

UNIVERSITÀ DEGLI STUDI DI TORINO

SCUOLA DI SCIENZE DELLA NATURA

Corso di Laurea Magistrale in Fisica



UNIVERSITÀ
DI TORINO

Tesi di Laurea Magistrale

Closure Tests of Transverse Momentum Distributions

Relatore:
Andrea Signori

Correlatore:
Emanuele Roberto Nocera

Candidato:
Kamil Laurent

Controrelatore:
Simon David Badger

Anno Accademico 2023/2024

Abstract

The subject of this thesis is the phenomenological determination of Transverse Momentum Distributions (TMDs), which describe the structure of hadrons (like protons and neutrons) in 3D momentum space. They are crucial for understanding the dynamical properties arising from their elementary constituents, quarks and gluons. We present a statistical analysis of two fitting frameworks for determining unpolarized TMDs, referred to as PV19 and MAPTMD22. These frameworks utilize parametric regression to infer TMDs through fits to experimental data from Semi-Inclusive Deep Inelastic Scattering (SIDIS, $\ell p \rightarrow \ell h X$) and Drell-Yan ($pp \rightarrow \ell^- \ell^+ X$) cross sections. The robustness of these methodologies is evaluated via closure testing, which assesses three critical aspects: the flexibility of the framework in capturing the true functional form of the TMDs, the contribution of uncertainties induced by the framework to the total uncertainty, and the statistical faithfulness of the TMD uncertainty estimates. This work contributes to the broader quest for precision in high energy physics, by developing tools to improve the reliability of the predicted TMDs and their associated uncertainties.

To find something out,
it is better to perform some careful experiments
than to carry on deep philosophical arguments.

Richard P. Feynman, *The Feynman Lectures on Physics*

Acknowledgments

I would like to thank my mother, Laura, my father, Ginetto, and my sister, Amalia, for their constant presence and support throughout these years of study.

A special thank to my supervisors, Andrea Signori and Emanuele Roberto Nocera, for their invaluable help, their care, their precision, and their guidance. This work would not have been possible without their mentorship. Working with them has been a pleasure, both personally and scientifically.

Lastly, I am grateful to my friends Rebecca, Gionata, Luca, and Mattia, with whom I spent countless hours discussing physics and bold ideas.

Contents

1	Introduction	1
2	From Parton Model to TMDs	5
2.1	QCD Lagrangian and Asymptotic Freedom	5
2.2	The Parton Model	7
2.3	Operator definition of TMDs	14
2.4	SIDIS and Drell-Yan Factorization	16
2.5	TMD Evolution Equations	22
3	TMD Extraction from Experimental Data	27
3.1	The Inverse Problem	27
3.2	Experimental Datasets	28
3.3	PV19 and MAPTMD22 Models	32
3.4	Input Parametrization	33
3.5	The Fitting Methodology	36
4	Closure Testing	40
4.1	Closure Tests	40
4.2	Level 0 Test	44
4.3	Level 1 Test	46
4.4	Level 2 Test	47
4.5	Multi Closure Tests	49
4.6	L0 Pseudo-Data Generation	53
4.7	Data Uncertainties	56
4.8	Data Fluctuation	57
4.9	Numerical Setup of the Tests	58
5	Results	62
5.1	Choice of the Minimizer	62
5.2	L0 Results	63
5.3	Uncertainty Characterization of MAPTMD22 and PV19	72
5.4	Bias-Variance Tradeoff and Uncertainty Faithfulness	95
6	Conclusions	99
6.1	Closure Test Results	99
6.2	Future Applications of This Work	101

CONTENTS

Bibliography	107
---------------------	------------

Chapter 1

Introduction

The focus of this thesis is on Transverse Momentum Distributions (TMDs) and their determination through experimental data. TMDs provide a detailed picture of the internal structure of hadrons—such as protons and neutrons—by describing how quarks and gluons move within them in a 3D momentum space [1]. Specifically, we will consider two types of TMDs: the unpolarized Transverse-Momentum Dependent Parton Distribution Functions, referred to as TMD PDFs, and the unpolarized TMD Fragmentation Functions (FFs). The TMD PDFs describe the distribution of quarks and gluons inside hadrons in 3D momentum space, while the TMD FFs represent the dynamical transition from a colored partonic state to a colorless hadronic state during the hadronization process. The accurate and precise determination of TMDs is crucial for enhancing our understanding of the dynamics of hadronic structure and for studying the phenomena of confinement and hadronization. Accurate TMDs can improve the precision of theoretical predictions in high-energy physics, which is essential for testing the Standard Model and for searching for potential new physics beyond it.

