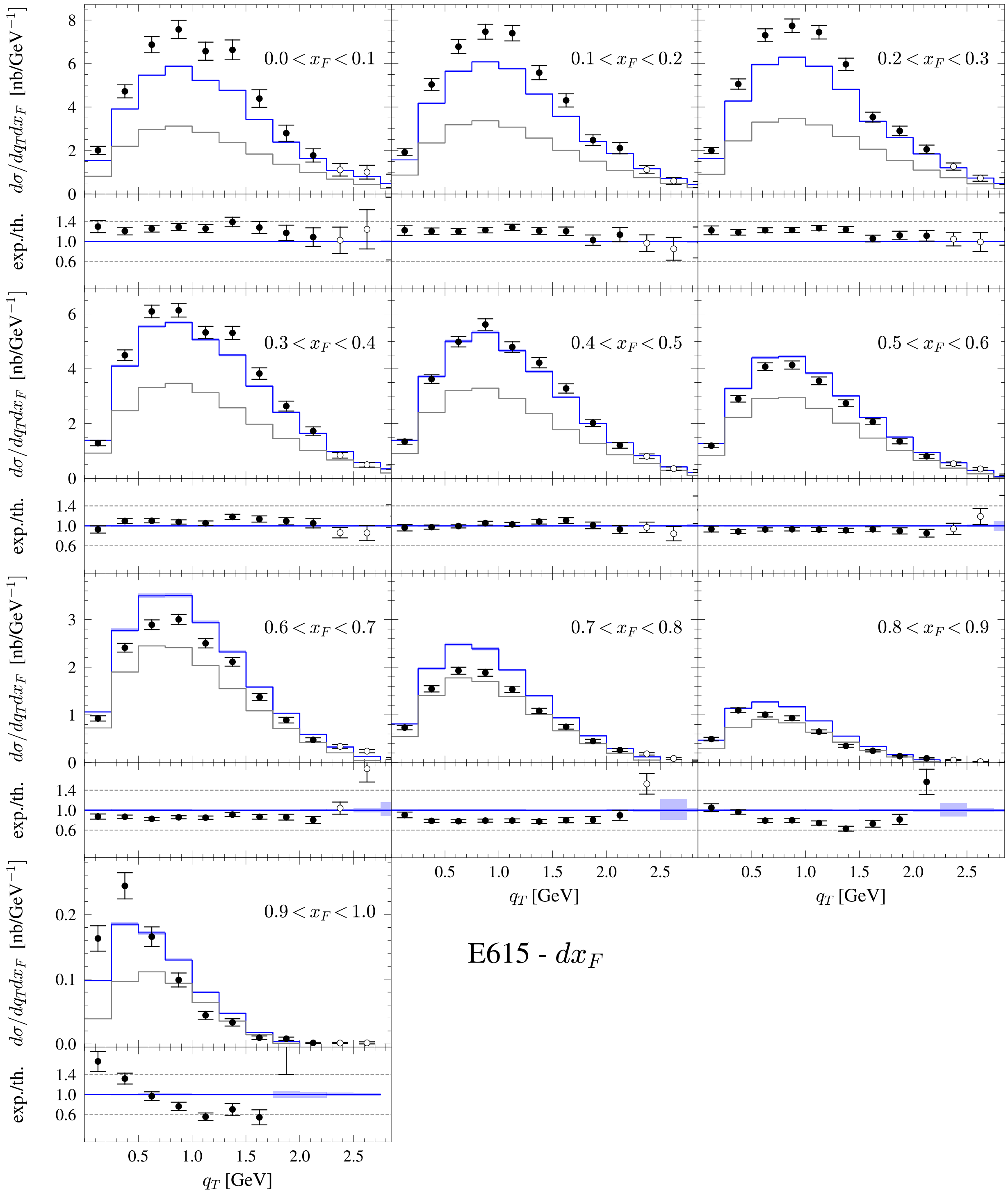


Figure 21: Comparison of the theory prediction with E615 experiment differential in x_F. The blue (grey) line is the prediction with(out) shift as defined in eq. (4.18). The filled (empty) dots are data which are (not) included in the fit. The blue band shows the uncertainty for the prediction. It is computed with simultaneous variation of data and the collinear pion TMDPDF.

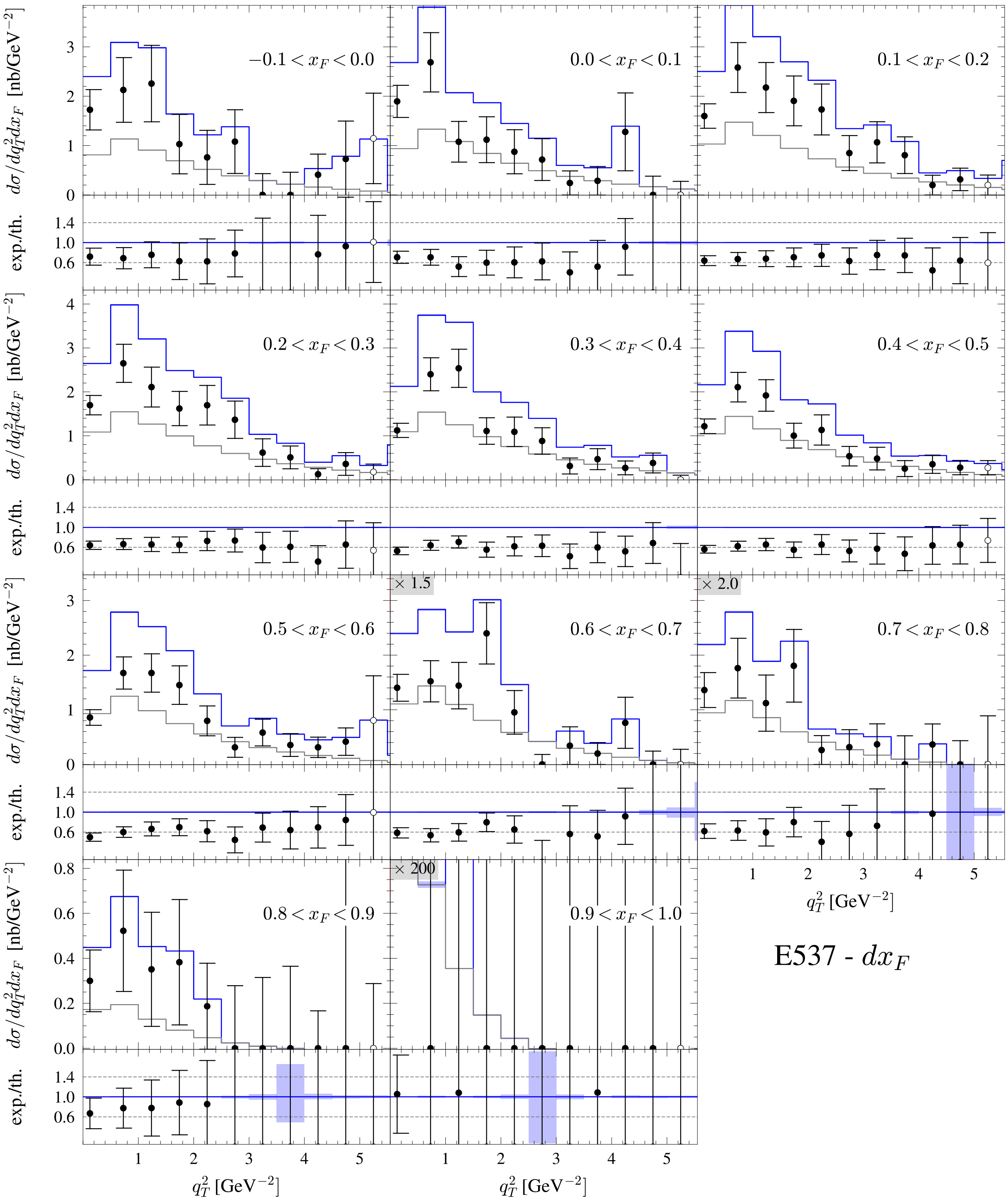


Figure 22: Comparison of the theory predicion with E537 experiment differential in x_F. The blue (grey) line is the prediction with(out) shift as defined in eq. (4.18). Note that none of these data were included in the fit. The filled dots pass the cuts to be hypothetically relevant for a fit, the empty dots are data which would be discarded. It is computed with simultaneous variation of data and the collinear pion TMDPDF.

V Treatment of TMDs in Pythia8

As a general motivation to this section we refer to the Electron Ion Collider (EIC) [86]. It is a planned collider that will allow for new measurements of (SI)DIS events at a range of centre-of-mass energies $\sqrt{s} = 20 - 100\,\text{GeV}$. This will be an extension of the kinematic region available from HERMES [18] ($\sqrt{s} = 7\,\text{GeV}$) and COMPASS [19, 20] ($\sqrt{s} = 17\,\text{GeV}$). EIC is also going to provide data for polarised cross sections, which will allow for a better determination of more functions characteristic to structure of nucleons (e.g. the Sivers function $f_{1T}^{\perp}$ in the context of TMDs).

To predict the measurables before the experiment exists, we can use theoretical models, or Monte Carlo Event generators (MCEGs). These are simulation software combining perturbative physics and models for non perturbative processes, e.g. hadronization, and ultimately are tuned to (mostly) high energy collision experiments. The MCEGs are indispensable for modern particle physics, as they are required in tuning detectors for experiment [87, 88], or to tune jet-finding algorithms [89, 90], to name just two of the many applications.

The most commonly used software is the `Pythia8` package. It is the `C++` successor of the `Pythia6` [91] package. Many other MCEG like `Cascade3` [92], `Herwig7` [93] or `Sherpa` [94] intrinsically use part of the Pythia package themselves. All of these packages are general purpose MCEGs, i.e., they are not designed to simulated only a single or a small handful of reactions, but in principle can simulate any scattering in the Standard Model.

As we are interested in TMD distributions, we are curious about what can be expected from EIC in that direction. To this end this study is determined to understand to what extent the MCEGs, namely `Pythia8`, are able to generate TMD relevant pseudo data in SIDIS.

In this section we compare to the data provided by the HERMES and COMPASS collaborations, and discuss possibilities to modify the the models in `Pythia8` for the k_T generation based on insights of TMD phenomenology. The outlook is to provide a prediction for TMD relevant data at kinematics of the EIC with the help of (a tuned version of) `Pythia8`.

A Parts of the Event Generation in `Pythia8`

In this section we sketch a single event in `Pythia8`, following the structure of the event generation and not a "time-equivalent" axis of the process. In this context we lay a special focus on two related aspects, which contribute to the transverse momentum of the particles. For most steps in the generation we point out what this means for our example of SIDIS, or what aspect herein is most important to this work. What we discuss can be found in more detail in the manual [95].

Hard matrix element. An event in MCEGs is a single reaction of interest - like in an experiment. The core of the event is the hard scattering. It is the partonic reaction and processed with the matrix element(s) among all *allowed* processes. The user can list hard reactions of interest, and the program will only take into account these types of reactions. This does not spoil any analysis, because if we want to consider only one particular process we would discard all other reactions later on anyway[36]. In the case of SIDIS description for electron in the initial and final state, we only take into account the electroweak process $k, p \to k', p'$. With this we have a matrix element from a two particle initial state to a two particle final state, where the momentum is transferred only by a single photon or Z-boson. The hard process depends on the hard scale Q, which allows to sample a flavour and an x value; for the details on this we refer to the manual.

Parton showers. The initial and final states of hard matrix element are then both treated with parton shower algorithms. They take into account an (possibly arbitrarily large) amount of radiation the particles can

[36]The MCEG can provide an estimate for the luminosity depending on the selected processes, such that the cross section can be determined. For multiplicities we compare against here, this is irrelevant because the luminosity cancels in the ratio between SIDIS and DIS cross sections.

emitt before and after the hard reaction. In the event generation in `Pythia8`, the initial state radiation (ISR) and the final state radiation (FSR) are treated separately, and can be also switched on or off independently. They can be understood as an overlap between the hard process, a matrix element, and the PDF evolution. In our cause, these mechanisms are relevant as they modify the (else) collinear momentum of the hard parton.

The **FSR** is somewhat more intuitively motivated than the ISR, thus we comment on its mechanism first. We consider the particles which are products of the hard matrix element at a large virtual scale[37] t. The hardest process emits only a small number of particles, and cannot describe the generation of a final state of a very large number of final states directly. The FSR adds corrections to this by considering splitting processes. These are related to $1 \to 2$ matrix elements of the kind $\mathcal{M}\left(q \to qg\right)$. The concept of these branchings is not unique to the treatment of QCD, in this overview we however omit the QED case which follows the same concept. In QCD branchings we thus have the probability of a parton of type a to branch into two partons of type b and c, formulated in differential form as

$$\mathrm{d}\mathcal{P}_a(z,t) = \frac{\mathrm{d}t}{t}\frac{\alpha_s(t)}{2\pi}\sum_{b,c} P_{a\to bc}(z)\mathrm{d}z. \tag{5.1}$$

This expression corresponds to the probability density for splitting at the scale t and to distribute the momentum of a such that b receives a fraction z and c has the remaining part of $1-z$. The larger branching scales would be associated to smaller space-time scales, and hence are expected to occur closer to the hardest matrix element than splittings at a lower scale. The scale can be understood as an ordering quantity, that orders the emissions in "time". A spontaneous emission at t_1 happens "before" an emission at scale t_2 for $t_1 > t_2$. Lower energy scales are closer to the physical final state of the event.

The splitting kernels are *essentially* the same kernels as in leading order perturbation theory (DGLAP) eq. (2.29)-eq. (2.32), yet boundary terms are not implemented in the same way. In `Pythia8` the splitting functions are implemented as

$$P_{q\to qg}(z) = \frac{4}{3}\frac{1+z^2}{1-z}, \tag{5.2}$$

$$P_{g\to gg}(z) = 3\frac{\left[1-z\left(1-z\right)\right]^2}{z\left(1-x\right)}, \tag{5.3}$$

$$P_{g\to q\bar{q}}(z) = \frac{1}{2}\left[z^2+\left(1-z\right)^2\right]. \tag{5.4}$$

In contrast to the DGLAP kernels, the `Pythia8`[38] implementation does not require terms that regulate the behaviour at the boundaries $z = 0, 1$. The last channel is normalized to $q\bar{q}$ only and needs to be summed over all (kinematically allowed) flavours.

With the splitting to a specified distribution of the momentum z in eq. (5.1), the differential branching probability is obtained by integration over the (scale dependent) region of z,

$$\mathrm{d}\mathcal{P}_a(t) = \frac{\mathrm{d}t}{t}\frac{\alpha_x(t)}{2\pi}\int_{z_{\min}(t)}^{z_{\max}(t)} \mathrm{d}z P_{a\to bc}(z). \tag{5.5}$$

Conversely, the differential probability not to fragment is $1-\mathrm{d}\mathcal{P}_a$, and thus the probability for a particle a

[37]We label the scale evolution as t for two reasons: The classical label of Q refers to the hard scale, which the resolution scale is related but not identical to. Also, it can be pictured as a *virtual* time evolution.

[38]Or MCEGs in general [95]

not to split over a certain range of scale $t_i > t_f$ is given by

$$\Pi_a\left(t_i, t_f\right) = \exp\left(-\int_{t_f}^{t_i} \mathrm{d}\mathcal{P}_a(t)\right), \tag{5.6}$$

called a *Sudakov* factor.

The generation of transverse momentum due to each $1 \to 2$ splitting is implicitly constrained. The momentum of the produced particles b and c is transverse to the momentum of a by conservation of four-momentum and their masses. We can make this explicit by considering the "light-cone" coordinates of the particle a. The components are

$$p^{\pm} = p_0 \pm p_z, \tag{5.7}$$

and if we consider the z axis to be aligned with the momentum of a, its virtual mass is expressed in $+$ and $-$ components

$$p_a^+ p_a^- = m_a^2. \tag{5.8}$$

The virtual mass of the particle a is the relevant branching scale t. If the particle splits to two particles and the momentum along the $+$ direction is split as in eq. (5.1), momentum conservation yields a gained transverse momentum for both particles b and c of

$$p_\perp^2 = z'(1-z')m_a^2 - (1-z')m_b^2 - z'm_c^2. \tag{5.9}$$

The labels z' used here indicate the difference from z previously used, because the fraction z' is distributed along the $+$ direction, whereas z is the fraction regarding the total momentum. The latter corresponds to a fully collinear splitting. The splitting is not implemented via the z' "light-cone" notation in `Pythia8`, however this shows the relation between the evolution scale and the transverse momentum of the branching.

The **ISR** is the evolution that relates the hard scale and the soft $\sim$ hadronic scale of the initial state which works in analogy to the FSR. The evolution of a initial state parton of the hadron along the scale to the hard process, however, is random, and not deterministic. Thus it is possible to obtain a parton at the hard scale, which does not contribute to the hard process, and the event has to be discarded. Not a problem in principle, but a loss of computational time: All the computational effort to determine the configuration of the hard particle would come before it could be determined whether the processes is relevant and should further be considered or discarded. It is thus reasonable to use the PDFs to provide a parton of flavour and momentum configuration at the hard scale in the first place. From this, the question arises why one now needs ISR, while the PDFs evolution already takes care of that; the PDFs provide us indeed with the correct flavour and momentum configuration for the hard matrix element already. However, the radiation of particles is absorbed in the PDFs evolution, which provides us with the hard particle for the partonic hard process only, and in that sense, information about the radiated particles is lost. In perturbative factorisation, there is no problem with this. The factorisation of a process specifies its initial state as well as the part of the final state of interest, leaving a large unspecified final state. In the `Pythia8` all momenta of all particles need to be considered.

In `Pythia8` the ISR evolves *backwards* in the "virtual time" from the hard scale down to an initial scale. This idea was presented in the literature [96] already in the 1980s by Sjöstrand. As we understand from [96] we relate the outmost particles of the parton shower to be the aligned in direction with the hadron they belong to. Their direction of motion coincides with that of their hadron, up to a small random transverse motion of the the parton.

This initial transverse momentum is labelled *primordial* k_T and discussed later in greater detail.

We conclude the section about the initial and final state parton showers and refer to the manual for a detailed description of their formal foundation and precise details about their implementation. The statement we want to stress here is that the parton showers are in analogy to the scale (DGLAP) evolution, and the branching results in a generation of transverse momentum.

Multi parton interaction. The FSR radiation of the final state of the matrix elements too changes the (transverse) momentum of the parton, which, in DIS, is the same quark type as the initial parton. The mechanism of FSR in `Pythia8` interferes with multi-parton-interaction (MPI), which considers a second (or more) sub-ordered hard event. A study on MPIs can be found in [97], where it is explained that in hadron-hadron collisions it is expected to have several processes which appear to occur on perturbative scales; to our cause, however, this mechanism is not relevant. In SIDIS we only take into account the hardest process, and the hadron forming from the hard parton. We argue that this fact combined with the relatively low centre-of-mass energy MPI does not play a relevant role for the SIDIS cross section.

Hadronization. The mechanisms relevant after the hard scattering and the parton shower corrections can be summarized as hadronization. In `Pythia8`, the *Lund String Model* treats any QCD colour splitting with the assignment of *unique* colour indices. This is implemented QCD in the limit $N_c \to \infty$, $\alpha_s N_c = \text{const.}$. This limit has theoretical foundation [98] and determines between which particles confining potentials arise, ultimately leading to hadronisation. As this part is not subject to our main interest, for a more detailed discussion of hadroisation in `Pythia8` and the Lund String Model [99], we refer to the manual.

B Standard code data description

We first compare a list of different general purpose MCEGs with which we generate pseudo data for the kinematic regions and cuts used in the data from the HERMES and the COMPASS collaboration. The p_T dependent data from COMPASS are also used for in the SIDIS fits for TMDs in the earlier section IV.C. We collect all technical aspects on the event generation in appendix E and do not discuss them here.

For the HERMES experiment we generate the pseudo data with `Pythia6`and `Pythia8`. We find that both programs are not able to describe the data on a quantitative level. A qualitative description is arguably given; as can be seen in figure 23. There is a significant underestimation of the data by the MCEGs. We study the HERMES data in the different binning the data are provided in. This includes binning in the single variable z, and two dimensional bins of the pairs of variables (Q^2, z), $(x + z)$, $(z + p_T)$. For all of them we are not able to achieve the reproduction of the data with the MCEGs we test (`Pythia8` and `Pythia6`). There are bins where the description is adequate, which is demonstrated in the figure 23 for the larger z values. In the figure it can also be seen that, for the two bin with smaller z, the pseudo data follow the slope of the experimental data by underestimate them by a factor of ~ 2. This selected plot should summarize our observations; the quantitative description of the data measured by the HERMES collaboration is not given by `Pythia8` or `Pythia6` either.

We reason this with the argument that the HERMES experiment with $\sqrt{s} = 7$ is too low energy to have a chance for a good description with the MCEGs we use in our study, which are tuned to higher energy experiments. Because we observe these quantitative differences in the one and two fold differential multiplicites, we decide to disregard the data in the further discussion.

The COMPASS collaboration provides SIDIS multiplicities as three fold [19] and four fold differential [20]. The three fold multiplicities are given in bins in the variables in y, x_B and z, and in the publication also y-integrated comparisons can be found. We mention that the data we present for the two-fold differential multiplicities we compute ourselves from the unintegrated data available. The data is presented like in the two differential from in the publication, but not directly accessible in this format.

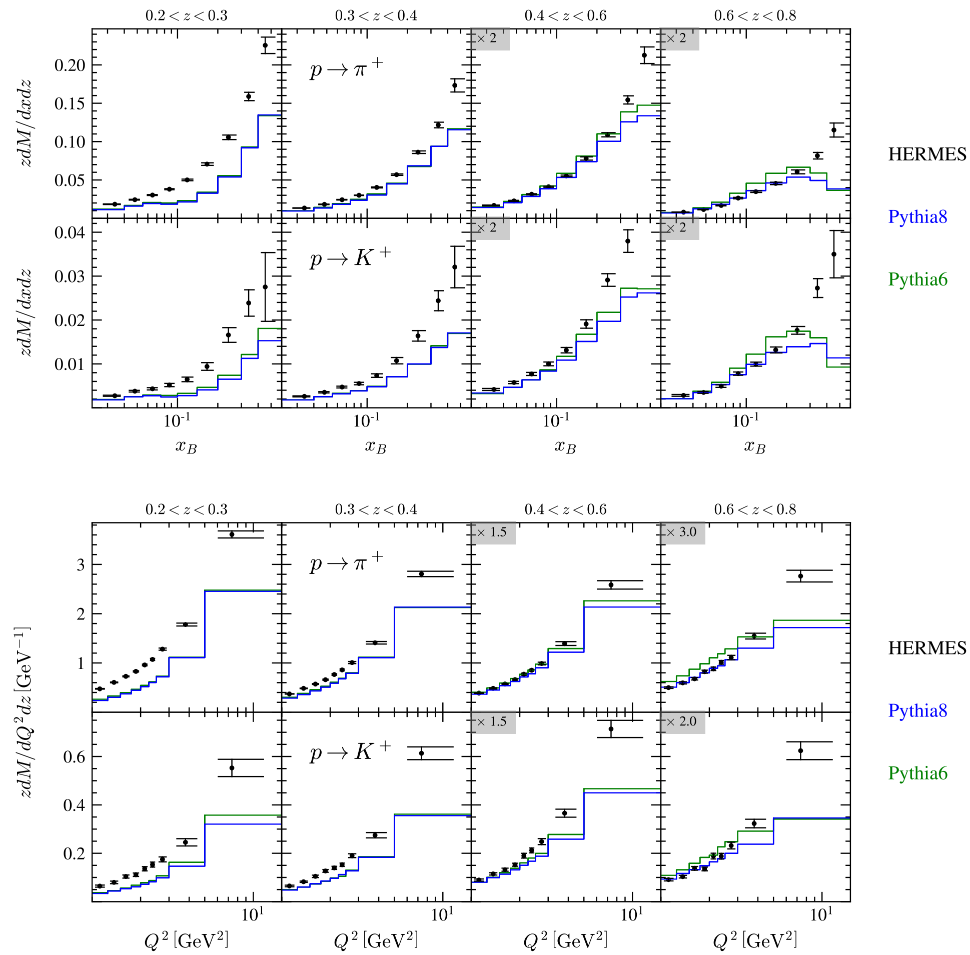


Figure 23: Two fold differential SIDIS multiplicity of meson production measured by HERMES, compared to results from `Pythia6` and `Pythia8`. The upper panel shows the multiplicity in (x, z) bins, the lower panel is in (Q^2, z) bins. For reasons of presentation some figures are rescaled by a factor indicated in grey.

The four fold multiplicity is the most relevant to this work; it depends on the final state hadrons transverse momentum p_T, and is further binned in the variables Q^2, x_B and z. In order to configure our setup, we compare our generated pseudo data also for the p_T unobserved data. In this step we discuss the quality of the description in three variables. This helps to distinguish between effects from taking into account p_T and issues with the description present at the level of an analysis in the three other variables.

In figure 24 we present the results obtained from different MCEGs for two fold differential multiplicity in the variables x and z. Qualitatively, all of them are able to describe the data. Quantitatively the description is not very good. We can see that we have up to 50% deviation from the experimental data and the deviation is larger than the experimental uncertainty. In the analysis we show the results from different MCEGs with the same PDF - for `Cascade3` a TMDPDF set is required as an input function. The shown MCEG results are generated with `HERAPDF15LO_EIG` [68]. A variation of the input PDF does not significantly change the result. For this compare the generation of pseudo data with `NNPDF3.1_nnlo_as_0118_1000` and `CT10` [100], and find that the input PDF is not responsible for the quality of the description of the data.

The overall description from MCEG side is better than for the data from HERMES, arguably due to the higher centre-of-mass energy.

The same data with their dependence on y taken into account in bins are as well compared to all MCEGs. The results of this comparison are shown in figure 25. We observe that the data are described the best by `Cascade3` and `Herwig7` . `Pythia8` shows the strongest deviation from the data, which appear to be systematically overestimated by the program. All programs show a tendency to produce multiplicities below/above/below the data for low/medium/high values in z as shown in the figure. For largest x_B bins, the description is worse for all programs than for low and intermediate bins in x_B.

In the following we compare to p_T dependent SIDIS data measured by COMPASS. They are the most relevant data for this study. The previous analyses should ensure that the MCEG are tuned and used correctly, and also are necessary to understand the magnitude of the discrepancy between MCEG pseudata and experimental data for the p_T integrated/unobserved measurements.

The transverse momentum of the final state hadron is transverse with respect to the direction of the virtual photon. For the analysis we thus decompose the momentum of the final state hadron in the lab frame as

$$\vec{p}_H = \frac{\vec{p}_H \vec{q}}{\vec{q}^2}\vec{q} + \vec{p}_T. \tag{5.10}$$

Then p_T^2 is the modulus squared of the three component vector in the decomposition. At a late stage of this work it was noticed that the definition of transverse direction is given via the relation eq. (2.54). Thus the definition in four momenta should be used instead of the decomposition in three components. In the case of a fixed target, however, the momentum $P = (M, \vec{0})$ only has an energy component. Demanding transversity with respect to the photon q and the proton P, thus, is equivalent to the transversity to the photon in the spatial components. A fortunate coincidence! We explicitly verified that an implementation of either of the definitions of the transverse momentum yields the same result on the example of the `LEPTO` events[39].

As for the other configurations, we compare all MCEGs against the data. Here we also compare to the MCEG `LEPTO` [101], which was used as a reference MCEG by the COMPASS collaboration [20]. A selection of representative kinematic regions is shown in figure 26 and 28. To clearly show the slope and the relative agreement with the data, we display the kinematic bins in two panels. In the upper panel, we show the absolute experimental values and the predictions from the MCEGs. The latter are shifted each by a differrent factor to separate them in the plot for a clear display. In the lower panel, we normalize all values to the experimental data, to better display the relative

[39]See next paragraph.

differences.

Our observations are the following. The qualitative description of the p_T slope is reproduced by all MCEGs, yet there are qualitative differences. The best description is achieved by `LEPTO`. For the discussed bins, `Pythia6`and `Cascade3`seem to overestimate the low p_T generation and underestimate the generation for p_T in the larger region. `Herwig7`shows the opposite behaviour, underestimating low p_T production while overestimating the cross section for the larger p_T values. For most bins and these three MCEGs, the p_T integration over the bin could balance the over- and underestimation of the data. Indeed, we also observe that for some kinematic regions the amount of events becomes small enough such that statistical fluctuations become visible. This could be cured by simply running more statistics, or by merging to larger p_T bins. The point we want to stress here is that these fluctuations do not blur the tendencies we observe and describe.

C Intrinsic transverse momentum treatment in `Pythia8`

After investigating the default description of the data, we discuss one part of the event generation in `Pythia8` that was not adressed before. The parton showers are associated to the scale evolution before and after the hard event, contributing also to the transverse momentum of the hard parton. They model the evolution of the momentum (and flavour) of the parton from the initial to the hard scale. The initial transverse momentum of the evolution cascade of branchings is called *primordial* k_T.

In the following we discuss how `Pythia8` generates the *primordial* k_T and how we can access it. The width σ of the Gaussian distribution according to which the primordial k_T fo the hard collision is sampled is modelled as

$$\sigma(Q) = \frac{\sigma_{\text{soft}} Q_{1/2} + \sigma_{\text{hard}} Q}{Q_{1/2} + Q} \frac{m}{m_{1/2} y_{\text{damp}}}. \tag{5.11}$$

The second fraction is introduced for technical reasons and should not have an influence on the observables [95]. Thus we focus only on the first factor in the following discussion, which we denote by σ^*. The first factor consists of two parameters, σ_{soft} and σ_{hard} which are the values for $Q = 0$ and $Q = \infty$, respectively. The transition rate depends on the parameter $Q_{1/2}$, which defines the scale for the midpoint between the two limits.

In order to understand what enters the hard process we examine an event tree of `Pythia8`. In an event tree, all particles of the reaction are listed, including both virual and final(observable) particles. The particles are listed with their quantum numbers[40] and their four momentum. For the electron-nucleon scattering we find, that the particle `event[4]`[41] is the quark which scatters with the electron. We can influence the momentum components from the routine in the source file `BeamRemnants.cc`. To this end we extract the k_T of the `event[4]`. With this default description (eq. (5.11)) the resulting distribution is expected to be Gaussian; however the ISR strongly distorts the resulting picture. With ISR switched off[42] we can obtain the k_T of the hard quark resulting from eq. (5.11).

The purpose of this study is to change the model eq. (5.11) such that we improve the resulting distribution of final state p_T in SIDIS. From TMDPDF phenomenology we can have naive ideas how the distribution should look like. The aspect we discuss in this study is the question, whether an x dependent model can improve the description of the data. This reflects that the ansatzes for TMDPDFs are parametrized as k_T and also x dependent[43] functions, see e.g. table 1.

Ultimately, it is also possible our adaptions of the following work could be (partially) compensated by tuning of the two existing parameters σ_{soft} and σ_{hard}. Indeed we will see that the adaption of the Gaussian width is helpful

[40] They are listed by an identification index.

[41] In the nomenclature of `Pythia8` and `event` of an index is a particle that exists or existed during the event, and we use this labeling here, too.

[42] The ISR and FSR can independently be switched off by a flag in the program file.

[43] The parametrization is in Fourier conjugate space to k_T.

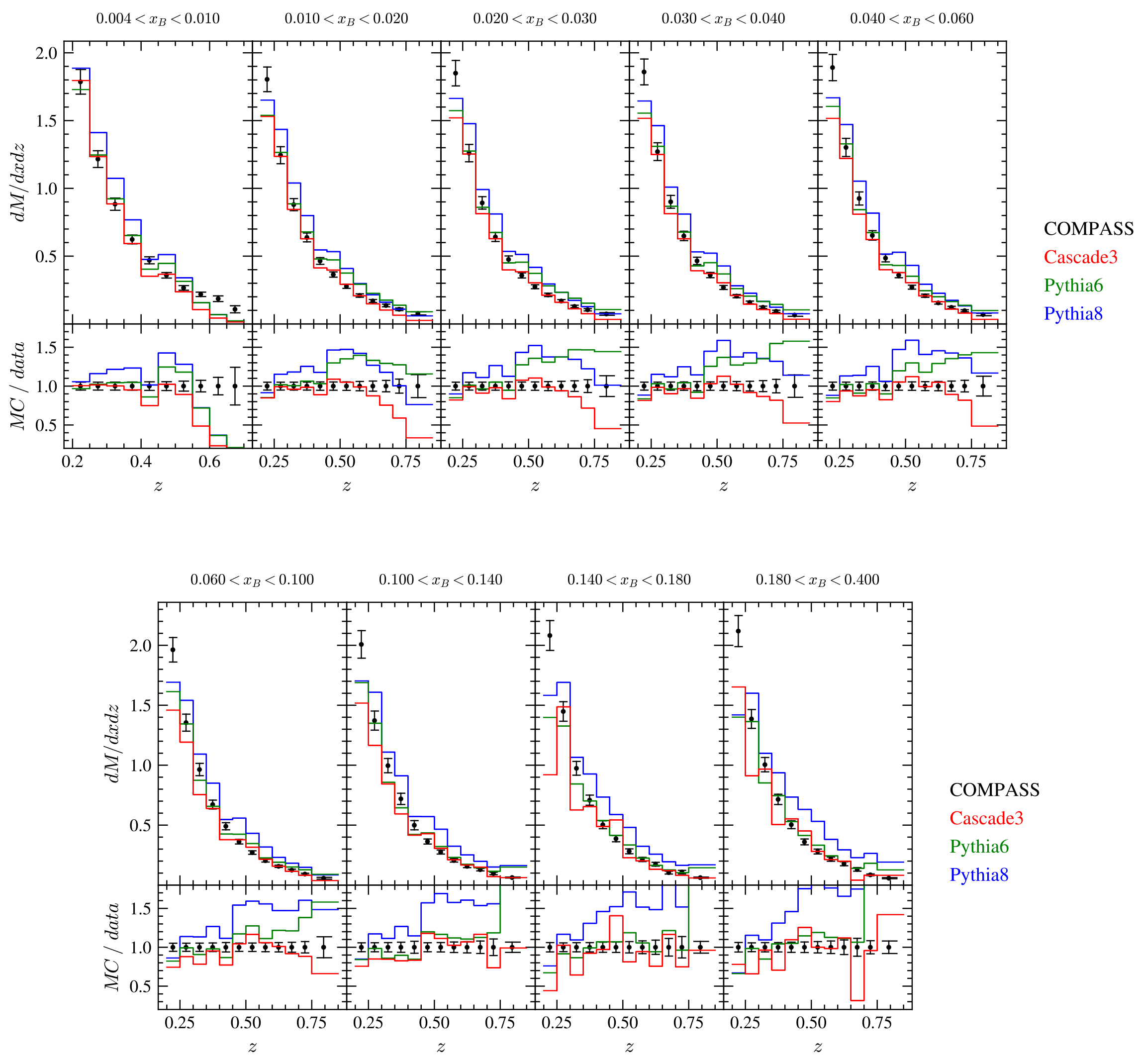


Figure 24: Two fold differential multiplicity of charged (+) pion production measured by COMPASS, compared to results from different MCEGs. `Pythia6`, `Pythia8` and `Herwig7` pseudo data are generated using the PDF `HERAPDF15LO_EIG` [68]. `Cascade3` results are generated with the TMDPDF set `PB-NLO-HERAI+II-2018-set1`.

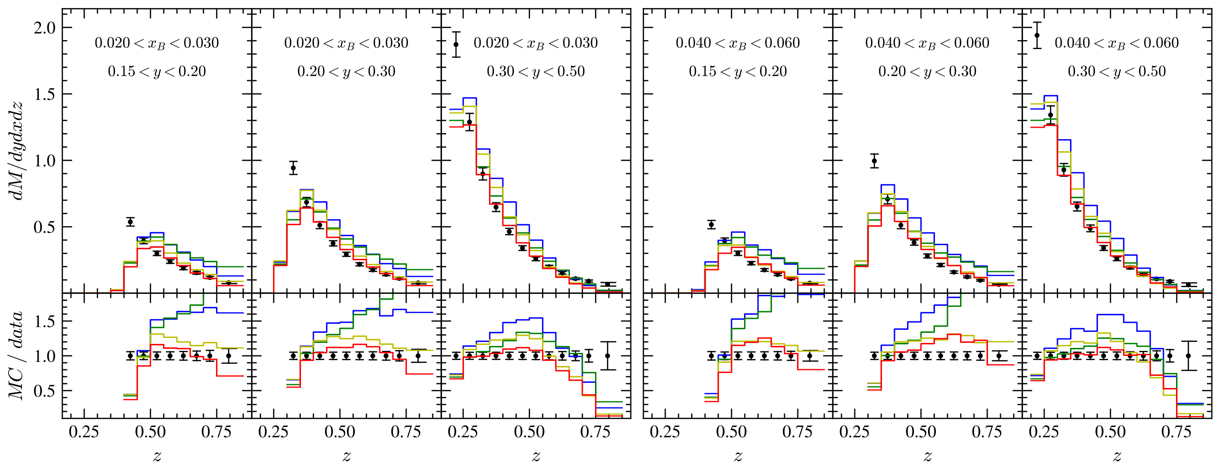


Figure 25: Three fold differential multiplicity of charged pion production differential in x_B and y bins, and plotted against their z bins, measured by COMPASS and compared to results from different MCEGs. `Pythia6`, `Pythia8` and `Herwig7` pseudo data are generated using the PDF `HERAPDF15LO_EIG` [68]. `Cascade3` results are generated with the TMDPDF set `PB-NLO-HERAI+II-2018-set1`.

to improve the qualitative description of the data for small Q^2. At the same time, we regard these parameters to be part of a tuned program which is valid in a large kinematic range. We do not want to effectively spoil the algorithm at larger $Q^2 \sim 10\,\mathrm{GeV}^2$ and beyond. Thus our modification of the k_T width would not only depend on the momentum fraction x but also on Q^2, to regulate the modification at larger Q. The Q dependence on the model, however, is not a new feature.

D Modifications of the source code

In the following we present two modifications which overall improve the description of the p_T dependent SIDIS data. The precise implementation is given in appendix E to separate our discussion here from too technical aspects. One of the two suggestions does not affect the current description of `Pythia8` for large Q scales, as it converges to the default description for large resolution scales, but has an effect in the low Q^2 region. The second modification we suggest, as an alternative to the first one, overwrites the current parameter $\sigma_{\text{hard}} = 1.8 \to 1.1$. Thus it is possible that it significantly affects the generated observables even for large Q values.

From studying the description of the p_T dependent profile by `Pythia8`, we observe the tendency that for fixed Q^2 and z, the lower x bins show a downward slope in the relative value comparison, while the larger x bins have the contrary behaviour with an upwards slope, see figure 26-28.

To construct a way of understanding which modifications are expected have a positive impact on the description of data, we draw a naive picture. We think about the final state p_T as being the sum of the transverse momentum generated throughout the event generation (ISR, hard matrix element, FSR, hadronizsation), a part that is given and we do not influcen, and the primordial k_T. Even though these add as vectors, a large p_T^2 can be naively related to a large modulus of k_T, and a small p_T^2 to a small intial transverse momentum. We also need to analyse the behavoir of the data description (figure 26figure 28), and what it means for the distribution of the *primordial* k_T. If `Pythia8` overshoots the data for small p_T^2 and underestimates the cross section for larger p_T^2 (downward slope), it indicates that the distribution of initial transverse momentum is accumulated around small values – the width of the Gaussian eq. (5.11) is too small. If the data in contrast is underestimated for small p_T^2 and overestimated

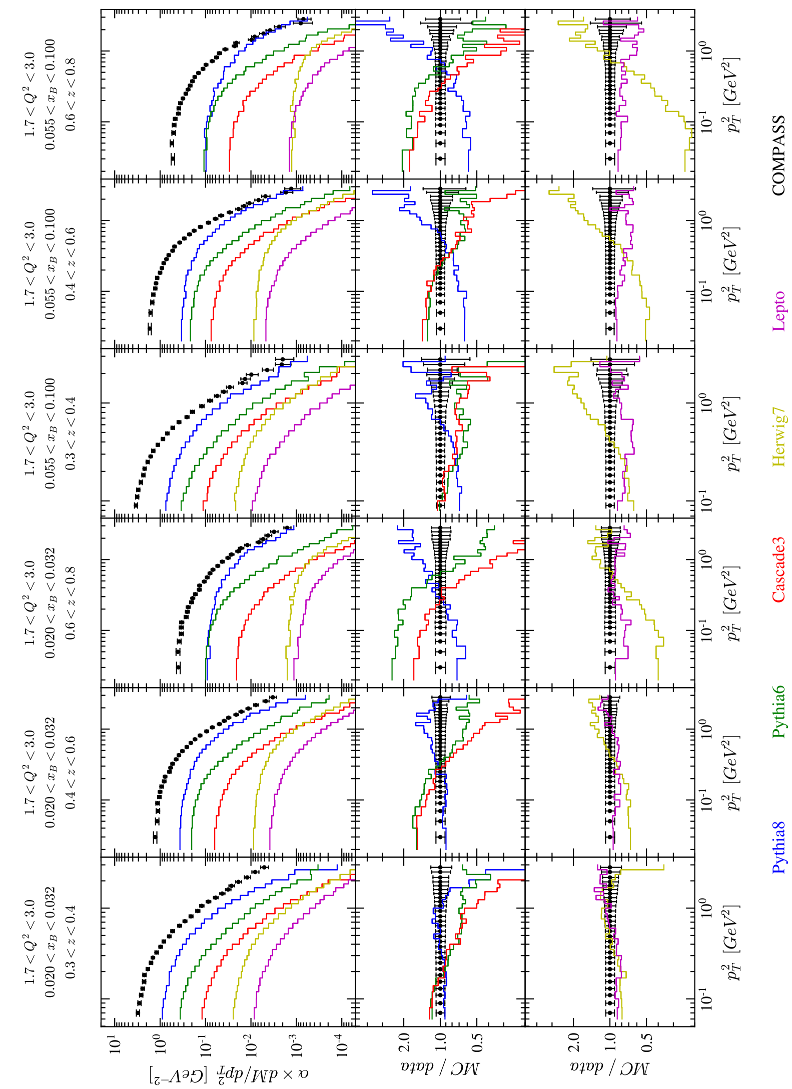


Figure 26: Four fold differential multiplicity of charged hadron production differential in bins of Q^2, x_B and z bins, and plotted against their p_T bins, measured by COMPASS and compared to results from different MCEGs. The upper row shows the absolute values of multiplicites. The MCEG data are shifted by a constant factor α for better visualization. The value of α is $\sqrt{0.1}$ for `Pythia8` and increased by a single power respectively for the other MCEGs, in their respective order. `Pythia6`, `Pythia8` and `Herwig7` pseudo data are generated using the PDF `HERAPDF15LO_EIG` [68]. `Cascade3` results are generated with the TMDPDF set `PB-NLO-HERAI+II-2018-set1`.

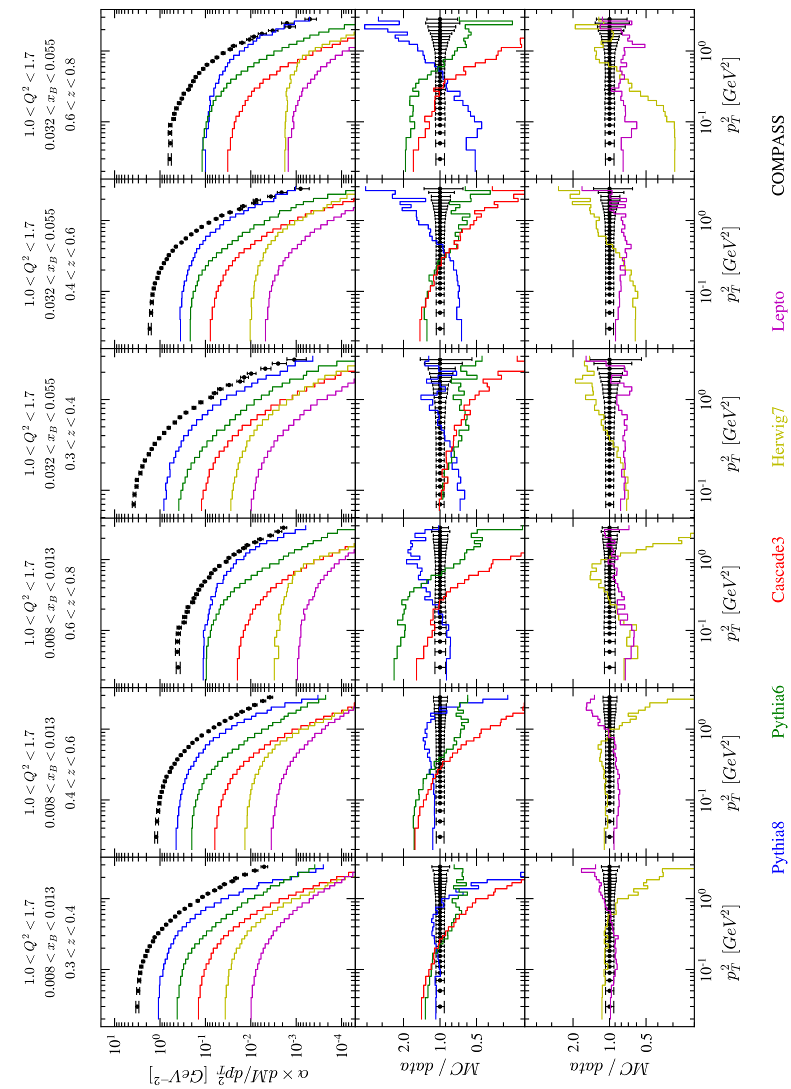


Figure 27: Four fold differential multiplicity of charged hadron production differential in bins of Q^2, x_B and z bins, and plotted against their p_T bins, measured by COMPASS and compared to results from different MCEGs. The upper row shows the absolute values of multiplicites. The MCEG data are shifted by a constant factor α for better visualization. The value of α is $\sqrt{0.1}$ for `Pythia8` and increased by a single power respectively for the other MCEGs, in their respective order. `Pythia6`, `Pythia8` and `Herwig7` pseudo data are generated using the PDF `HERAPDF15LO_EIG` [68]. `Cascade3` results are generated with the TMDPDF set `PB-NLO-HERAI+II-2018-set1`.