The phenomenological determination of TMDs relies on factorization theorems, which are a class of theorems that ensure, under specific conditions, that the cross sections of certain high-energy physics processes can be divided into perturbative and non-perturbative contributions. The perturbative part of the hadronic processes, referred to as the "hard process," can be calculated order by order in perturbative Quantum Chromodynamics (pQCD). The non-perturbative contribution, also known as the "soft process," cannot be predicted using perturbation theory but can be determined through fits to experimental data. We use experimental measurements from Semi-Inclusive Deep Inelastic Scattering (SIDIS, $\ell p \rightarrow \ell h X$) and Drell-Yan ($pp \rightarrow \ell^- \ell^+ X$) cross sections to infer the functional forms of the TMDs via parametric regression techniques. Factorization theorems have been proven for both Drell-Yan and SIDIS cross sections [2].

In this work, we perform a statistical analysis of two fitting frameworks constructed for determining TMDs from experimental data. The resulting TMDs predicted by these two frameworks, which we refer to as the PV19 and MAPTMD22 frameworks, can be found in Refs [3, 4], respectively. Both PV19 and MAPTMD22 were developed by

the MAP (Multi-dimensional Analyses of Partonic Distributions) collaboration and are available for use within the Nanga Parbat suite, which is publicly accessible at <https://github.com/MapCollaboration/NangaParbat>.

The principal limitation of the current fitting methodologies are the following:

1. The inability to assess whether the fitted TMDs are close to the real TMDs we are searching for. This assessment cannot be made by comparing the real cross sections with the predictions calculated using the TMDs inferred from the dataset, as cross sections depend non-trivially on TMDs.
2. The methodology used to infer TMDs from experimental data introduces a component of uncertainty to the predicted TMDs that is not related to the experimental uncertainty. We cannot quantify how much of the total TMD uncertainties arises from the fitting framework used to predict those TMDs.
3. We do not know whether the methodology is capable of reproducing faithful uncertainty estimates and whether it is optimized in terms of the bias-variance tradeoff.

The goal of this thesis is to test the robustness of the PV19 and MAPTMD22 methodologies, specifically addressing the three points mentioned above, through closure testing. Closure testing involves generating pseudo-datasets using known a priori TMD functional forms and performing fits on such pseudo-data with the fitting frameworks we aim to test [5, 6]. This procedure allows us to test three crucial aspects of a fitting framework, that are discussed in the paragraphs below.

First, it permits to directly compare the TMDs used to generate the pseudo-datasets, referred to as the *real* TMDs, with those obtained from fitting the pseudo-data, the *fitted* TMDs, thus assessing the methodology’s ability to capture the underlying real law that we are searching.

Closure tests also permit a qualitative evaluation of the contributions of uncertainties introduced by the fitting frameworks to the total uncertainty. This is achieved by using the fitting framework to fit datasets with different levels of statistical fluctuations and comparing the uncertainties obtained from fits on such differently fluctuated datasets. The levels of fluctuations are Level 0 (no statistical fluctuations), Level 1 (one Monte Carlo replica of the Level 0 dataset), and Level 2 (a large number of Monte Carlo replicas of the Level 1 dataset). By controlling the levels of fluctuation in the data points, we can introduce and remove experimental uncertainty (Level 2) and compare TMD uncertainties with and without the contribution from experimental uncertainty.

Finally, we determine whether the frameworks are optimized and produce faithful uncertainty estimates. In a multi-closure test, we utilize the tested fitting framework to execute a large number of fits, where each fit corresponds to running the framework on a substantial number of Level 2 datasets generated by fluctuating a single Level 1 dataset. This approach enables us to compute the bias and variance of each fit and to

count the fraction of real points (artificially generated) that lie within the one-sigma band of the predicted points from a fit. The averages of bias and variance across the fits are employed to assess the optimization of the bias-variance tradeoff, while counting the real points that fall within the error bars of the predicted ones is used to evaluate the faithfulness of the uncertainties.

In the following, we present an outline of the chapters of this thesis and their content.