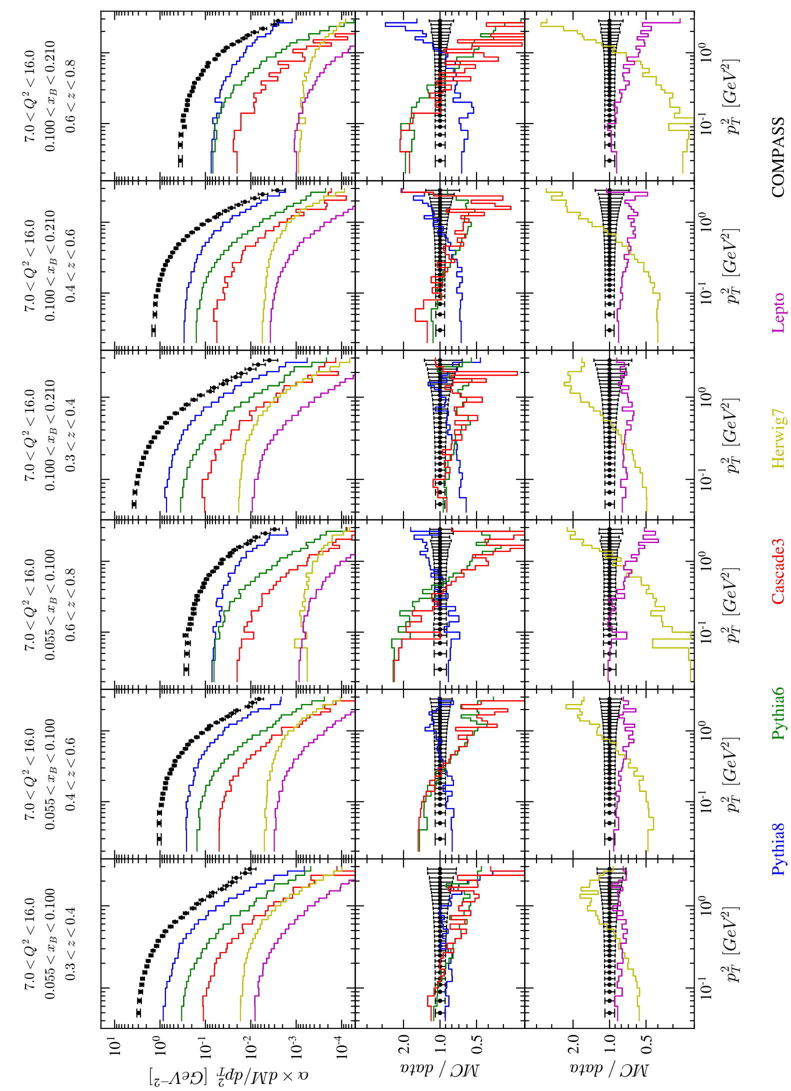


Figure 28: Part two.
Four fold differential multiplicity of charged hadron production differential in bins of Q^2, x_B and z bins, and plotted against their p_T bins, measured by COMPASS and compared to results from different MCEGs. The upper row shows the absolute values of multiplicites. The value of α is $\sqrt{0.1}$ for `Pythia8` and increased by a single power respectively for the other MCEGs, in their respective order. `Pythia6`, `Pythia8` and `Herwig7` pseudo data are generated using the PDF `HERAPDF15LO_EIG` [68]. `Cascade3` results are generated with the TMDPDF set `PB-NLO-HERAI+II-2018-set1`.

for large p_T^2 (upward slope), this means that the distribution is too spread out; the width $\sigma(Q)$ should be reduced. Our argumentation can be summarized in the diagramatic structure

- smaller $x \sim$ downward slope $\rightarrow \langle k_T \rangle \sim \sigma$ is too small;
- larger $x \sim$ upwards slope $\rightarrow \langle k_T \rangle \sim \sigma$ is too large.

We acknowledge that this is a naive approach. Yet it is motivated by the argument that all the vectorial addition, such that the direction of the photon in the hardest process, are all in random angles and statistically should somewhat average out.

Asymptotic ansatz. With this we motivate an ansatz that decreases the width with increasing x, and has the boundary condition that $\sigma(Q)$ in its default implementation is matched for large Q. We implement this with replacing eq. (5.11) by

$$\sigma^*(Q, x) = \sigma_{\text{hard}} - \gamma \frac{\sqrt{(1+x)^2 - 1}}{Q}, \tag{5.12}$$

$$\gamma = 2.2\,\text{GeV}. \tag{5.13}$$

The structure of the second term in eq. (5.12) enhances the x dependence in contrast to a linear model – the precise power can also be considered as a parameter. This enhances sensitivity to the small absolute value changes at the small x regions while also apply to larger x, where in this study we are limited to the region $0.032 < x < 0.4$. The prefactor γ controls the variation and is in units of energy; the precise value is determined by several tries, but is not obtained by a involved fine tuning procedure, or a fit. The value of γ being larger than the limit for large Q is not an issue, since the x values too are small and thus balanced.

The impact of the changes we propose can be seen in figure 29, where we show several groups of kinematical bins in (Q^2, x, z). In each of these groups we fix two bins and vary the third, the multiplicities are then shown in their p_T^2 profile.

To judge the quality of the improvement it is necessary to specify the aspects we consider. Since the agreement with the data is very poor in general, we do not use a χ^2 test function here. Instead, we assess the qualitative description of the data. We consider the agreement to the data, but also the integrated matching – if the data are underestimated for small p_T values and overestimated for large p_T, this may be still an overall better agreement with the data than a constant over/underestimation of the data, even if the relative factor is closer to one. In these criteria, we see an improvement in various bins in the representatively chosen bins.

In the figure 29 we can see that for larger values in Q^2 we show here, we have no significant difference with the default `Pythia8` results. This is an important aspect as it shows we modify only the low Q region, and thus our modfification would not interfere with processes at large virtualities. The asymptotic model is improves the bins on the left in the first and last row, while being more aligned to the default description in general.

We also see, that the description worsens in some bins. Nevertheless we clearly see an impact of the modifications, and thus stress the potential of a continuation of this study by providing a fine tuned model with which we can provide a significant quantitative improvement on the description of the COMPASS data.

E Discussion

We have concluded a study of the agreement between MCEGs and experimental data relevant in the TMD regime of SIDIS. Considering the HERMES data we cannot reproduce the experimental data with either `Pythia8` or `Pythia6`, which are the herefore considered MCEGs, and we do not expect a significant difference in the description with the other MCEGs we used for the main comparison (`Cascade3` and `Herwig7`).

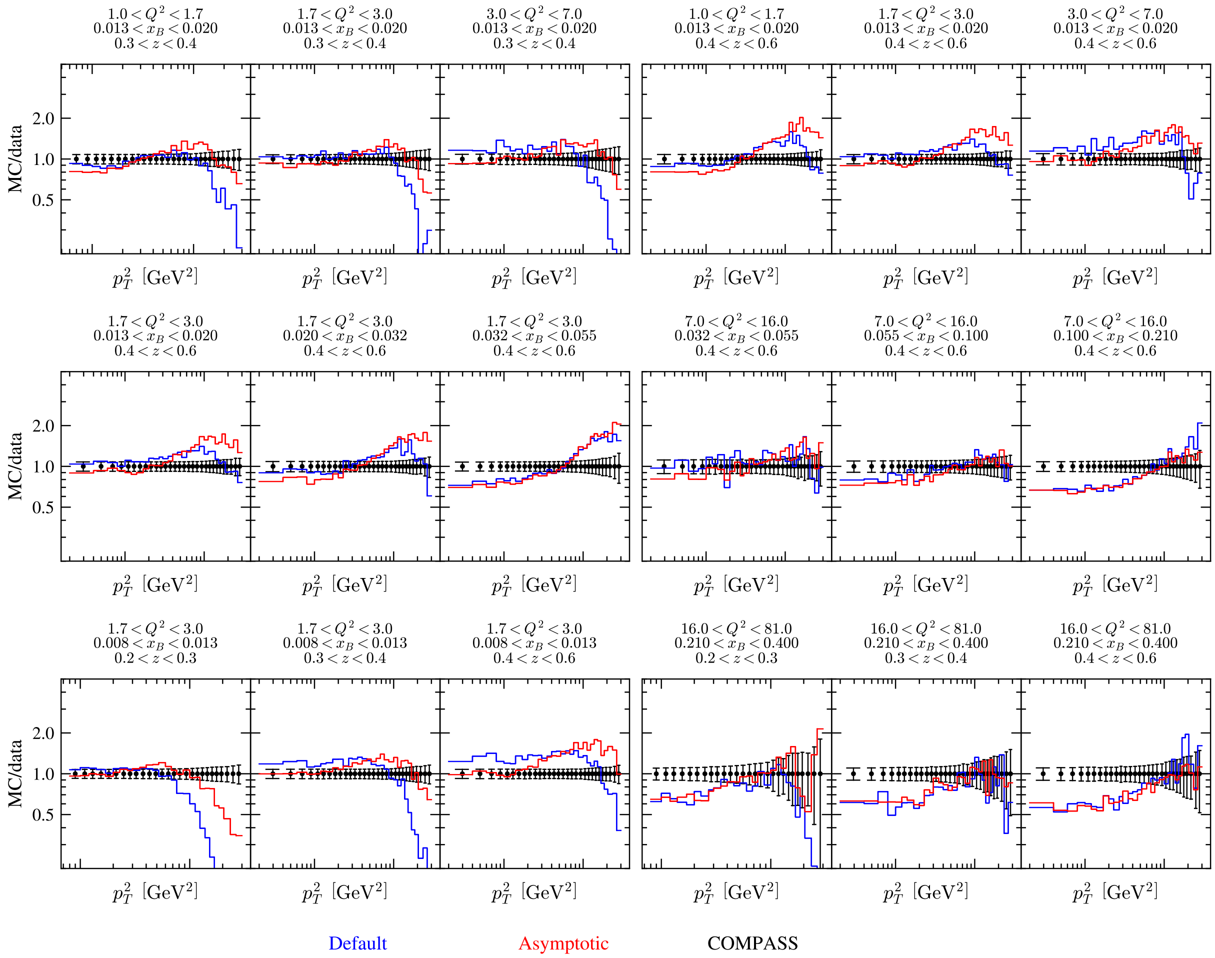


Figure 29: The default `Pythia8` description to the four-fold differential multiplicities compared to the two suggested modifications of *primordial* k_T in `Pythia8`. The bins are denoted over the respective figure. Each individual plot shows the p_T^2 dependence of the MCEG results and the data. The first panel shows two different groups of fixed x_B and z, where the Q^2 range is increased from left to right within a group. In the second and third panel, we demonstrate the same variation for variable x_B and variable z with other kinematic fixed, respectively.

For the data from COMPASS, we are able to reproduce the data qualitatively; the p_T dependent four fold multiplicities are however not being described from any of the above mentionded codes. The library `LEPTO` is suited the best for describing the SIDIS data we compare to. For `Pythia8` we demonstrate that the code can be tuned for a better description of the data. We however acknowledge that the mechanisms we study here are not the only parts in the code that are most relevant for the observables we are interested in. The changes to the `Pythia8` intrinsic algorithm we describe show that we can achive a much better agreement with data in many bins, but at the same time we also obtain a worse description in other bins. A modification which leads to an overall improvement has not been found. At the same time, these modifications need to be cross checked to guarantee that the performance of the program is not influenced in other kinematical regions and or processes. We stress the shape of the modification, which turns to the default implementation for large Q. At the vpercent for Q=10.

A tune of the MCEG, however, cannot be tested by only working in one rather extreme outlier of the final state phase space. Furthermore there are many more factors and parts of the code which affect the final state p_T. We here list some of them which are linked to the final state transverse momentum in a direct way. All remaining partons of the hadron which do not interfere in the hard process itself are labelled as *beam remnants*. They too have an undefined component of transverse momentum $k_{r,T}$, which is generated obeying a Gaussian distribution, like for the hard parton. In `Pythia8`the implementation generates the $k_{r,T}$ at hadronic scale according to the width parameter σ eq. (5.11) for all remnants the width of this guassian is set to a constant parameter `primordialKTremnant`$= 0.4$.

The total transverse momentum of the hadron is the sum of the transverse momenta of its constituents, thus there is a constraint on the sum of the transverse momenta like there is on the sum of the collinear components. To implement this constraint, `Pythia8`regulates the transverse momenta of all partons, including the hard parton. This is how also the generation of $k_{r,T}$ for *beam remnants* affects the initial state k_T for the hard parton, and thus indirectly also the final state p_T.

The feedback between the hard and the remaining quarks could be reduced by changing this routine. One option would be to demand the remnant partons have to compensate for the hard parton, leaving the latter unchanged. Alternatively, the total transverse momentum would be - on average - zero, if all partons shared the same transverse momentum distribution via the same Gaussian width. Thus, the feedback still exists, but possibly would disappear with growing statistics.

Because of the convolution of initial and final state effects, the SIDIS p_T dependent multiplicity is difficult to study. Further investigation of whether a tuning of primordial k_T influenced by TMD phenomenology is helpful should be studied taking into account the DY process as well. The parameters then should be extracted by a tuning package, such as `Professor` [102], unless one can guarantee that the implemented modification does not have a strong influence on other applications of `Pythia8`.

VI Conclusions

In this study, we discuss the frontier of the TMD framework on the formal and the phenomenological side. The computation of one-loop corrections to TMD operators for both gluon and quark distributions has been accomplished within the scope of this thesis. The loop correction for the twist (2+1) gluon operator has been finished only recently, and therefore we make it public for the first time in this work. To our best knowledge, we are the first to derive the results for these diagrams. At this point it is still unclear what can be said about the impact of these results in terms of phenomenology; TMD distributions of gluons are not tackled in extractions, as they do not appear at the leading power in the factorisation approach to the cross section.

In this work we have provided an overview of the framework of TMD extractions, and explicitly provide a state of the art extraction of TMD distributions. With the concluded work ART23, we have demonstrated the benefit of a flavour dependent ansatz for TMDPDFs in the standard method of collinear PDF matching. The success of this ansatz is based on the simple parametrization of a non perturbative function; flavour separation reduces the complexity of the models, and thus the total number of free parameters is not larger than for studies without flavour separation in our model. ART23 provides an outstanding description of the Drell-Yan production in the TMD factorization framework. The quality of the fit is marked by the distribution of the χ^2 replicas, with a consistent treatment of the sources of uncertainties (experimental and theoretical) via simultaneous variation of data and PDFs.

In this thesis, we have presented two further studies, building on the developments in ART23. The first of which consists of combining the flavour dependent ansatz in TMDPDFs with a simultaneous fit of TMDFFs for the pion and kaon. In this combined fit, we provide a state of the art analysis at the highest perturbative accuracy available on TMD related DY and SIDIS measurements. We are able to simultaneously describe data for both processes with very good fit quality, and extract the TMDFFs, as well as another version of TMDPDFs for protons and the RAD. The most crucial aspect we point out here is the adjustment of ART23 by treating of quarks and their antiquarks equally for small momentum fractions x. This modification alone empowers our framework to simultaneously achieve the fit to SIDIS and DY pair production by pion-nucleon collosion with very good quality, which is the second study that is founded on the progress made in ART23. In both of these extractions the NP ansatzes retain a simple functional form that requires only a few parameters, that is, for the TMDFFs we retain the total amount of parameters of the SV19 fit even though we distinguish between pions and kaons, and the pion TMDPDFs with the same functional form as the proton TMDPDFs provides a good description of the data available.

In both of these newly added fits, theoretical extractions can be improved by new constraints from experiments. For the pion DY fit, the COMPASS collaboration is going to provide new data and hence the kinematical range of the data available is increased. With this, the theory prediction can be tested, most importantly however, it can also be improved. Currently we fit to only one set of data due to the availability theirof. This is a limiting factor in phenomenology regarding the pion, which is expected to be removed by the announced release of data measured by the COMPASS collaboration. The fits to SIDIS data likewise should improve with data from the EIC, since for now both experiments used in TMDFFs determinations are fixed target experiments at small centre-of-mass energies.

We stress that the results of a combined fit to DY and SIDIS data are presented are not yet public. They belong to ongoing research projects and further analysis might lead to (minor) adjustments before the results are published. For the pion TMDPDFs we can provide a good fit, however, we stress that in order to provide a complete analysis of the uncertainties provided, the uncertainties from the proton TMDPDFs should also be taken into account.

As we argued, knowledge from TMD phenomenology can be used to improve MCEGs. We put these considerations to the test by looking specifically at the `Pythia8` library. We provide an overview of the default performance of several MCEGs for the data from HERMES and COMPASS. We discuss concrete ideas that improve the agreement between Pythia8 and the data measured by the COMPASS collaboration on a qualitative level. Due to the design of these modifications, we expect a benefit only in the region of small hard scales Q, such that the update we work on could eventually be implemented in Pythia8 without interference with its performance on large Q processes, namely for applications for hadron-hadron collisions. The first implementation of these ideas demonstrates the potential of the ansatz. Nevertheless we stress that this work is incomplete. Before a final suggestion of a tune for Pythia8 can be made, a more extensive analysis of all processes is needed.

Appendix A Diagram Results

In this appendix we collect the remaining results of chapter III diagram by diagram.
The diagrams are

$$A \qquad \frac{-g^2 C_A}{8\pi^{\frac{d}{2}}\epsilon}\int_0^1 \mathrm{d}u\frac{u}{\overline{u}}\left(\mathcal{X}^A_{\mu,\nu}\left(z_{21}^u,z_2\right)-\mathcal{X}^A_{\mu,\nu}\left(z_1,z_2\right)\right) \tag{A.1}$$

$$A \qquad \frac{-g^2 C_A}{16\pi^{\frac{d}{2}}\epsilon}\left[\left(1-\ln\left(\frac{-\mathrm{i}p_+^{z_1}}{\delta_L}\right)\right)\mathcal{X}^A_{\mu,\nu}\left(z_1,z_2\right)-\int_0^1 \mathrm{d}u\frac{u}{\overline{u}}\left(\mathcal{X}^A_{\mu,\nu}\left(z_{21}^u,z_2\right)-\mathcal{X}^A_{\mu,\nu}\left(z_1,z_2\right)\right)\right] \tag{A.2}$$

$$B \qquad \frac{g^2 C_A}{16\pi^{\frac{d}{2}}\epsilon}\int_0^1 \mathrm{d}u\frac{u}{\overline{u}}\left(\mathcal{X}^A_{\mu,\nu}\left(z_1,z_{12}^u\right)-\mathcal{X}^A_{\mu,\nu}\left(z_1,z_2\right)\right) \tag{A.3}$$

$$B' \qquad \frac{-g^2 C_A}{16\pi^{\frac{d}{2}}\epsilon}\left(1-\ln\left(\frac{-\mathrm{i}p_+^{z_2}}{\delta_L}\right)\right)\mathcal{X}^A_{\mu,\nu}\left(z_1,z_2\right) \tag{A.4}$$

$$A2 \qquad \frac{-g^2 C_A}{8\pi^{\frac{d}{2}}\epsilon}\mathcal{X}^A_{\mu,\nu}\left(z_1,z_2\right) \tag{A.5}$$

$$B2 \qquad \frac{-g^2 C_A}{8\pi^{\frac{d}{2}}\epsilon}\mathcal{X}^A_{\mu,\nu}\left(z_1,z_2\right) \tag{A.6}$$

$$A* \qquad \frac{g^2 C_A}{16\pi^{\frac{d}{2}}\epsilon}\left[\int_L^{z_3}\frac{\mathrm{d}\sigma}{z_1-z_3}\left(\mathcal{X}^A_{\mu,\nu}\left(z_1,\sigma n\right)-\int_0^1 \mathrm{d}u\mathcal{X}^A_{\mu,\nu}\left(z_{31}^u,\sigma n\right)\right)\right] \tag{A.7}$$

$$B* \qquad \frac{-g^2 C_A}{16\pi^{\frac{d}{2}}\epsilon}\left[\int_{z_3}^{z_1}\mathrm{d}\sigma\mathcal{X}^A_{\mu,\nu}\left(\sigma n,z_3 n\right)-\int_L^{z_3}\mathrm{d}\sigma\mathcal{X}^A_{\nu,\mu}\left(z_3 n,\sigma n\right)\right. \tag{A.8}$$

$$\left.\int_0^1\frac{\mathrm{d}u}{z_3-z_1}\left(\int_{z_13^u}^{z_1}\mathrm{d}\sigma\mathcal{X}^A_{\mu,\nu}\left(\sigma n,z_{13}^u n\right)-\int_L^{z_{13}^u}\mathrm{d}\sigma\mathcal{X}^A_{\nu,\mu}\left(z_{13}^u,\sigma n\right)\right)\right] \tag{A.9}$$

$$D \qquad \frac{-g^2 C_A}{4\pi^2\varepsilon}\int_0^1 \mathrm{d}u(\overline{u}-u)\int_L^u \mathrm{d}w\left(\mathcal{X}^E_{\mu,\nu}\left(z_{21}^u n,z_{21}^w n\right)-\mathcal{X}^E_{\nu,\mu}\left(z_{21}^u n,z_{21}^w n\right)\right) \tag{A.10}$$

$$C \qquad \int[\mathrm{d}\alpha\mathrm{d}\beta\mathrm{d}\gamma]\,\frac{-\mathrm{i}g^2 C_A f^{A'D'E}\chi\psi}{32\pi^2\epsilon\Lambda} \tag{A.11}$$

$$\left(g_{\mu\nu}A_\alpha^{A'}(z_3\psi' n+z1\overline{\psi}' n)A_\alpha^{D'}(z_1\chi' n+z_3\overline{\chi}' n)\left\{\mathcal{F}_1\partial^2_{\psi+}+\mathcal{F}_2\partial^2_{\chi+}+\mathcal{F}_3\partial_{\psi+}\partial_{\chi+}\right\}\right. \tag{A.12}$$

$$+A_\nu^{A'}(z_3\psi' n+z1\overline{\psi}' n)A_\mu^{D'}(z_1\chi' n+z_3\overline{\chi}' n)\left\{\mathcal{G}_1\partial^2_{\psi+}+\mathcal{G}_2\partial^2_{\chi+}+\mathcal{G}_3\partial_{\psi+}\partial_{\chi+}\right\} \tag{A.13}$$

$$+A_\mu^{A'}(z_3\psi' n+z1\overline{\psi}' n)A_\nu^{D'}(z_1\chi' n+z_3\overline{\chi}' n)\left\{\mathcal{H}_1\partial^2_{\psi+}+\mathcal{H}_2\partial^2_{\chi+}+\mathcal{H}_3\partial_{\psi+}\partial_{\chi+}\right\} \tag{A.14}$$

The short hand notation $z_{ij}^w = z_i + w(z_j - z_i)$ and p_+^x denotes which argument the momentum operator acts upon. Other diagrams do not contribute because of their Lorentz structure. In eq. (3.1) the vector n is light like and the Lorentz indices of the field strength tensor in the TMD operator are transverse to n. The consequence is that

$$A1 = A1' = B1 = B1' = E1 = E2 = E3 = 0 \tag{A.15}$$

The structure functions appearing in diagram C are listed here.

$$\mathcal{F}_1 = f_1 + F_1 \tag{A.16}$$

$$\mathcal{F}_3 = f_3 + F_3 \tag{A.17}$$

$$\mathcal{G}_1 = g_1 + G_1 \tag{A.18}$$

$$\mathcal{G}_3 = g_3 + G_3 \tag{A.19}$$

$$\mathcal{H}_1 = h_1 + J_{\delta_2^2} \tag{A.20}$$

$$\mathcal{H}_3 = h_3 + J_{\delta_1^1 \delta_2^1} \tag{A.21}$$

$$\mathcal{H}_4 = z_1 J^{z_1}_{\delta_2^3} + z_2 J^{z_2}_{\delta_2^3} \tag{A.22}$$

$$\mathcal{H}_6 = z_1 J^{z_1}_{\delta_1^1 \delta_2^2} + z_2 J^{z_2}_{\delta_1^1 \delta_2^2} \tag{A.23}$$

where

$$f_1 = \frac{5}{2\Lambda^3}\left(\varphi^3 + \varphi^2\psi\right) \tag{A.24}$$

$$f_2 = \frac{5}{2\Lambda^3}\left(\varphi^3 + \varphi^2\chi\right) \tag{A.25}$$

$$f_3 = \frac{5}{2\Lambda^3}\left(2\varphi^3 + \varphi^2(\chi + \psi) + \varphi\chi\psi\right) \tag{A.26}$$

$$g_1 = -\frac{2}{\Lambda^3}\left(\varphi^3 + \varphi^2\psi\right) \tag{A.27}$$

$$g_2 = -\frac{2}{\Lambda^3}\left(\varphi^3 + \varphi^2\chi\right) \tag{A.28}$$

$$g_3 = -\frac{2}{\Lambda^3}\left(2\varphi^3 + \varphi^2(\chi + \psi) + \varphi\chi\psi\right) \tag{A.29}$$

$$h_1 = \frac{3}{\Lambda^3}\left(2\varphi^3 - 2\varphi^2\chi - 4\varphi\chi\psi - \varphi\psi^2 - 2\chi\psi^2\right) \tag{A.30}$$

$$h_2 = \frac{3}{\Lambda^3}\left(2\varphi^3 - 2\varphi^2\psi - 4\varphi\chi\psi - \varphi\chi^2 - 2\psi\chi^2\right) \tag{A.31}$$

$$h_3 = \frac{6}{\Lambda^3}\left(2\varphi^3 - \varphi^2(\chi + \psi) - 2\varphi\chi\psi\right) \tag{A.32}$$

and from $3d/2+1$

$$F_1 = -\frac{\varphi^2}{\Lambda^4}\left(5\varphi^2(\chi+\psi)+\varphi(6\chi\psi+5\psi^2)\right) \tag{A.33}$$

$$F_2 = -\frac{\varphi^2}{\Lambda^4}\left(5\varphi^2(\chi+\psi)+\varphi(6\chi\psi+5\chi^2)\right) \tag{A.34}$$

$$F_3 = -\frac{\varphi}{\Lambda^4}\left(10\varphi^3(\chi+\psi)+\varphi^2(\chi^2+4\chi\psi+\psi^2)-2\varphi(\chi^2\psi+\chi\psi^2)-3\chi^2\psi^2\right) \tag{A.35}$$

$$G_1 = \frac{4\varphi^3}{\Lambda^4}\left(\varphi(\chi+\psi)+\chi\psi\right) \tag{A.36}$$

$$G_2 = \frac{4\varphi^3}{\Lambda^4}\left(\varphi(\chi+\psi)+\chi\psi\right) \tag{A.37}$$

$$G_3 = \frac{4\varphi}{\Lambda^4}\left(2\varphi^3(\chi+\psi)+2\varphi^2\chi\psi-\varphi^1(\chi^2\psi+\chi\psi^2)-\chi^2\psi^2\right) \tag{A.38}$$

$$H_1 = \frac{2}{\Lambda^4}\Bigg(-6\varphi^4(\chi+\psi)+\varphi^3(\chi^2+8\chi\psi-3\psi^2)+\varphi^2(5\chi^2\psi+4\chi\psi^2+3\psi^3) \tag{A.39}$$

$$+\varphi(7\chi^2\psi^2+6\chi\psi^3)+3\chi^2\psi^3\Bigg) \tag{A.40}$$

$$H_2 = \frac{2}{\Lambda^3}\left(-6\varphi^3+\varphi^2(\psi-3\chi 2)+\varphi(3\chi^2+4\chi\psi)+7\chi^2\psi\right) \tag{A.41}$$

$$H_3 = \frac{-4\varphi^2}{\Lambda^4}\Bigg(+16\varphi^2(\chi+\psi)+\varphi(16\chi\psi-10\Lambda)+\Lambda(\chi+\psi)\Bigg) \tag{A.42}$$

The structures with small font originate from $3d/2$, structures with capital Font from $3d/2+1$. The functions have the symmetry properties

$$f_1(\chi,\psi) = f_2(\psi,\chi) \qquad f_3(\chi,\psi) = f_3(\psi,\chi). \tag{A.43}$$

The above holds as well for F_i, g_i, G_i, h_i. The structures with index 3 are symmetric by themselves.

In order to compute the two d dim. integrals with one formula, we worked out the expressions for the appearing denominator structures for up to six terms appearing in the numerator. They are

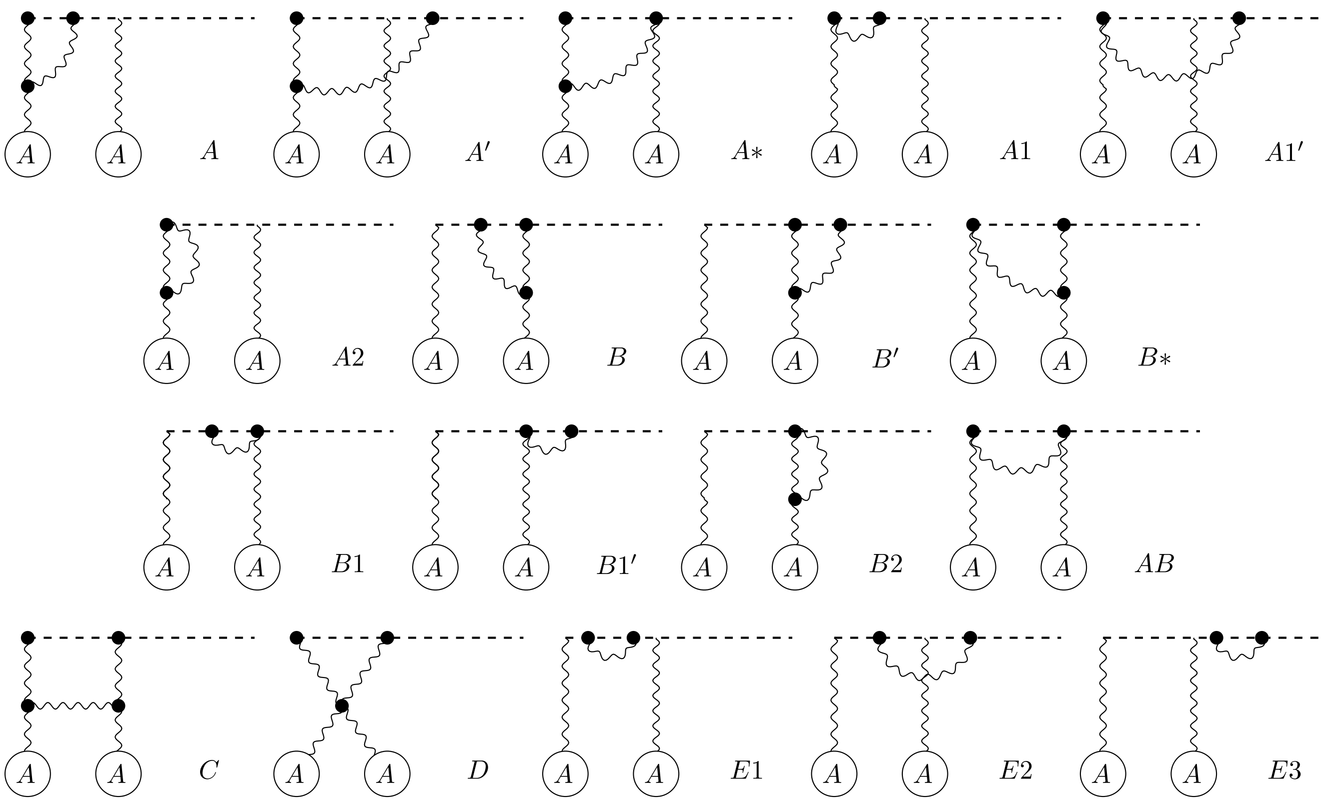


Figure 30: One loop corrections to the gluon TMD operator of twist 2. Black dots indicate the location of quantum fields (vertices), the encircled A represents the remaining gluon background field.

$$\int \mathrm{d}^d x \int \mathrm{d}^d y \frac{1}{(-ax^2 - by^2 + 2cxy + m^2)^n} = \frac{\mathrm{i}^2(-1)^0}{\pi^{-d}}(ab - c^2)^{-\frac{d}{2}} \frac{\Gamma(n-d)}{\Gamma(n)}(m^2)^{d-n}, \tag{A.44}$$

$$\int \mathrm{d}^d x \int \mathrm{d}^d y \frac{x^\mu x^\nu}{(-ax^2 - by^2 + 2cxy + m^2)^n} = \frac{\mathrm{i}^2(-1)^1}{\pi^{-d}} b^{-\frac{d}{2}}(a - \frac{c^2}{b})^{-\frac{d}{2}-1} \frac{g^{\mu\nu}}{2} \frac{\Gamma(n-d-1)}{\Gamma(n)}(m^2)^{d-n+1}, \tag{A.45}$$

$$\int \mathrm{d}^d x \int \mathrm{d}^d y \frac{x^\mu y^\nu}{(-ax^2 - by^2 + 2cxy + m^2)^n} = \frac{\mathrm{i}^2(-1)^1}{\pi^{-d}} c(ab - c^2)^{-\frac{d}{2}-1} \frac{g^{\mu\nu}}{2} \frac{\Gamma(n-d-1)}{\Gamma(n)}(m^2)^{d-n+1}, \tag{A.46}$$

$$\int \mathrm{d}^d x \int \mathrm{d}^d y \frac{x^\mu x^\nu x^\theta x^\varphi}{(-ax^2 - by^2 + 2cxy + m^2)^n} = \frac{\mathrm{i}^2(-1)^2}{\pi^{-d}} b^{-\frac{d}{2}}(a - \frac{c^2}{b})^{-\frac{d}{2}-2} \frac{E^{\mu\nu\theta\varphi}}{4} \frac{\Gamma(n-d-2)}{\Gamma(n)}(m^2)^{d-n+2}, \tag{A.47}$$

$$\int \mathrm{d}^d x \int \mathrm{d}^d y \frac{x^\mu x^\nu x^\theta y^\varphi}{(-ax^2 - by^2 + 2cxy + m^2)^n} = \frac{\mathrm{i}^2(-1)^2}{\pi^{-d}} c b^{-\frac{d}{2}-1}(a - \frac{c^2}{b})^{-\frac{d}{2}-2} \frac{E^{\mu\nu\theta\varphi}}{4} \frac{\Gamma(n-d-2)}{\Gamma(n)}(m^2)^{d-n+2}, \tag{A.48}$$

$$\int \mathrm{d}^d x \int \mathrm{d}^d y \frac{x^\mu x^\nu y^\theta y^\varphi}{(-ax^2 - by^2 + 2cxy + m^2)^n} = \frac{\mathrm{i}^2(-1)^2}{\pi^{-d}}(ab - c^2)^{-\frac{d}{2}-2} \left(c^2 \frac{E^{\mu\nu\theta\varphi}}{4} + (ab - c^2)\frac{g^{\mu\nu} g^{\theta\varphi}}{4}\right) \tag{A.49}$$

$$\frac{\Gamma(n-d-2)}{\Gamma(n)}(m^2)^{d-n+2}, \tag{A.50}$$

$$\int \mathrm{d}^d x \int \mathrm{d}^d y \frac{x^\mu x^\nu x^\xi x^o x^\rho x^\sigma}{(-ax^2 - by^2 + 2cxy + m^2)^n} = \frac{\mathrm{i}^2(-1)^3}{\pi^{-d}} b^3(ab - c^2)^{-\frac{d}{2}-3} \frac{E_6^{\mu\nu\xi o\rho\sigma}}{8} \frac{\Gamma(n-d-3)}{\Gamma(n)}(m^2)^{d-n+3}, \tag{A.51}$$

$$\int \mathrm{d}^d x \int \mathrm{d}^d y \frac{x^\mu x^\nu x^\xi x^o x^\rho y^\sigma}{(-ax^2 - by^2 + 2cxy + m^2)^n} = \frac{\mathrm{i}^2(-1)^3}{\pi^{-d}} c b^2(ab - c^2)^{-\frac{d}{2}-3} \frac{E_6^{\mu\nu\xi o\rho\sigma}}{8} \frac{\Gamma(n-d-3)}{\Gamma(n)}(m^2)^{d-n+3}, \tag{A.52}$$

$$\int \mathrm{d}^d x \int \mathrm{d}^d y \frac{x^\mu x^\nu x^\xi x^o y^\rho y^\sigma}{(-ax^2 - by^2 + 2cxy + m^2)^n} = \frac{\mathrm{i}^2(-1)^3}{\pi^{-d}} b(ab - c^2)^{-\frac{d}{2}-3} \left(c^2 \frac{E_6^{\mu\nu\xi o\rho\sigma}}{8} + (ab - c^2)\frac{E_4^{\mu\nu\xi o} g^{\rho\sigma}}{8}\right) \tag{A.53}$$

$$\frac{\Gamma(n-d-3)}{\Gamma(n)}(m^2)^{d-n+3}, \tag{A.54}$$

$$\int \mathrm{d}^d x \int \mathrm{d}^d y \frac{x^\mu x^\nu x^\xi y^o y^\rho y^\sigma}{(-ax^2 - by^2 + 2cxy + m^2)^n} = \frac{\mathrm{i}^2(-1)^3}{\pi^{-d}} c(ab - c^2)^{-\frac{d}{2}-3} \tag{A.55}$$

$$\left(c^2 \frac{E_6^{\mu\nu\xi o\rho\sigma}}{8} + (ab - c^2)\left\{\frac{E_4^{\mu\nu\xi o} g^{\rho\sigma}}{8} + [o \leftrightarrow \rho] + [o \leftrightarrow \sigma]\right\}\right) \frac{\Gamma(n-d-3)}{\Gamma(n)}(m^2)^{d-n+3}. \tag{A.56}$$

The parameter m is inserted as an infrared regulator and can be set to zero in the scenario studied in chapter III.

Appendix B Additional material for ART23 extraction

In this appendix we provide additional material to chapter IV, namely on the TMDPDF determination by a fit to DY data in section B.

A Run of statistical fit

To obtain a reliable prediction with uncertainties, the computation needs to be repeated over the ensemble of parameter sets, each of which are obtained from the best fit to a replica of data and PDF. With this, the uncertainties of the experimental data and the collinear PDF are progressed and together contribute to the uncertainties provided for the extracted TMDPDFs and the RAD.

It is explained in [32] in greater detail how the replicas of experimental data are generated. Their generation is implemented in `artemide-Dataprocessor` directly. To generate replicas of PDFs (or FFs), we follow the instructions given in the literature [103] to convert PDFs (FFs) with Hessian uncertainty information to a set of Monte Carlo replicas. If they are already presented in such a form, we use them as provided. For this work we generated Monte Carlo replicas for the DSS kaon FFs. The replicas for `MSHTPDF20` are already used in [10] and for the pion FFs we use the official replicas of the DSS collaboration; also for the pion TMDPDF in section IV.D we rely on the official replicas to `JAM21PionPDFnlo` from the JAM collaboration.

The statistical fit was run on the cluster of Regensburg University, *Athene*. Due to the structure of clusters the fit script can only run until a walltime is reached; in our case we have a runtime limit of 168 hours per individual fit. A mentionworthy, but most likely very small effect of this is that replica-fits which converge very slowly could hit the wall time limit and thus disappear. The effect can be argued to be negligible, especially since extreme outliers are discarded. In the work of ART23 there is no issue with this. In the ongoing work of ART24, this problem had to be tackled due to a runtime of a single fit which, on average, was well beyond the wall time.

For the main productive run[44] 1186 replcia fits were generated. After discarding a few outliers[45], 1000 representative replicas are stored. The result of ART23 is computed from these. They are publicly available in `artemide-Dataprocessor` [74].

B Figures

In the following we provide additional figures of the main result of ART23. We list all the data plots for the extraction of ART23 and also a selection of figures of the obtained TMDPDFs. The plots in which we compare of theory and experiment figure 31 - 35. The visualized extraction of the TMDPDFs can be seen in figure 36 and 39. In all data plots the filled points are datapoints used in the extraction. The empty datapoints are non included in the fit. In the upper panel there is the absolute value of the cross section, in the lower panel the values are normalized to the average theory prediction for a better visualization of the uncertainties.

C Extraction with NNPDF3.1

We compare our extraction to a repetition in which we solely change the input collinear PDF to `NNPDF3.1`. To this end, we here compare in the resulting distribution of replica fit in terms of χ^2, the RAD and a selection of TMDPDFs in figures. We also provide the numbers for the parameters with the extraction with `NNPDF3.1`. They

[44]For comparison, e.g. usage `NNPDF3.1` instead of `MSHT20PDF`, we rely on a smaller amount of replica fit. These serve as a comparison only, and we find that several hundered replica are enough to obtain a stable and "smooth" distributions for the parameters.

[45]We discard fits converge to parameters that are far away from the common minimum. I.e. we constrain most of the parameters for the TMDPDF to be greater than 10^{-4}. The only exception is λ_2^{sea} which we impose to be smaller than 30. This is affects less than 0.5% of the replica-fits we obtained.

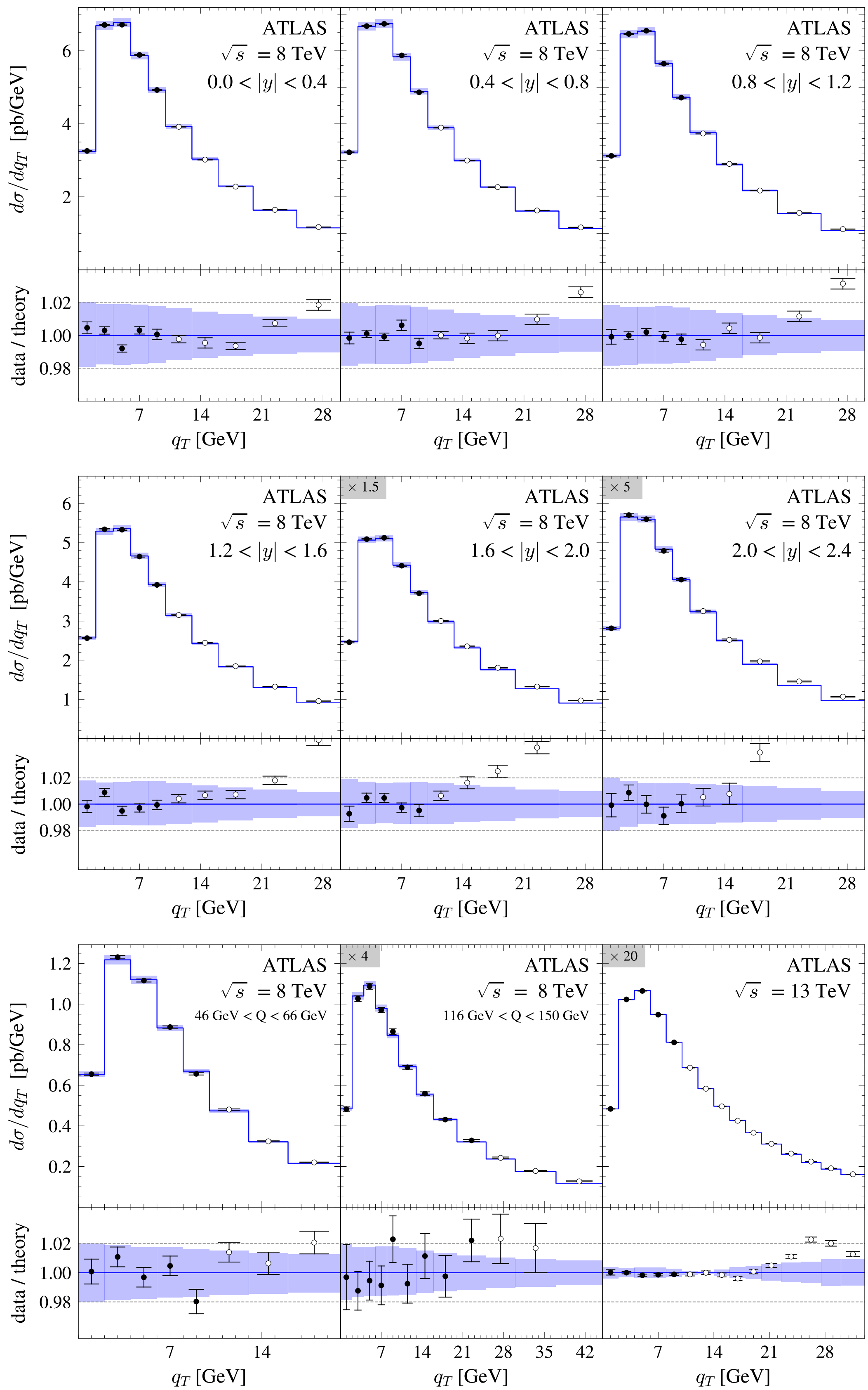


Figure 31: Differential cross section for Z/γ^* boson production measured by ATLAS [104] in different rapidity regions y.