- In Chapter 2, we discuss the physics of TMDs. We start with the simplistic parton model, which considers one-dimensional Parton Distribution Functions (the collinear PDFs) without QCD corrections, to explain the basic assumptions that remain valid for factorization theorems and TMDs. We then define Transverse Momentum Distributions (TMDs) and illustrate how they can be used to describe SIDIS and Drell-Yan cross sections through factorization, including a discussion of their divergences and the evolution equations arising from their renormalization.
- In Chapter 3, we focus on the structure of the MAPTMD22 and PV19 fitting frameworks and how they utilize parametric regression to solve the *inverse problem*, which is the problem of finding an input u (the TMD functional forms) for a known model $\mathcal{G}$ (the factorized SIDIS and Drell-Yan cross sections), given a set of results $\mathcal{D} = \mathcal{G}(u)$ (the experimental dataset).
- In Chapter 4, we explore closure testing. We present the general idea, explaining the levels of fluctuations (Level 0, Level 1, and Level 2) and how they can be employed to test the frameworks' flexibility and characterize the final TMD uncertainty. We then delve into multi-closure tests, defining two statistical estimators (the bias-variance ratio, $\mathcal{R}_{bv}$, and the quantile estimator, $\xi_{n\sigma}$) that provide a quantitative assessment of the bias-variance tradeoff optimization and the faithfulness of the TMD uncertainties. Finally, we outline the practical implementation of closure testing in the fitting frameworks we aim to evaluate.
- In Chapter 5, we present the results obtained from performing closure tests on the MAPTMD22 and PV19 frameworks. The results include, for both methodologies, the choice of the minimizer, the assessment of the frameworks' flexibility, the qualitative estimation of the contributions of uncertainties to the total TMD uncertainties, and the quantitative evaluation of the robustness of the frameworks in terms of faithful uncertainty estimates and the bias-variance tradeoff.
- In Chapter 6, we present our conclusions regarding the results obtained from this analysis. We discuss potential directions for improving TMD determination from experimental data based on the closure test results and explore how the computational framework developed in this work can be utilized to validate and optimize future fitting methodologies.

In summary, this thesis aims to improve the determination of Transverse Momentum Distributions (TMDs) by investigating the robustness of the fitting methodologies

PV19 and MAPTMD22 through closure testing. This work contributes to the broader quest for precision in high-energy physics by developing tools that improve the reliability of the predicted TMDs and their associated uncertainty estimates.

Chapter 2

From Parton Model to TMDs

In this chapter, we will explore the basic assumptions of the parton model and construct the definition of transverse-momentum-dependent parton distribution functions (TMD PDFs), which are the physical quantities we aim to infer from experimental data.

2.1 QCD Lagrangian and Asymptotic Freedom

Hadrons, such as protons, are composed of partons, quarks and gluons, which are in a confined state at low energy scale, where for low energy we mean below 1 GeV. The interactions between quarks and gluons are described by quantum chromodynamics (QCD) [7], a Yang-Mills theory that promotes global $SU(3)$ invariance to a local invariance (*i.e.* the QCD Lagrangian is invariant under local color transformation). The QCD Lagrangian is presented in equation (2.1), including only one quark of mass m and excluding ghost fields for simplicity [8],

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu} + \bar{\psi}^i(i\mathcal{D}_j^k - m\delta_j^k)\psi_k, \quad (2.1)$$

where $a = 1, 2, \dots, 8$ is the color index; $\psi, \bar{\psi}$ represent the fermion fields (quarks). The gauge-fields A_μ^a (gluons) are enclosed in the gauge-field strength tensor

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf^{abc}A_\mu^b A_\nu^c, \quad (2.2)$$

and in the covariant derivative D_μ in

$$\mathcal{D} = \gamma^\mu D_\mu = \gamma^\mu(\partial_\mu - igA_\mu^a T^a). \quad (2.3)$$

Expanding all terms of the Lagrangian we obtain the strong interaction terms:

$$\mathcal{L} = \mathcal{L}_0 + gA_\mu^a \bar{\psi}\gamma^\mu T^a \psi + gf^{abc}(\partial_\mu A_\nu^a)A^{b\mu}A^{c\nu} - g^2 f^{eab}f^{ecd}A_\mu^a A_\nu^b A^{c\mu}A^{d\nu}. \quad (2.4)$$

It is interesting to note that the terms in equation (2.4) represent different types of interaction. The first term is the non-interacting contribution, the second term represent the interaction between two leptonic fields and a gluon, and the third and fourth terms represent the interactions between three and four gluon fields, respectively. All

these terms are weighted by powers of the gauge coupling g .

From the QCD Lagrangian, we can deduce the Feynman rules for calculating amplitudes of elementary processes involving quarks and gluons in terms of the strong coupling, $\alpha_s = \frac{g^2}{4\pi}$. If α_s were a small constant, we would have some dominant processes (those with a small number of vertices) and some negligible higher order contributions, similar to the case of quantum electrodynamics (QED) where the amplitudes of processes get smaller with an increasing number of vertices in the corresponding Feynman diagram. Unfortunately, the strong coupling is small only at high energy scales, preventing us from developing a fully perturbative approach. In particular, the dependence of α_s from the energy scale (μ) comes from the renormalization group (RG) equation for the strong coupling [9]:

$$\mu^2 \frac{d\alpha_s}{d\mu^2} = \beta(\alpha_s) = -(b_0\alpha_s^2 + b_1\alpha_s^3 + b_2\alpha_s^4 + \dots) \quad (2.5)$$

If one considers only the one-loop β -function coefficient $b_0 = (33 - 2n_f)/(12\pi)$, with n_f being the number of quark flavors, an analytical solution exist for the strong coupling,

$$\alpha_s(\mu) = \frac{1}{b_0 \ln(\frac{\mu^2}{\Lambda^2})}, \quad (2.6)$$

where $\Lambda \sim 200$ MeV is an integration constant that represent a pole in the analytical solution, that means the strong coupling diverges rapidly for energies below 1 GeV. This phenomenon is known as asymptotic freedom. In figure 2.1 is shown the dependence of the strong coupling from the energy scale μ computed at five-loops and compared with experimental results. At high energy (short distance interactions) the strong coupling becomes very small. As the particles get closer to each other their interaction decreases, but as the distance increases (or, equivalently, the energy decreases), the strong coupling α_s increases.

If we know the value of the strong coupling at a certain energy scale μ , we can use equation (2.5) to compute the value of α_s for a process occurring at another energy $Q \gg 1$ GeV, obtaining:

$$\alpha_s(Q) = \alpha_s(\mu) \frac{1}{1 + \frac{b_0\alpha_s(\mu) \ln(\frac{Q}{\mu})}{2\pi}} \quad (2.7)$$

Asymptotic freedom is crucial for describing hadronic interactions, because it ensures that we can use a perturbative approach to compute the amplitude of short range processes order by order in perturbation theory. However, for a complete description of hadronic interactions, we cannot consider only short distance interactions. We need to introduce some non-perturbative functions to describe the part of the process that cannot be predicted due to non-perturbative effects. This approach was first introduced in the parton model, which will be discussed in the next paragraph.

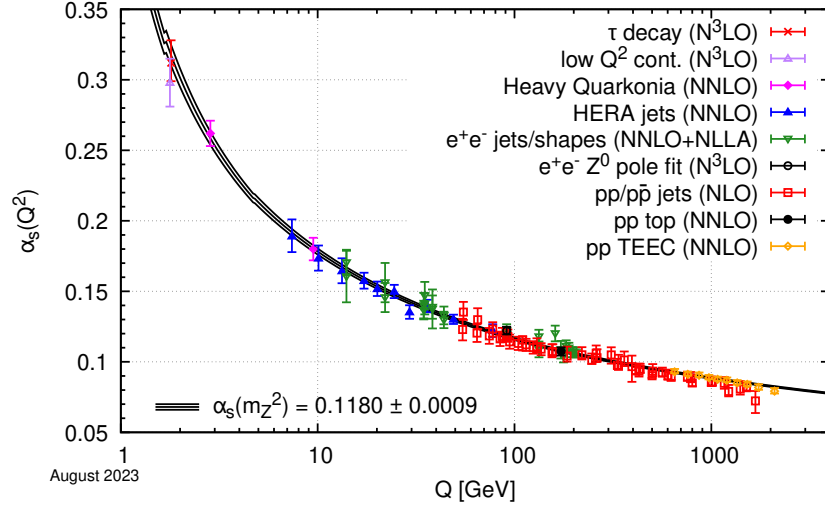


Figure 2.1: Energy dependence of the strong coupling $\alpha_s(\mu)$, image from Ref. [9].