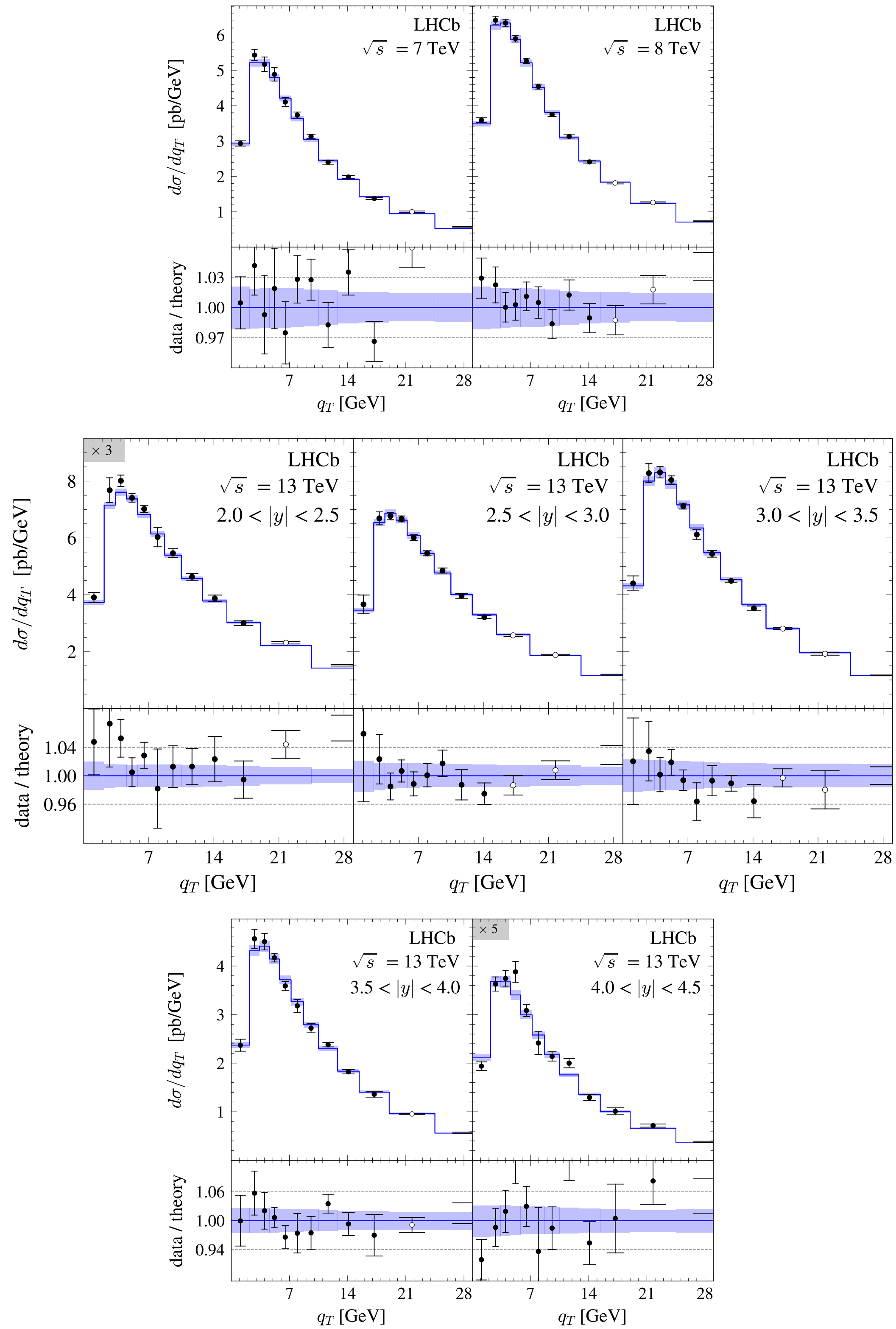


Figure 32: Differential cross section for Z/γ^* boson production measured by LHCb [105, 106, 107, 108] in different rapidity regions y.

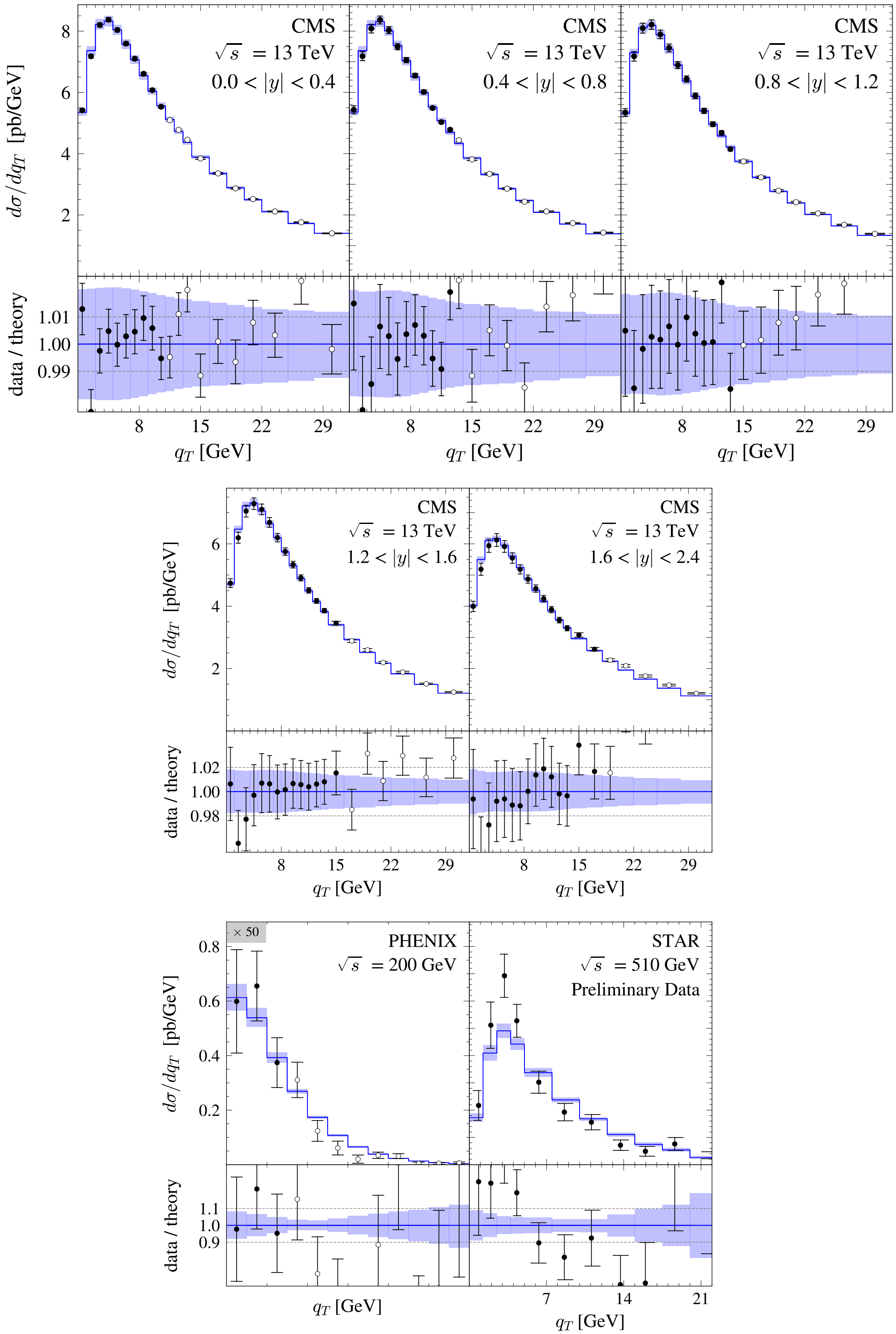


Figure 33: Differential cross section for Z/γ^* boson production measured by CMS (in different rapidity regions y), PHENIX [109] and STAR [110], as indicated in the plots.

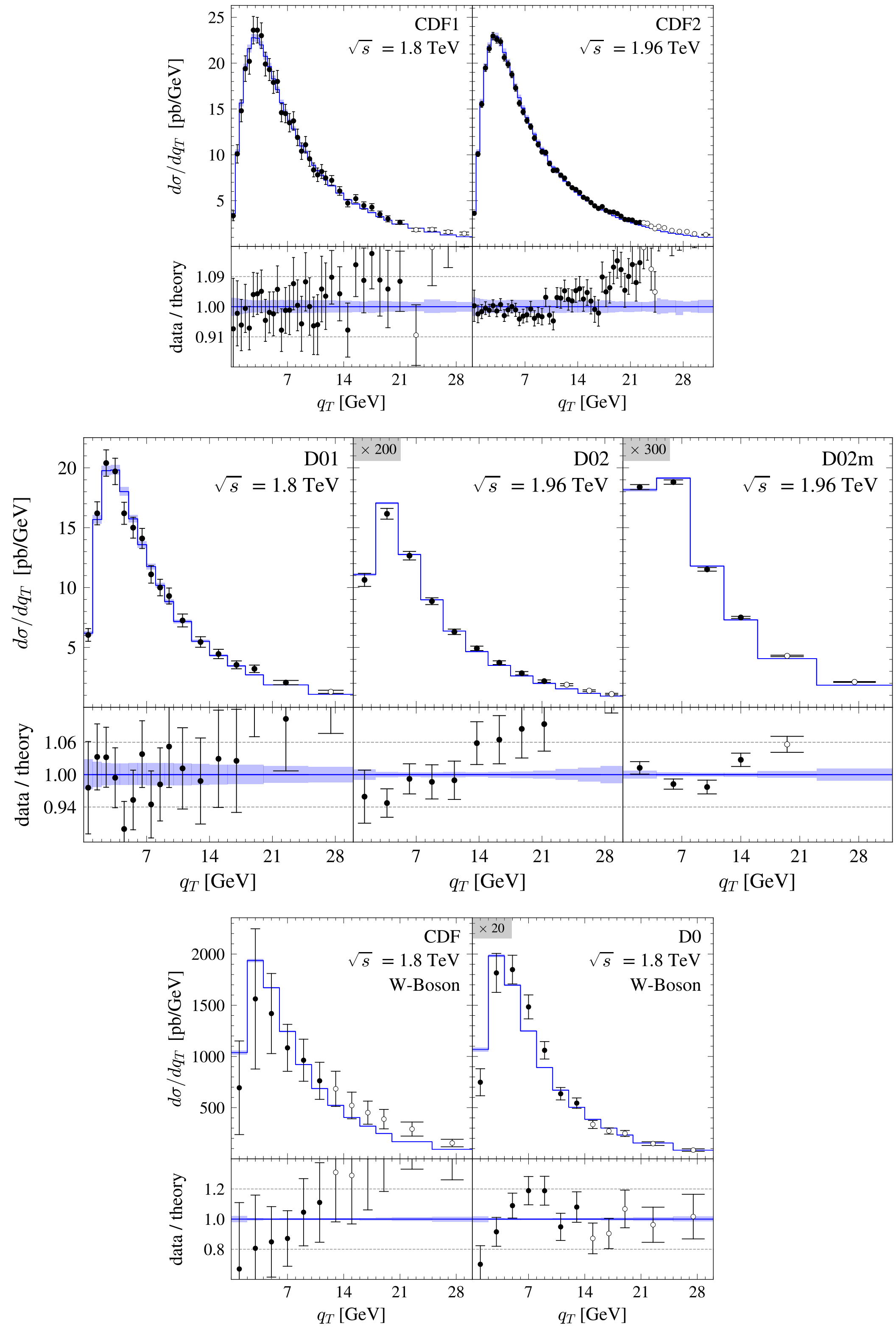


Figure 34: Differential cross section for Z/γ^* (upper rows) and W (lower row) boson production measured by CDF [111, 112, 113] and D0 [114, 115, 116, 117].

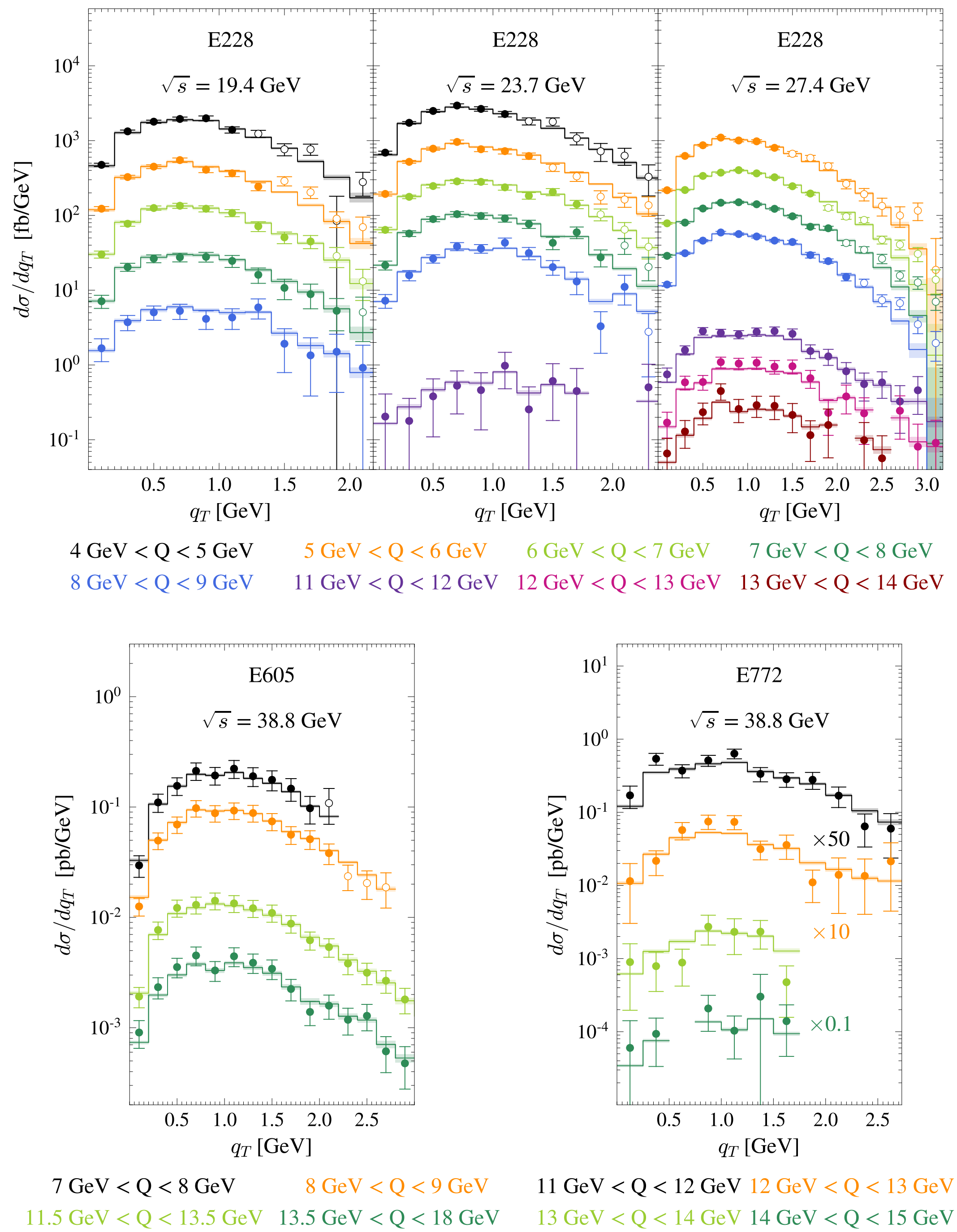


Figure 35: Differential cross section for the DY process at different values of Q measured by the experiments E228 [118], E772 [119] and E605 [120].

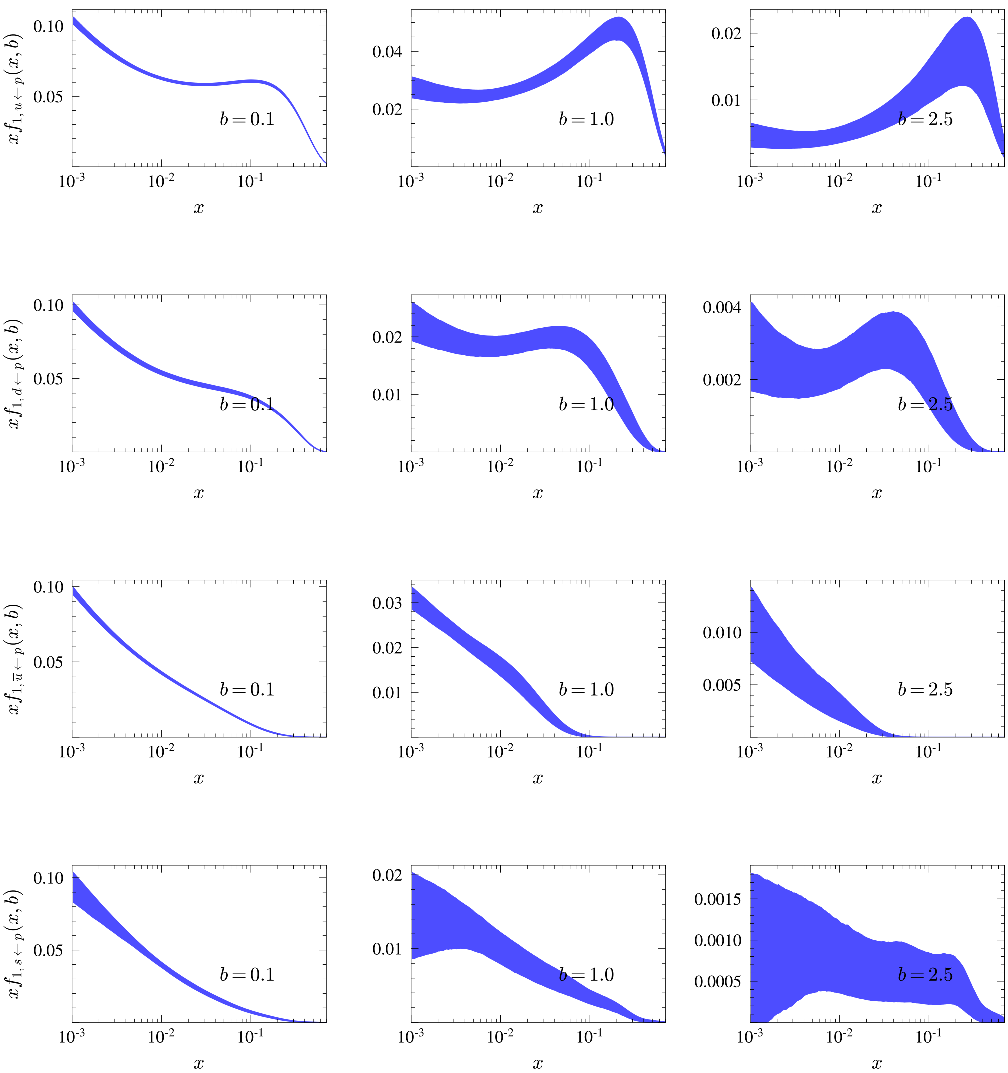


Figure 36: The TMDPDF obtained with uncertainty band computed at the optimal scale. The flavours are indicated in the label of the y axis. Recall that the optimal scale is $\boldsymbol{b}$ dependent.

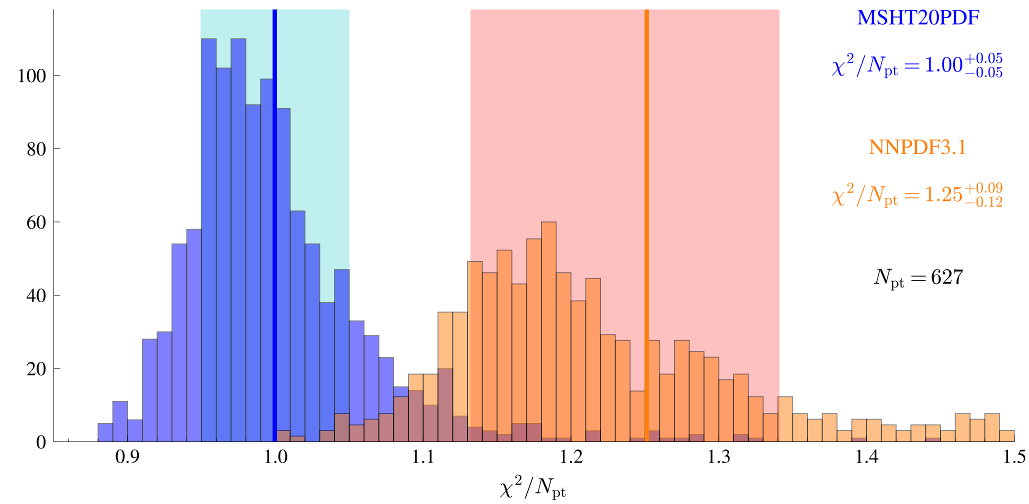


Figure 37: χ^2 distributions for replica fits using `MSHT20PDF` or `NNPDF3.1`. The `NNPDF3.1` distribution is normalized to the same total number of replica as `MSHT20PDF` for a better comparison. The average value and the 68% CI are given with a solid line and a rectangle in pale colour for both distributions, respectivly.

are

$$B_{\rm NP} = 1.82^{+0.37}_{-0.29}\ \text{GeV}, \qquad c_0 = 5.03^{+1.02}_{-0.81} \cdot 10^{-2}, \qquad c_1 = 1.04^{+0.09}_{-0.15} \cdot 10^{-1}. \tag{B.1}$$

for the RAD. These parametres are quite different from those obtained with MSHT20PDF, see eq. (4.20).

For the TMDPDFs the parameters are

$$\begin{aligned} \lambda_1^u &= 1.11^{+0.15}_{-0.17}, & \lambda_2^u &= 2.69^{+1.79}_{-2.53}, \\ \lambda_1^d &= 0.99^{+0.14}_{-0.17}, & \lambda_2^d &= 3.60^{+4.36}_{-3.57}, \\ \lambda_1^{\bar u} &= 0.37^{+0.16}_{-0.15}, & \lambda_2^{\bar u} &= 75.0^{+14.8}_{-13.2}, \\ \lambda_1^{\bar d} &= 0.37^{+0.30}_{-0.32}, & \lambda_2^{\bar d} &= 17.7^{+6.91}_{-14.2}, \\ \lambda_1^r &= 1.88^{+0.34}_{-0.29}, & \lambda_2^r &= 2.71^{+2.44}_{-2.69}. \end{aligned} \tag{B.2}$$

The distribution of χ^2 compared for the two different collinear PDFs is shown in figure 37. Clearly, the extraction with NNPDF3.1 has its average at a higher value of χ^2, but also the distribution is more smeared out. The RAD is shown in figure 38, the comparison of TMDPDFs in figure 39. For both we can see that there is mostly agreement within uncertainties. However, the NNPDF3.1 extraction has an overall larger statistical uncertainty. The small disagreement between both extractions in the RAD region from $\boldsymbol{b} \sim 0.5 - 0.9\,\text{GeV}^{-1}$ should be pointed out, as the RAD is a universal function and should be best constrained in the regime of small $\boldsymbol{b}$.

Appendix C Plots for SIDIS and FFs

In the following we provide additional figures to the description of the SIDIS data. The description of the positively charged hadron production for COMPASS data has already been shown in figure 18 and the plots for pion

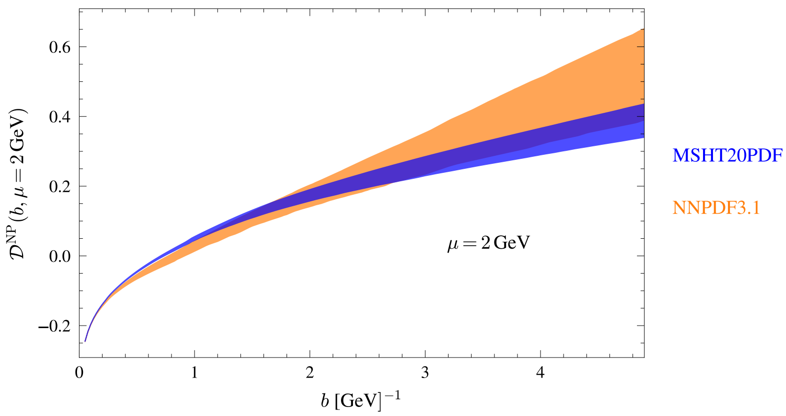


Figure 38: The RAD extracted with `MSHT20PDF` and `NNPDF3.1`.

production at HERMES is given in 19. The remaining figures are the the negatively charged hadron production for COMPASS (figure 40) and the kaon production (figure 41) listed in this section. In the figures filled points are used in the extraction, while empty datapoints are not included in the fit.

We provide figures for the extracted TMDFFs to illustrate both their shape and the respective uncertainties. They are displayed in figure 42 for the pion FFs and figure 43 for the kaon FFs.

Appendix D Data plots for the pion induced DY process

Here we list the remaining data plots for the pion-induced DY production, discussed in section D. Filled data points fullfill the kinematic constraints to be considered in the fit. Note that only the data from E615 have been used in the fit. The data presented in figure 45 have not been used in the fit.

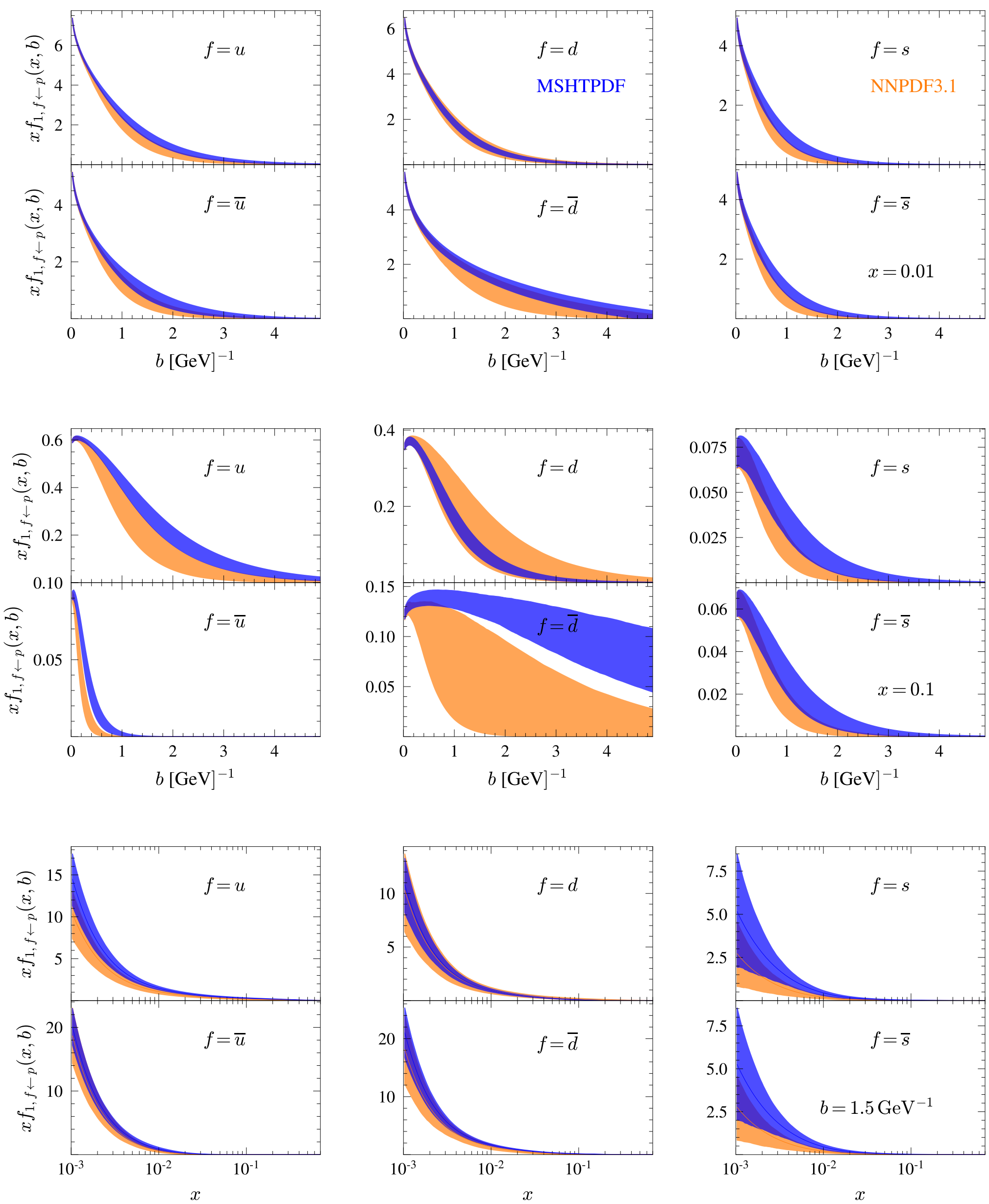


Figure 39: A comparison between resulting TMDPDFs with MSHT20PDF (blue) and NNPDF3.1 (orange). The upper two panels show the $\boldsymbol{b}$ dependence for different values of x, the lower panel shows the x dependence for fixed $\boldsymbol{b}$. Note that this figure depicts the TMDPDFs at the *optimal* scale, which itself is $\boldsymbol{b}$ dependent.

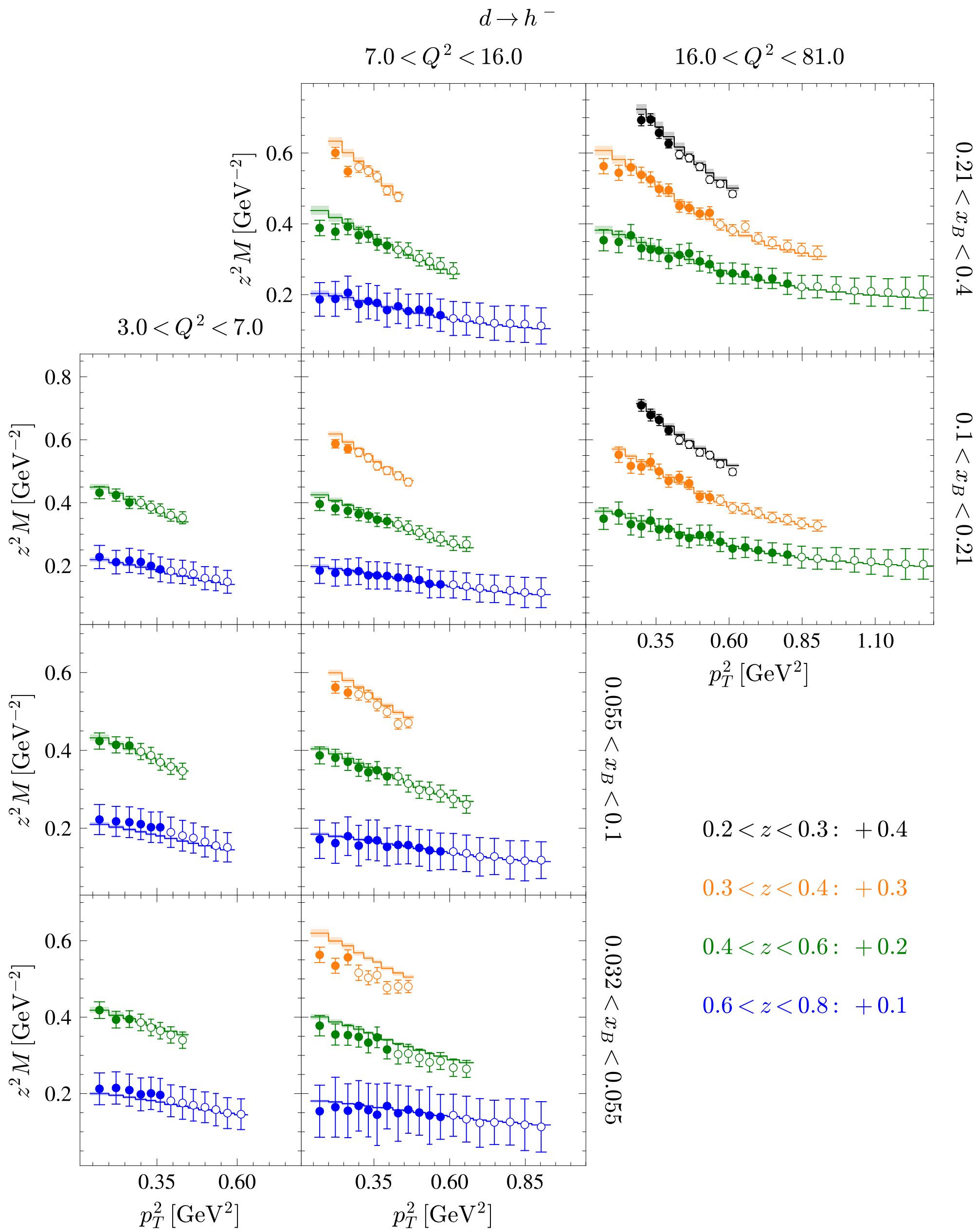


Figure 40: Theory prediction negatively charged hadron production in SIDIS for the COMPASS experiment. The data are given as multiplicity in bins of the variables Q, x, z, and the p_T profile is plotted. To show multiple kinematic regions of z we use an offset as indicated in the figure.

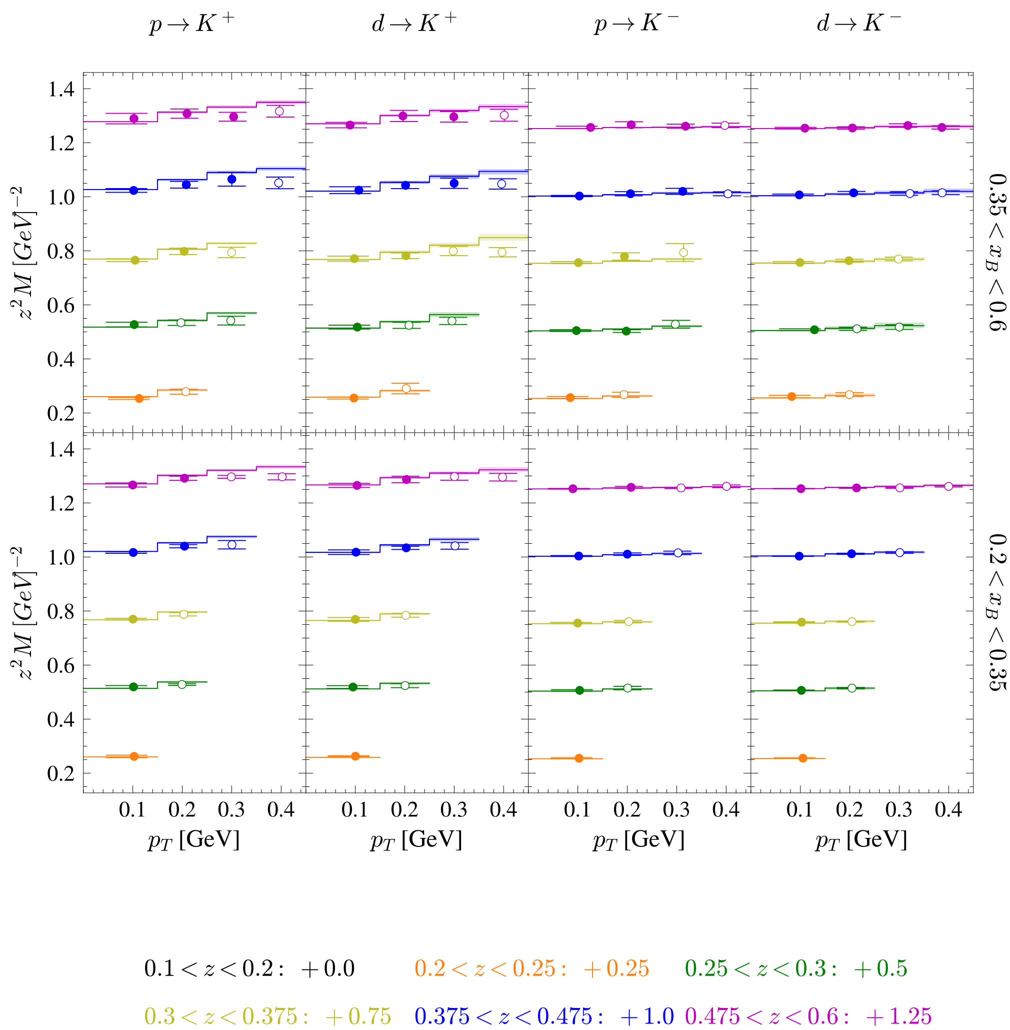


Figure 41: Theory prediction for charged kaon production in SIDIS for the HERMES experiment. The data are given as multiplicity in bins of the variables x, z, and the p_T profile is plotted. To show multiple kinematic regions of z we use an offset as indicated in the figure.

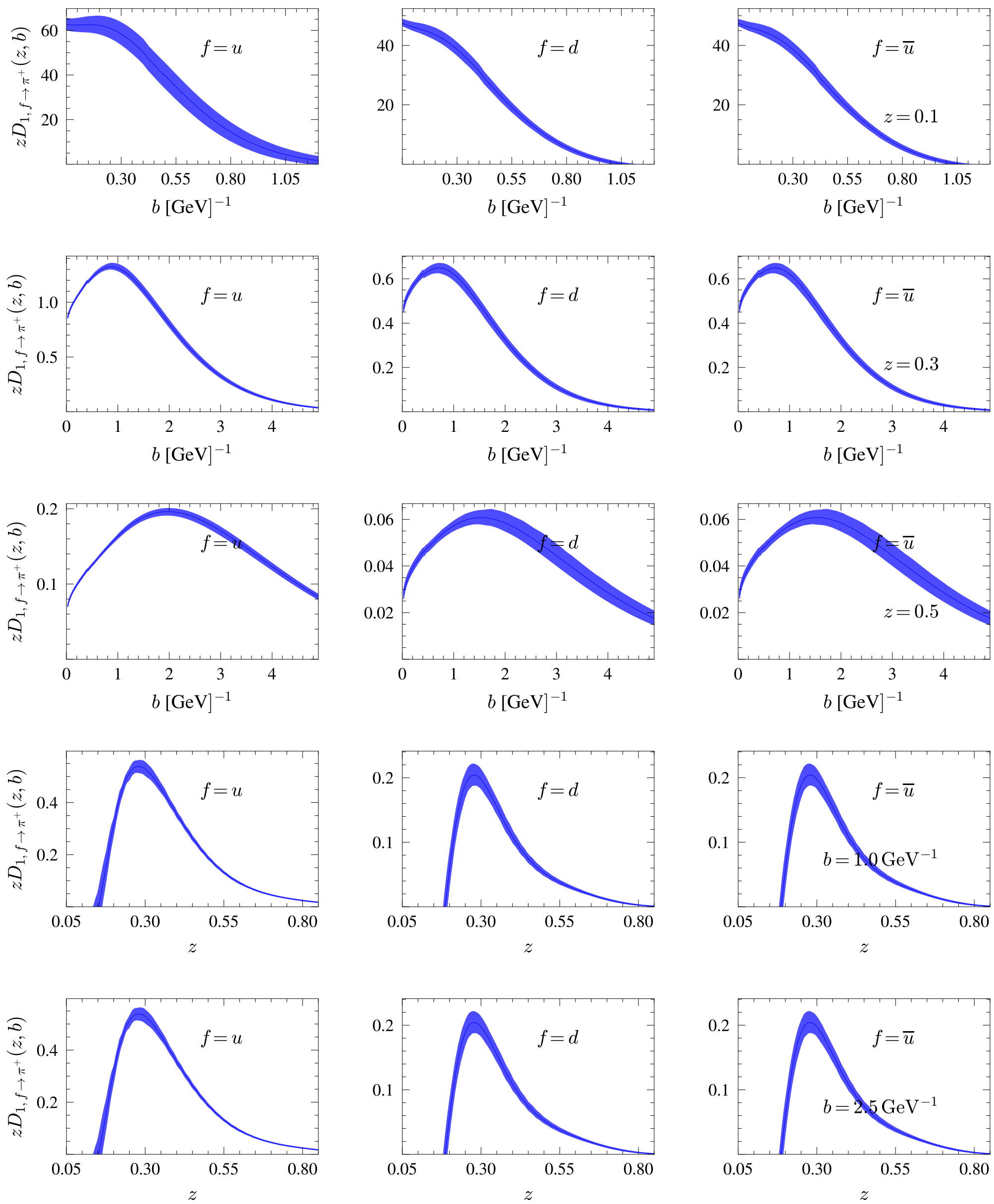


Figure 42: The TMDFF obtained with uncertainty band obtained at the optimal scale. In the upper panel we display different flavours for fixed z, the lower panel shows the TMDFFs for fixed b. Note that the optimal scale is $\boldsymbol{b}$ dependent.

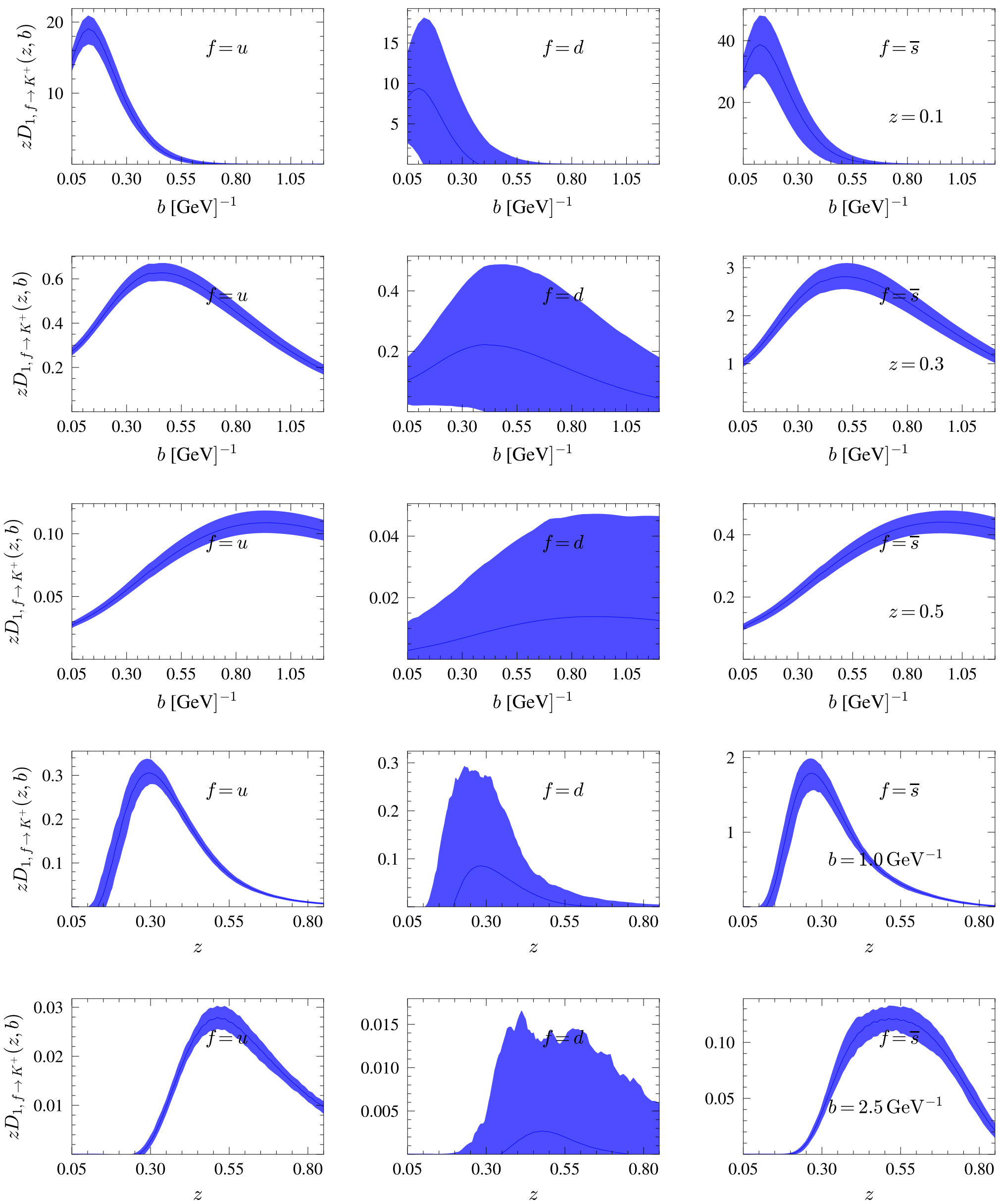


Figure 43: The TMDFF obtained with uncertainty band obtained at the optimal scale. In the upper panel we display different flavours for fixed z, the lower panel shows the TMDFFs for fixed b. Note that the optimal scale is $\boldsymbol{b}$ dependent.

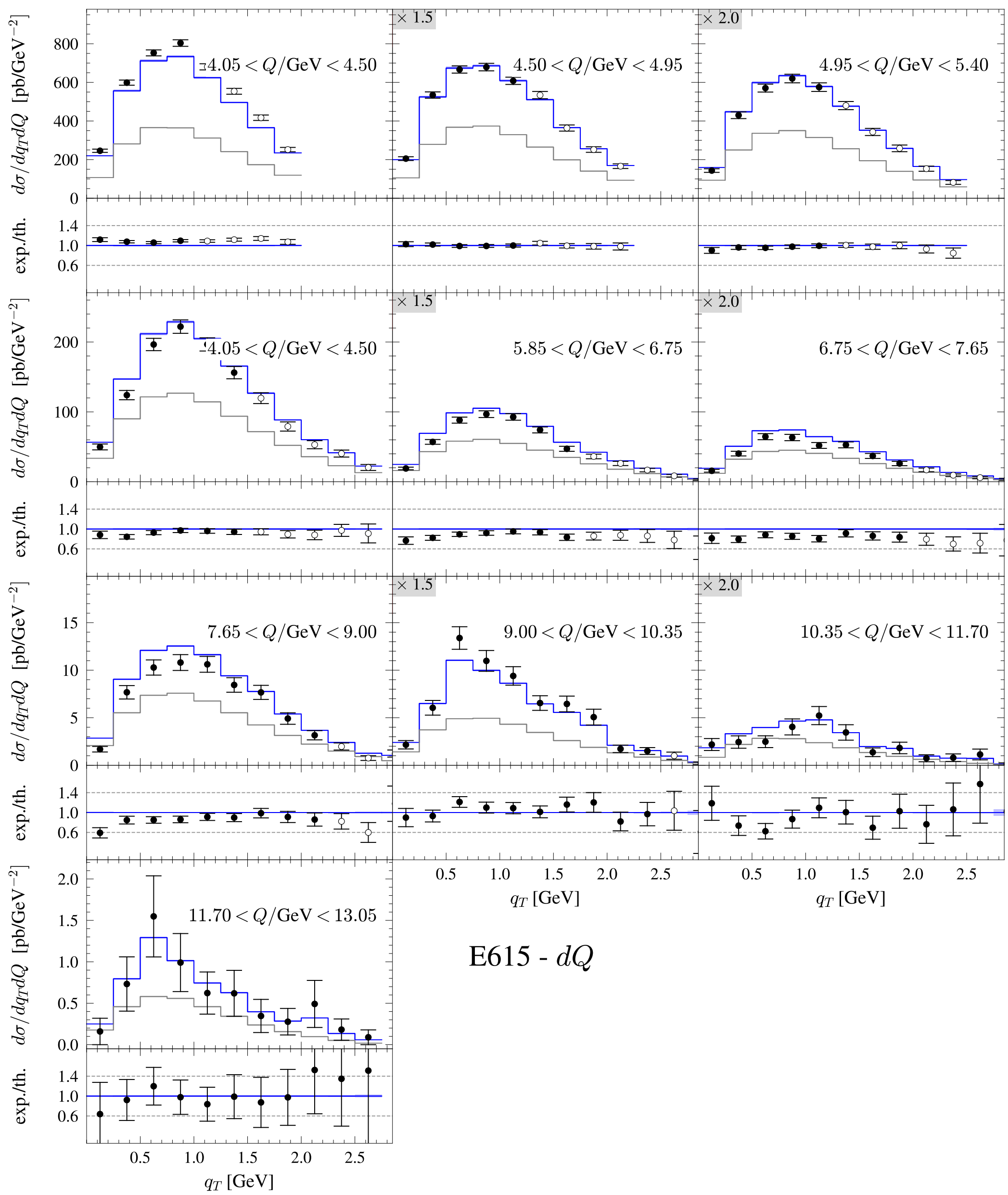


Figure 44: Differential cross section for E615 in bins of Q.

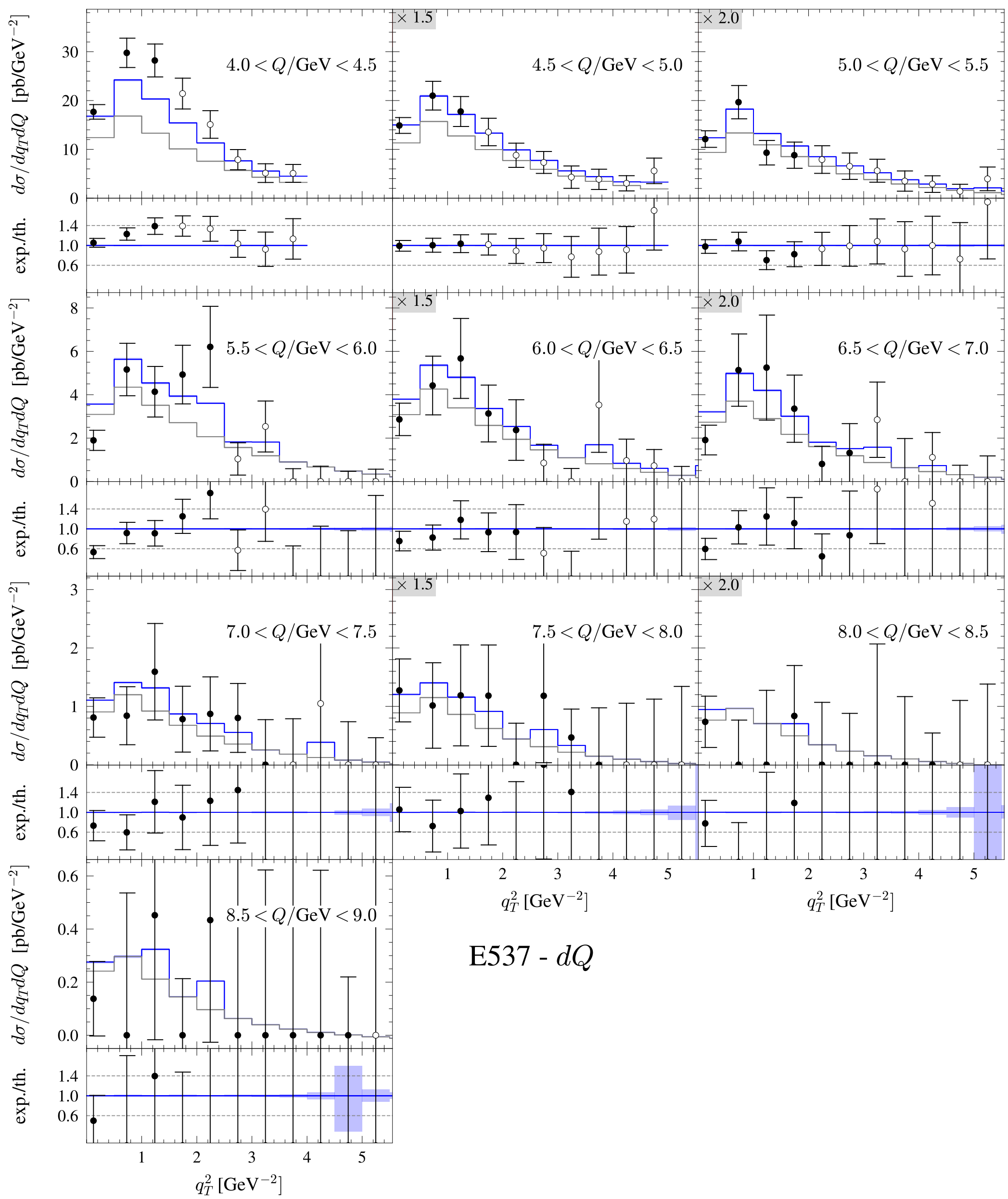


Figure 45: Differential cross section for E537 in bins of Q.

Appendix E Comments on implementations of event generation and their analysis

A Event Generation

Here, were share all the information relevant to understand and reproduce the pseudo data we show in chapter V, namely figure 23-29. For the most technical aspects we focus on `Pythia8`, for all other MCEGs we summarize the boundary conditions.

Regarding the comparison to the COMPASS data, for all MCEGs the following is correct. 10 million MC events are generated for the deuterium target; which consist of data generated with proton target and data with neutron target to the same amount. We have an electron beam with 160 GeV as beam A and the proton (neutron) target at rest[46]. The MCEGs `Pythia6` and `LEPTO` were used in their configuration available at the computers of the Brookhaven National Laboratory (BNL). `Pythia8`, `Herwig7` and `Cascade3` are manually configured and run on a local machine.

For `Pythia8` we use the following user hooks:

```
pythia.readString("Beams:frameType = 2");
// BeamA = proton.
pythia.readString("Beams:idA = 2212"); // a proton, a neutron would be 2112
pythia.settings.parm("Beams:eA", 0);

// BeamB = electron. As mentioned an electron beam is used instead of a muon beam.
pythia.readString("Beams:idB = 11");
pythia.settings.parm("Beams:eB", 160);

// parton radiation
pythia.readString("PartonLevel:ISR = off"); //check.
//pythia.readString("PartonLevel:FSR = off");
// The following two user hooks were crucial to cover the full phase space in HERMES.
// Special thanks to I. Helenius' hint on this!
pythia.readString("PhaseSpace:mHatMin = force 0.1");
pythia.readString("PhaseSpace:pTHatMindiverge = force 0.1");

// Set up DIS process within some phase space.
// Neutral current (with gamma/Z interference).
pythia.readString("WeakBosonExchange:ff2ff(t:gmZ) = on");
pythia.readString("HardQCD:all = off"); // default : off
// Uncomment to allow charged current.
//pythia.readString("WeakBosonExchange:ff2ff(t:W) = on"); // No charged current in our analysis.
// Phase-space cut: minimal Q2 of process.
// cuts later Q2 is restricted to Q2>1 making this cut obsolte.
pythia.settings.parm("PhaseSpace:Q2Min", 0.8);
```

[46] In the experiment, it is muon-nucleon scattering. Because we were not able to use muon-beams in `Pythia6`, for consistency reason we use electron beams in all simulations. With `Pythia8` a comparison was done with the results from muon beams. No significant difference was observed.

```
// The minimum invariant mass.
pythia.settings.parm("PhaseSpace:mHatMax",-1.); //  (default = -1.)
// The maximum invariant mass. A value below mHatMin means there is no upper limit.
// Set dipole recoil on. Necessary for DIS + shower.
pythia.readString("SpaceShower:dipoleRecoil = on");
// Allow emissions up to the kinematical limit,
// since rate known to match well to matrix elements everywhere.
pythia.readString("SpaceShower:pTmaxMatch = 2");
// QED radiation off lepton not handled yet by the new procedure.
pythia.readString("PDF:lepton = off");
//Allow quarks to radiate photons, i.e. branchings q → q gamma; on/off = true/false
pythia.readString("TimeShower:QEDshowerByL = off"); // (default = on)


pythia.readString("Random:setSeed = on"); //   (default = off)
pythia.readString("Random:seed = 1337");

pythia.readString("PDFinProcess:nQuarkIn = 5"); // 0 to 5, quark and antiquarks simultaneously

pythia.readString("PDF:useHard = on");
pythia.readString("PDF:extrapolate = on");
pythia.readString("PDF:pSet = LHAPDF6:" + pdfset + "/" + to_string(pdfMember));
pythia.readString("PDF:extrapolate = on");
// Q2 = -tHat..
pythia.readString("SigmaProcess:factorScale2 = 6");
pythia.readString("SigmaProcess:renormScale2 = 6");
```

With these user hooks the process is specified.

B Event Analysis

The analyses were (mainly[47]) done in `Rivet` [121]. `Rivet` is the standard library used by the community. Its main benefits are understood as the following:

- `Rivet` provides an environment that allows only to analyse the final layers of an event. Inside the code of a MCEG it is possible to access all the information of an event, but in an experimental analysis the event has to be reconstructed from signals. In order to achieve a more trustworthy comparison between experiment and MCEG, `Rivet` sets up an environment that allows to perform kinematic restrictions on accepted signals (cuts) as they are possible in an experimental anylsis.
- the `Rivet` library includes most of the standard definitions of operations on four-momenta and provides access most standard kinematics by *projections* to the process. In our application we use the DIS projection (a SIDIS projection is not (yet) available). The standard fiducial cuts one could impose in an analysis are directly built in and accessible via commands, which are documented online. This saves an enormous amount of time and avoids mistakes in definitions in the basic kinematic definitions in the analysis. The package allows to easily

[47] The analysis for comparing to the data from HERMES is implemented inside the event generation routine for `Pythia8` and in a routine reading from the event files of `Pythia6` at an early stage of this project

use already created and validated analyses directly. In this way, the analyses from different pseudodata are more efficient and consistent among each other[48].

- `Rivet` analyses can be called from command line on a saved event tree in `.hepmc`[122] format, or performed along the event generation in case the generated events do not need to be stored themselves. To process these data it relies on the library `HEPMC`.

To summarize, the existence of the standard analysis software in high energy physics Monte Carlo Generators (HEPMC[49]) of which we highlight the `Rivet` and the `HEPMC` packages enhance the speed and the reliability with which analyses can be written. We use the `Rivet` framework to compute all kinematics and a custom `C++` library for the histograms.

[48] We also plan to publish our analysis once the project is finished; a technical step necessary for this is to implement the histogramming in `Rivet` built-in functions.

[49] See the online reference in the digital version or `https://ep-dep-sft.web.cern.ch/project/hepmc`

References

[1] David H. Sloan and Ernest O. Lawrence. "The Production of Heavy High Speed Ions without the Use of High Voltages". In: *Phys. Rev.* 38 (11 Dec. 1931), pp. 2021–2032. DOI: 10.1103/PhysRev.38.2021. URL: https://link.aps.org/doi/10.1103/PhysRev.38.2021.

[2] J. D. Bjorken. "THEORETICAL IDEAS ON HIGH-ENERGY INELASTIC ELECTRON - PROTON SCATTERING". In: Mar. 1969, pp. 41–51.

[3] John C. Collins, Davison E. Soper, and George Sterman. *Factorization of Hard Processes in QCD*. 2004. arXiv: hep-ph/0409313 [hep-ph].

[4] R. Angeles-Martinez et al. "Transverse Momentum Dependent (TMD) parton distribution functions: status and prospects". In: *Acta Phys. Polon. B* 46.12 (2015), pp. 2501–2534. DOI: 10.5506/APhysPolB.46.2501. arXiv: 1507.05267 [hep-ph].

[5] V. M. Braun, Yao Ji, and Jakob Schoenleber. "Deeply Virtual Compton Scattering at Next-to-Next-to-Leading Order". In: *Phys. Rev. Lett.* 129.17 (2022), p. 172001. DOI: 10.1103/PhysRevLett.129.172001. arXiv: 2207.06818 [hep-ph].

[6] Xiangdong Ji. "Viewing the Proton through "Color" Filters". In: *Phys. Rev. Lett.* 91 (6 Aug. 2003), p. 062001. DOI: 10.1103/PhysRevLett.91.062001. URL: https://link.aps.org/doi/10.1103/PhysRevLett.91.062001.

[7] A.V. Belitsky, Xiangdong Ji, and Feng Yuan. "Quark imaging in the proton via quantum phase-space distributions". In: *Phys. Rev. D* 69 (7 Apr. 2004), p. 074014. DOI: 10.1103/PhysRevD.69.074014. URL: https://link.aps.org/doi/10.1103/PhysRevD.69.074014.

[8] Alessandro Bacchetta et al. "Unpolarized transverse momentum distributions from a global fit of Drell-Yan and semi-inclusive deep-inelastic scattering data". In: *JHEP* 10 (2022), p. 127. DOI: 10.1007/JHEP10(2022)127. arXiv: 2206.07598 [hep-ph].

[9] Valentin Moos et al. "Extraction of unpolarized transverse momentum distributions from fit of Drell-Yan data at N^{4}LL". In: (May 2023). arXiv: 2305.07473 [hep-ph].

[10] Marcin Bury et al. "PDF bias and flavor dependence in TMD distributions". In: *JHEP* 10 (2022), p. 118. DOI: 10.1007/JHEP10(2022)118. arXiv: 2201.07114 [hep-ph].

[11] Alessandro Bacchetta et al. "Extraction of partonic transverse momentum distributions from semi-inclusive deep-inelastic scattering, Drell-Yan and Z-boson production". In: *JHEP* 06 (2017). [Erratum: JHEP 06, 051 (2019)], p. 081. DOI: 10.1007/JHEP06(2017)081. arXiv: 1703.10157 [hep-ph].

[12] Ignazio Scimemi and Alexey Vladimirov. "Non-perturbative structure of semi-inclusive deep-inelastic and Drell-Yan scattering at small transverse momentum". In: *JHEP* 06 (2020), p. 137. DOI: 10.1007/JHEP06(2020)137. arXiv: 1912.06532 [hep-ph].

[13] Roman N. Lee et al. "Quark and Gluon Form Factors in Four-Loop QCD". In: *Phys. Rev. Lett.* 128.21 (2022), p. 212002. DOI: 10.1103/PhysRevLett.128.212002. arXiv: 2202.04660 [hep-ph].

[14] S. Bailey et al. "Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs". In: *Eur. Phys. J. C* 81.4 (2021), p. 341. DOI: 10.1140/epjc/s10052-021-09057-0. arXiv: 2012.04684 [hep-ph].

[15] Ignacio Borsa et al. "Towards a Global QCD Analysis of Fragmentation Functions at Next-to-Next-to-Leading Order Accuracy". In: *Phys. Rev. Lett.* 129.1 (2022), p. 012002. DOI: 10.1103/PhysRevLett.129.012002. arXiv: 2202.05060 [hep-ph].

[16] R. J. Hernández-Pinto et al. "Global extraction of the parton-to-kaon fragmentation functions at NLO in QCD". In: *J. Phys. Conf. Ser.* 912.1 (2017). Ed. by I. Bautista et al., p. 012043. DOI: 10.1088/1742-6596/912/1/012043. arXiv: 1709.07415 [hep-ph].

[17] A. Glazov. *Recent Results from HERA*. 2009. arXiv: 0911.0159 [hep-ex].

[18] A. Airapetian et al. "Multiplicities of charged pions and kaons from semi-inclusive deep-inelastic scattering by the proton and the deuteron". In: *Phys. Rev. D* 87 (2013), p. 074029. DOI: 10.1103/PhysRevD.87.074029. arXiv: 1212.5407 [hep-ex].

[19] C. Adolph et al. "Multiplicities of charged kaons from deep-inelastic muon scattering off an isoscalar target". In: *Phys. Lett. B* 767 (2017), pp. 133–141. DOI: 10.1016/j.physletb.2017.01.053. arXiv: 1608.06760 [hep-ex].

[20] M. Aghasyan et al. "Transverse-momentum-dependent Multiplicities of Charged Hadrons in Muon-Deuteron Deep Inelastic Scattering". In: *Phys. Rev. D* 97.3 (2018), p. 032006. DOI: 10.1103/PhysRevD.97.032006. arXiv: 1709.07374 [hep-ex].

[21] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley, 1995. ISBN: 978-0-201-50397-5.

[22] Francisco J. Ynduráin. *The Theory of Quark and Gluon Interactions*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Springer-Verlag Berlin Heidelberg 1993,1999, 2006. DOI: 849/UO5700Y53(3).

[23] R. P. Feynman. "The behavior of hadron collisions at extreme energies". In: *Conf. Proc. C* 690905 (1969), pp. 237–258.

[24] John C. Collins. "Hard scattering factorization with heavy quarks: A General treatment". In: *Phys. Rev. D* 58 (1998), p. 094002. DOI: 10.1103/PhysRevD.58.094002. arXiv: hep-ph/9806259.

[25] O. V. Tarasov. "Anomalous Dimensions of Quark Masses in the Three-Loop Approximation". In: *Physics of Particles and Nuclei Letters* 17.2 (Mar. 2020), pp. 109–115. ISSN: 1531-8567. DOI: 10.1134/s1547477120020223. URL: http://dx.doi.org/10.1134/S1547477120020223.

[26] Guido Altarelli and G. Parisi. "Asymptotic Freedom in Parton Language". In: *Nucl. Phys. B* 126 (1977), pp. 298–318. DOI: 10.1016/0550-3213(77)90384-4.

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

[28] V. N. Gribov and L. N. Lipatov. "e+ e- pair annihilation and deep inelastic e p scattering in perturbation theory". In: *Sov. J. Nucl. Phys.* 15 (1972), pp. 675–684.

[29] M. Jezabek. "A Perturbative solution to Gribov-Lipatov equation". In: *Z. Phys. C* 56 (1992), pp. 285–288. DOI: 10.1007/BF01555526.

[30] M. Botje. "QCDNUM: Fast QCD evolution and convolution". In: *Computer Physics Communications* 182.2 (Feb. 2011), pp. 490–532. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2010.10.020. URL: http://dx.doi.org/10.1016/j.cpc.2010.10.020.

[31] A. Vogt. "Efficient evolution of unpolarized and polarized parton distributions with QCD-Pegasus". In: *Computer Physics Communications* 170.1 (July 2005), pp. 65–92. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2005.03.103. URL: http://dx.doi.org/10.1016/j.cpc.2005.03.103.

[32] Richard D. Ball et al. "A Determination of parton distributions with faithful uncertainty estimation". In: *Nucl. Phys. B* 809 (2009). [Erratum: Nucl.Phys.B 816, 293 (2009)], pp. 1–63. DOI: 10.1016/j.nuclphysb.2008.09.037. arXiv: 0808.1231 [hep-ph].

[33] Sebastian Ruder. "An overview of gradient descent optimization algorithms". In: (Sept. 2016). arXiv: 1609.04747 [cs.LG].

[34] D. de Florian, M. Stratmann, and W. Vogelsang. "QCD analysis of unpolarized and polarized Lambda baryon production in leading and next-to-leading order". In: *Phys. Rev. D* 57 (1998), pp. 5811–5824. DOI: 10.1103/PhysRevD.57.5811. arXiv: hep-ph/9711387.

[35] Daniel de Florian, Rodolfo Sassot, and Marco Stratmann. "Global analysis of fragmentation functions for pions and kaons and their uncertainties". In: *Physical Review D* 75.11 (June 2007). ISSN: 1550-2368. DOI: 10.1103/physrevd.75.114010. URL: http://dx.doi.org/10.1103/PhysRevD.75.114010.

[36] M. Leitgab et al. "Precision Measurement of Charged Pion and Kaon Differential Cross Sections in e+e- Annihilation at s=10.52 GeV". In: *Phys. Rev. Lett.* 111 (2013), p. 062002. DOI: 10.1103/PhysRevLett.111.062002. arXiv: 1301.6183 [hep-ex].

[37] J. Adams et al. "Forward neutral pion production in p+p and d+Au collisions at s(NN)**(1/2) = 200-GeV". In: *Phys. Rev. Lett.* 97 (2006), p. 152302. DOI: 10.1103/PhysRevLett.97.152302. arXiv: nucl-ex/0602011.

[38] John C. Collins, Davison E. Soper, and George F. Sterman. "Factorization of Hard Processes in QCD". In: *Adv. Ser. Direct. High Energy Phys.* 5 (1989), pp. 1–91. DOI: 10.1142/9789814503266_0001. arXiv: hep-ph/0409313.

[39] Ignacio Borsa et al. "Pion fragmentation functions at high energy colliders". In: *Phys. Rev. D* 105.3 (2022), p. L031502. DOI: 10.1103/PhysRevD.105.L031502. arXiv: 2110.14015 [hep-ph].

[40] Alessandro Bacchetta et al. "Semi-inclusive deep inelastic scattering at small transverse momentum". In: *JHEP* 02 (2007), p. 093. DOI: 10.1088/1126-6708/2007/02/093. arXiv: hep-ph/0611265.

[41] D. Boer, P.J. Mulders, and F. Pijlman. "Universality of T-odd effects in single spin and azimuthal asymmetries". In: *Nuclear Physics B* 667.1–2 (Sept. 2003), pp. 201–241. ISSN: 0550-3213. DOI: 10.1016/s0550-3213(03)00527-3. URL: http://dx.doi.org/10.1016/S0550-3213(03)00527-3.

[42] Valentin Moos and Alexey Vladimirov. "Calculation of transverse momentum dependent distributions beyond the leading power". In: (Aug. 2020). arXiv: 2008.01744 [hep-ph].

[43] S. D. Drell and Tung-Mow Yan. "Massive Lepton Pair Production in Hadron-Hadron Collisions at High-Energies". In: *Phys. Rev. Lett.* 25 (1970). [Erratum: Phys.Rev.Lett. 25, 902 (1970)], pp. 316–320. DOI: 10.1103/PhysRevLett.25.316.

[44] Thomas Becher and Matthias Neubert. "Drell-Yan Production at Small q_T, Transverse Parton Distributions and the Collinear Anomaly". In: *Eur. Phys. J. C* 71 (2011), p. 1665. DOI: 10.1140/epjc/s10052-011-1665-7. arXiv: 1007.4005 [hep-ph].

[45] Alexey Vladimirov, Valentin Moos, and Ignazio Scimemi. "Transverse momentum dependent operator expansion at next-to-leading power". In: *JHEP* 01 (2022), p. 110. DOI: 10.1007/JHEP01(2022)110. arXiv: 2109.09771 [hep-ph].

[46] Palash B. Pal. *Representation-independent manipulations with Dirac matrices and spinors*. 2015. arXiv: physics/0703214 [physics.ed-ph].

[47] Thomas Becher, Matthias Neubert, and Gang Xu. "Dynamical Threshold Enhancement and Resummation in Drell-Yan Production". In: *JHEP* 07 (2008), p. 030. DOI: 10.1088/1126-6708/2008/07/030. arXiv: 0710.0680 [hep-ph].

[48] J.C. Collins, Davison E. Soper, and George Sterman. "Transverse momentum distribution in Drell-Yan pair and W and Z boson production". In: *Nuclear Physics B* 250.1 (1985), pp. 199–224. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(85)90479-1. URL: https://www.sciencedirect.com/science/article/pii/0550321385904791.

[49] Ignazio Scimemi and Alexey Vladimirov. "Systematic analysis of double-scale evolution". In: *JHEP* 08 (2018), p. 003. DOI: 10.1007/JHEP08(2018)003. arXiv: 1803.11089 [hep-ph].

[50] Ignazio Scimemi and Alexey Vladimirov. "Matching of transverse momentum dependent distributions at twist-3". In: *The European Physical Journal C* 2018.8 (Sept. 2018), p. 19. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-018-6263-5. URL: https://doi.org/10.1140/epjc/s10052-018-6263-5.

[51] L. F. Abbott. "Introduction to the Background Field Method". In: *Acta Phys. Polon. B* 13 (1982), p. 33.

[52] Vladyslav Shtabovenko. "FeynCalc 9". In: *J. Phys. Conf. Ser.* 762.1 (2016). Ed. by Luis Salinas and Claudio Torres, p. 012064. DOI: 10.1088/1742-6596/762/1/012064. arXiv: 1604.06709 [hep-ph].

[53] I.I. Balitsky and V.M. Braun. "Evolution equations for QCD string operators". In: *Nuclear Physics B* 311.3 (1989), pp. 541–584. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(89)90168-5. URL: https://www.sciencedirect.com/science/article/pii/0550321389901685.

[54] F. Landry et al. "Tevatron Run-1 Z boson data and Collins-Soper-Sterman resummation formalism". In: *Phys. Rev. D* 67 (2003), p. 073016. DOI: 10.1103/PhysRevD.67.073016. arXiv: hep-ph/0212159.

[55] Filippo Delcarro et al. "Extraction of Transverse Momentum Distributions from semi-inclusive DIS and Drell-Yan data." In: *PoS* DIS2018 (2018), p. 219. DOI: 10.22323/1.316.0219.

[56] Valerio Bertone, Ignazio Scimemi, and Alexey Vladimirov. "Extraction of unpolarized quark transverse momentum dependent parton distributions from Drell-Yan/Z-boson production". In: *JHEP* 06 (2019), p. 028. DOI: 10.1007/JHEP06(2019)028. arXiv: 1902.08474 [hep-ph].

[57] M. Stratmann and W. Vogelsang. "Next-to-leading order evolution of polarized and unpolarized fragmentation functions". In: *Nucl. Phys. B* 496 (1997), pp. 41–65. DOI: 10.1016/S0550-3213(97)00182-X. arXiv: hep-ph/9612250.

[58] Miguel G. Echevarria, Ignazio Scimemi, and Alexey Vladimirov. "Transverse momentum dependent fragmentation function at next-to–next-to–leading order". In: *Phys. Rev. D* 93.1 (2016). [Erratum: Phys.Rev.D 94, 099904 (2016)], p. 011502. DOI: 10.1103/PhysRevD.93.011502. arXiv: 1509.06392 [hep-ph].

[59] Miguel G. Echevarria, Ignazio Scimemi, and Alexey Vladimirov. "Unpolarized Transverse Momentum Dependent Parton Distribution and Fragmentation Functions at next-to-next-to-leading order". In: *JHEP* 09 (2016), p. 004. DOI: 10.1007/JHEP09(2016)004. arXiv: 1604.07869 [hep-ph].

[60] Alessandro Bacchetta et al. "Global extraction of unpolarized quark TMDs at N3LL". In: *PoS* ICHEP2022 (2022), p. 800. DOI: 10.22323/1.414.0800.

[61] Richard D. Ball et al. "Parton Distribution Benchmarking with LHC Data". In: *JHEP* 04 (2013), p. 125. DOI: 10.1007/JHEP04(2013)125. arXiv: 1211.5142 [hep-ph].

[62] Claude Duhr, Bernhard Mistlberger, and Gherardo Vita. "Four-Loop Rapidity Anomalous Dimension and Event Shapes to Fourth Logarithmic Order". In: *Phys. Rev. Lett.* 129.16 (2022), p. 162001. DOI: 10.1103/PhysRevLett.129.162001. arXiv: 2205.02242 [hep-ph].

[63] Ian Moult, Hua Xing Zhu, and Yu Jiao Zhu. “The four loop QCD rapidity anomalous dimension”. In: *JHEP* 08 (2022), p. 280. DOI: 10.1007/JHEP08(2022)280. arXiv: 2205.02249 [hep-ph].

[64] Markus A. Ebert, Bernhard Mistlberger, and Gherardo Vita. “Transverse momentum dependent PDFs at N^{3}LO”. In: *JHEP* 09 (2020), p. 146. DOI: 10.1007/JHEP09(2020)146. arXiv: 2006.05329 [hep-ph].

[65] Ming-xing Luo et al. “Unpolarized quark and gluon TMD PDFs and FFs at N^{3}LO”. In: *JHEP* 06 (2021), p. 115. DOI: 10.1007/JHEP06(2021)115. arXiv: 2012.03256 [hep-ph].

[66] Juan Rojo et al. “The PDF4LHC report on PDFs and LHC data: Results from Run I and preparation for Run II”. In: *J. Phys. G* 42 (2015), p. 103103. DOI: 10.1088/0954-3899/42/10/103103. arXiv: 1507.00556 [hep-ph].

[67] Richard D. Ball et al. “Parton distributions from high-precision collider data”. In: *Eur. Phys. J. C* 77.10 (2017), p. 663. DOI: 10.1140/epjc/s10052-017-5199-5. arXiv: 1706.00428 [hep-ph].

[68] H. Abramowicz et al. “Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data”. In: *Eur. Phys. J. C* 75.12 (2015), p. 580. DOI: 10.1140/epjc/s10052-015-3710-4. arXiv: 1506.06042 [hep-ex].

[69] Tie-Jiun Hou et al. “New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC”. In: *Phys. Rev. D* 103.1 (2021), p. 014013. DOI: 10.1103/PhysRevD.103.014013. arXiv: 1912.10053 [hep-ph].

[70] I. Bubanja et al. “Center-of-mass energy dependence of intrinsic-k_T distributions obtained from Drell-Yan production”. In: (Apr. 2024). arXiv: 2404.04088 [hep-ph].

[71] F. James and M. Roos. “Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations”. In: *Comput. Phys. Commun.* 10 (1975), pp. 343–367. DOI: 10.1016/0010-4655(75)90039-9.

[72] *artemide*. stable version: https://github.com/VladimirovAlexey/artemide-public
in-production version: https://github.com/VladimirovAlexey/artemide-development. Accessed: 2023-04-30.

[73] Andy Buckley et al. “LHAPDF6: parton density access in the LHC precision era”. In: *Eur. Phys. J. C* 75 (2015), p. 132. DOI: 10.1140/epjc/s10052-015-3318-8. arXiv: 1412.7420 [hep-ph].

[74] *artemide-DataProcessor*. https://github.com/VladimirovAlexey/artemide-DataProcessor. Accessed: 2023-04-30.

[75] Georges Aad et al. “Measurement of the transverse momentum distribution of Drell–Yan lepton pairs in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Eur. Phys. J. C* 80.7 (2020), p. 616. DOI: 10.1140/epjc/s10052-020-8001-z. arXiv: 1912.02844 [hep-ex].

[76] M. Boglione, J. O. Gonzalez-Hernandez, and A. Simonelli. “Transverse momentum dependent fragmentation functions from recent BELLE data”. In: *Phys. Rev. D* 106.7 (2022), p. 074024. DOI: 10.1103/PhysRevD.106.074024. arXiv: 2206.08876 [hep-ph].

[77] Valerio Bertone, Stefano Carrazza, and Juan Rojo. “APFEL: A PDF evolution library with QED corrections”. In: *Computer Physics Communications* 185.6 (June 2014), pp. 1647–1668. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2014.03.007. URL: http://dx.doi.org/10.1016/j.cpc.2014.03.007.

[78] S. Alekhin et al. *HERAFitter, Open Source QCD Fit Project*. 2015. arXiv: 1410.4412 [hep-ph].

[79] E. Anassontzis et al. “High mass dimuon production in $\bar{p}n$ and $\pi^- n$ interactions at 125-GeV/c”. In: *Phys. Rev. D* 38 (1988), p. 1377. DOI: 10.1103/PhysRevD.38.1377.

[80] J. S. Conway et al. “Experimental Study of Muon Pairs Produced by 252-GeV Pions on Tungsten”. In: *Phys. Rev. D* 39 (1989), pp. 92–122. DOI: 10.1103/PhysRevD.39.92.

[81] Alexey Vladimirov. “Pion-induced Drell-Yan processes within TMD factorization”. In: *JHEP* 10 (2019), p. 090. DOI: 10.1007/JHEP10(2019)090. arXiv: 1907.10356 [hep-ph].

[82] Lorenzo Rossi. “Fits of unpolarized proton and pion TMDs at N3LL accuracy”. In: *Nuovo Cim. C* 46.4 (2023), p. 102. DOI: 10.1393/ncc/i2023-23102-2.

[83] F Gautheron et al. “COMPASS-II Proposal”. In: (May 2010).

[84] P. C. Barry et al. “Global QCD Analysis of Pion Parton Distributions with Threshold Resummation”. In: *Phys. Rev. Lett.* 127.23 (2021), p. 232001. DOI: 10.1103/PhysRevLett.127.232001. arXiv: 2108.05822 [hep-ph].

[85] P. C. Barry et al. “First Monte Carlo Global QCD Analysis of Pion Parton Distributions”. In: *Phys. Rev. Lett.* 121.15 (2018), p. 152001. DOI: 10.1103/PhysRevLett.121.152001. arXiv: 1804.01965 [hep-ph].

[86] R. Abdul Khalek et al. “Science Requirements and Detector Concepts for the Electron-Ion Collider”. In: *Nuclear Physics A* 1026 (Oct. 2022), p. 122447. ISSN: 0375-9474. DOI: 10.1016/j.nuclphysa.2022.122447. URL: http://dx.doi.org/10.1016/j.nuclphysa.2022.122447.

[87] Andy Buckley et al. “General-purpose event generators for LHC physics”. In: *Physics Reports* 504.5 (July 2011), pp. 145–233. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2011.03.005. URL: http://dx.doi.org/10.1016/j.physrep.2011.03.005.

[88] J. M. Campbell et al. “Event Generators for High-Energy Physics Experiments”. In: *Snowmass 2021*. Mar. 2022. arXiv: 2203.11110 [hep-ph].

[89] Torbjorn Sjostrand. “PYTHIA at HERA”. In: *Workshop on Physics at HERA*. Jan. 1992.

[90] Helen Caines. “Exploring Jet Properties in p-p Collisions at 200-GeV with STAR”. In: *Nucl. Phys. A* 830 (2009). Ed. by Paul Stankus et al., pp. 263C–266C. DOI: 10.1016/j.nuclphysa.2009.10.096. arXiv: 0907.3460 [nucl-ex].

[91] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. “PYTHIA 6.4 Physics and Manual”. In: *JHEP* 05 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175.

[92] S. Baranov et al. “CASCADE3 A Monte Carlo event generator based on TMDs”. In: *Eur. Phys. J. C* 81.5 (2021), p. 425. DOI: 10.1140/epjc/s10052-021-09203-8. arXiv: 2101.10221 [hep-ph].

[93] Johannes Bellm et al. “Herwig 7.0/Herwig++ 3.0 release note”. In: *Eur. Phys. J. C* 76.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].

[94] T. Gleisberg et al. “Event generation with SHERPA 1.1”. In: *JHEP* 02 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph].

[95] Christian Bierlich et al. *A comprehensive guide to the physics and usage of PYTHIA 8.3*. 2022. arXiv: 2203.11601 [hep-ph].

[96] Torbjörn Sjöstrand. “A model for initial state parton showers”. In: *Physics Letters B* 157.4 (1985), pp. 321–325. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(85)90674-4. URL: https://www.sciencedirect.com/science/article/pii/0370269385906744.

[97] Jeppe R. Andersen et al. “Exploring High-Purity Multiparton Scattering at Hadron Colliders”. In: *Physical Review Letters* 132.4 (Jan. 2024). ISSN: 1079-7114. DOI: 10.1103/physrevlett.132.041901. URL: http://dx.doi.org/10.1103/PhysRevLett.132.041901.

[98] G.'t Hooft. “A planar diagram theory for strong interactions”. In: *Nuclear Physics B* 72.3 (1974), pp. 461–473. ISSN: 0550-3213. DOI: `https://doi.org/10.1016/0550-3213(74)90154-0`. URL: `https://www.sciencedirect.com/science/article/pii/0550321374901540`.

[99] Silvia Ferreres-Solé and Torbjörn Sjöstrand. “The space–time structure of hadronization in the Lund model”. In: *The European Physical Journal C* 78.11 (Nov. 2018). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-018-6459-8`. URL: `http://dx.doi.org/10.1140/epjc/s10052-018-6459-8`.

[100] Hung-Liang Lai et al. “New parton distributions for collider physics”. In: *Physical Review D* 82.7 (Oct. 2010). ISSN: 1550-2368. DOI: `10.1103/physrevd.82.074024`. URL: `http://dx.doi.org/10.1103/PhysRevD.82.074024`.

[101] G Ingelman, A Edin, and J Rathsman. “LEPTO 6.5 — A Monte Carlo generator for deep inelastic lepton-nucleon scattering”. In: *Computer Physics Communications* 101.1–2 (Apr. 1997), pp. 108–134. ISSN: 0010-4655. DOI: `10.1016/s0010-4655(96)00157-9`. URL: `http://dx.doi.org/10.1016/S0010-4655(96)00157-9`.

[102] Andy Buckley et al. “Systematic event generator tuning for the LHC”. In: *The European Physical Journal C* 65.1–2 (Nov. 2009), pp. 331–357. ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-009-1196-7`. URL: `http://dx.doi.org/10.1140/epjc/s10052-009-1196-7`.

[103] Tie-Jiun Hou et al. “Reconstruction of Monte Carlo replicas from Hessian parton distributions”. In: *JHEP* 03 (2017), p. 099. DOI: `10.1007/JHEP03(2017)099`. arXiv: `1607.06066 [hep-ph]`.

[104] Georges Aad et al. “Measurement of the transverse momentum and ϕ^*_η distributions of Drell–Yan lepton pairs in proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector”. In: *Eur. Phys. J. C* 76.5 (2016), p. 291. DOI: `10.1140/epjc/s10052-016-4070-4`. arXiv: `1512.02192 [hep-ex]`.

[105] Roel Aaij et al. “Measurement of the forward Z boson production cross-section in pp collisions at $\sqrt{s} = 7$ TeV”. In: *JHEP* 08 (2015), p. 039. DOI: `10.1007/JHEP08(2015)039`. arXiv: `1505.07024 [hep-ex]`.

[106] Roel Aaij et al. “Measurement of forward W and Z boson production in pp collisions at $\sqrt{s} = 8$ TeV”. In: *JHEP* 01 (2016), p. 155. DOI: `10.1007/JHEP01(2016)155`. arXiv: `1511.08039 [hep-ex]`.

[107] Roel Aaij et al. “Measurement of the forward Z boson production cross-section in pp collisions at $\sqrt{s} = 13$ TeV”. In: *JHEP* 09 (2016), p. 136. DOI: `10.1007/JHEP09(2016)136`. arXiv: `1607.06495 [hep-ex]`.

[108] R. Aaij et al. “Precision measurement of forward Z boson production in proton-proton collisions at $\sqrt{s} = 13$ TeV”. In: *JHEP* 07 (2022), p. 026. DOI: `10.1007/JHEP07(2022)026`. arXiv: `2112.07458 [hep-ex]`.

[109] C. Aidala et al. “Measurements of $\mu\mu$ pairs from open heavy flavor and Drell-Yan in $p + p$ collisions at $\sqrt{s} = 200$ GeV”. In: *Phys. Rev. D* 99.7 (2019), p. 072003. DOI: `10.1103/PhysRevD.99.072003`. arXiv: `1805.02448 [hep-ex]`.

[110] Salvatore Fazio and Xiaoxuan Chu. private communication. Feb. 3, 2013.

[111] T. Affolder et al. “The transverse momentum and total cross section of e^+e^- pairs in the Z boson region from $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV”. In: *Phys. Rev. Lett.* 84 (2000), pp. 845–850. DOI: `10.1103/PhysRevLett.84.845`. arXiv: `hep-ex/0001021`.

[112] T. Aaltonen et al. “Transverse momentum cross section of e^+e^- pairs in the Z-boson region from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV”. In: *Phys. Rev. D* 86 (2012), p. 052010. DOI: `10.1103/PhysRevD.86.052010`. arXiv: `1207.7138 [hep-ex]`.

[113] F. Abe et al. “Measurement of the W P(T) distribution in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV”. In: *Phys. Rev. Lett.* 66 (1991), pp. 2951–2955. DOI: `10.1103/PhysRevLett.66.2951`.

[114] Victor Mukhamedovich Abazov et al. "Measurement of the Normalized $Z/\gamma^*- > \mu^+\mu^-$ Transverse Momentum Distribution in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV". In: *Phys. Lett. B* 693 (2010), pp. 522–530. DOI: 10.1016/j.physletb.2010.09.012. arXiv: 1006.0618 [hep-ex].

[115] V. M. Abazov et al. "Measurement of the shape of the boson transverse momentum distribution in $p\bar{p} \to Z/\gamma^* \to e^+e^- + X$ events produced at $\sqrt{s}$=1.96-TeV". In: *Phys. Rev. Lett.* 100 (2008), p. 102002. DOI: 10.1103/PhysRevLett.100.102002. arXiv: 0712.0803 [hep-ex].

[116] B. Abbott et al. "Measurement of the inclusive differential cross section for Z bosons as a function of transverse momentum in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV". In: *Phys. Rev. D* 61 (2000), p. 032004. DOI: 10.1103/PhysRevD.61.032004. arXiv: hep-ex/9907009.

[117] B. Abbott et al. "Measurement of the shape of the transverse momentum distribution of W bosons produced in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV". In: *Phys. Rev. Lett.* 80 (1998), pp. 5498–5503. DOI: 10.1103/PhysRevLett.80.5498. arXiv: hep-ex/9803003.

[118] A. S. Ito et al. "Measurement of the Continuum of Dimuons Produced in High-Energy Proton - Nucleus Collisions". In: *Phys. Rev. D* 23 (1981), pp. 604–633. DOI: 10.1103/PhysRevD.23.604.

[119] P. L. McGaughey et al. "Cross-sections for the production of high mass muon pairs from 800-GeV proton bombardment of H-2". In: *Phys. Rev. D* 50 (1994). [Erratum: Phys.Rev.D 60, 119903 (1999)], pp. 3038–3045. DOI: 10.1103/PhysRevD.50.3038.

[120] G. Moreno et al. "Dimuon production in proton - copper collisions at $\sqrt{s} = 38.8$-GeV". In: *Phys. Rev. D* 43 (1991), pp. 2815–2836. DOI: 10.1103/PhysRevD.43.2815.

[121] Christian Bierlich et al. "Robust Independent Validation of Experiment and Theory: Rivet version 3". In: *SciPost Physics* 8.2 (Feb. 2020). ISSN: 2542-4653. DOI: 10.21468/scipostphys.8.2.026. URL: http://dx.doi.org/10.21468/SciPostPhys.8.2.026.

[122] Matt Dobbs and Jørgen Beck Hansen. "The HepMC C++ Monte Carlo event record for High Energy Physics Available via the following web-address: http://home.cern.ch/mdobbs/HepMC/." In: *Computer Physics Communications* 134.1 (2001), pp. 41–46. ISSN: 0010-4655. DOI: https://doi.org/10.1016/S0010-4655(00)00189-2. URL: https://www.sciencedirect.com/science/article/pii/S0010465500001892.

Declaration of primary authorship

I hereby confirm that I have authored this dissertation independently and without use of others than the indicated sources. All passages which are literally or in general matter taken out of publications or other sources are marked as such.

Regensburg, 14. Juni 2024

Acknowledgements

I have to, but more importantly I want to, thank several persons for their support along the way. Without their help, non of the results would stand. I want to thank Kolja Kauder and Xiaoxuan Chu for their help with their technical support to run the `Pythia6` and `LEPTO` on the BNL computers. Several persons from the `Herwig7` collaboration helped me with finding a proper setup with the program, so I thank their whole collaboration. I am very grateful for the support and patience I have experienced from from Hannes Jung with regarding the installation of and event generation with `Cascade3`. Valerio Bertone has my gratitude for his unrequested (but even more appreciated) help with `apfel(++)`. Concerning `Pythia8`, I want to point out the help from Ilkka Helenius, who gave me technical support in the early stages of the project and pointed out the a necessary flag such that the HERMES and COMPASS data could be described; and I thank Leif Gellerson for discussions and helpful explanations about the structure of `Pythia8`. I want to thank Vladimir and Alexandr for accepting me as part of their group, ensuring helpful physical explanations and interesting political discussions over a cup of tea. I want to thank my team for the great collaboration in the past and I hope for a continuation thereof in the future.
I thank my colleague to Óscar for his help with the computation of the diagrams, which could not be concluded without the verification from another person. Without Óscar, I believe, the correct signs would represent a minority.

I thank Alexey very much for possibility to be involved in several projects, all of which I personally appreciate as chances to learn something different, but which I also understand the an offered support in gaining a foothold in academia.

I want to express my gratitude for the support I have experienced by my supervisor Pia. Pia has supported me by giving the projects a necessary guidance and structure and making it possible for me to go to various summer schools and conferences, both to the benefit of my personal and scientific development. I could always count on help when it came to problems I could not solve on my own. Last but not least I want to thank Pia for the unrelentles effort in supporting me writing the thesis. None of the mentioned can be taken for granted and I shall never forget it.

I also thank my friends and family for their support, most of all I thank my mother.