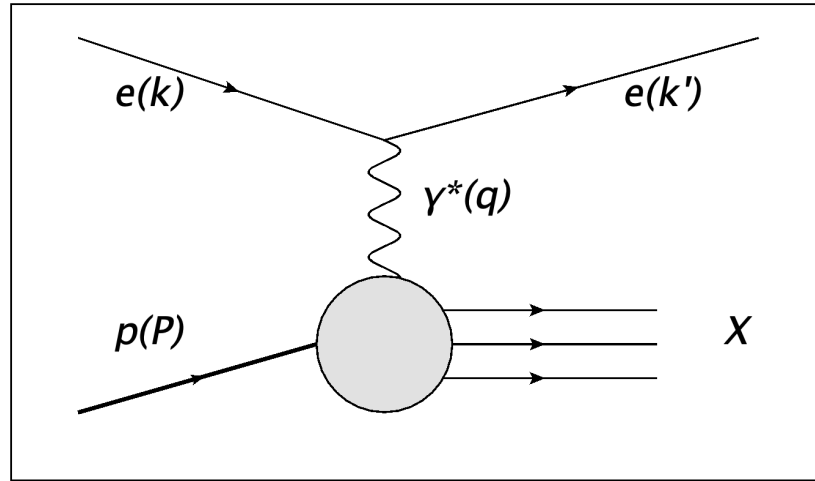
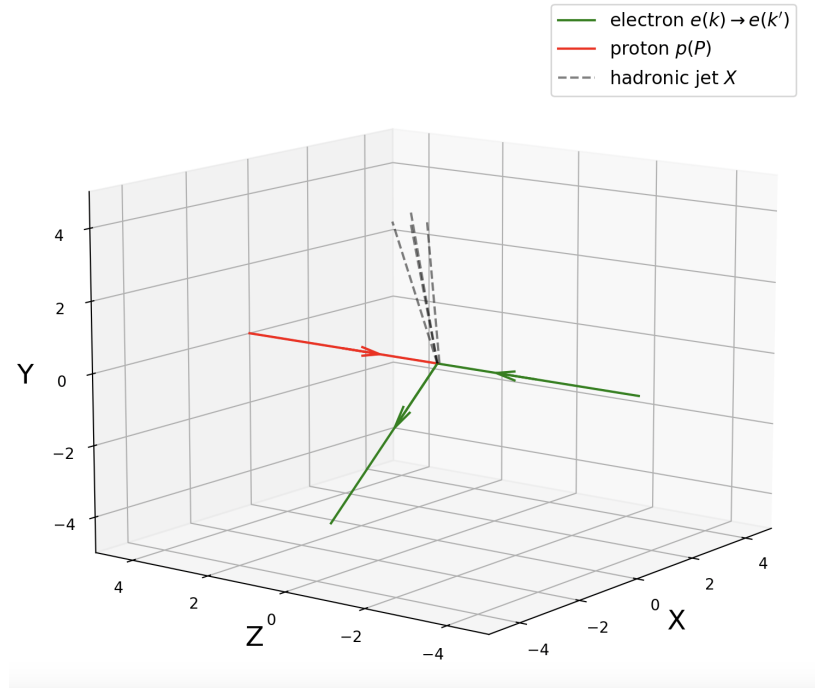
2.2 The Parton Model

Historically, the parton model was introduced before the theory of QCD to explain the results of the SLAC-MIT experiment [10], in which electrons and protons were scattered at energies up to 20 GeV. This scattering experiment led to the confirmation that the proton is not an elementary particle; instead, it is composed of point-like particles of spin $\frac{1}{2}$, which were named partons by Feynman in 1968. Partons were then identified with the quarks of QCD, and the term "partons" is now used to describe both quarks and gluons.

According to the parton model, introduced by Bjorken and Feynman, when an electron interacts with a proton at high energy, the interaction occurs between the electron and one of the proton's fundamental constituents, which carries a fraction ξ of the total longitudinal momentum of the proton. After the parton-electron interaction, the struck quark interacts with the other partons on a much larger time scale, and the proton breaks up.

To explain the basic assumptions of the parton model, consider a process of electron-proton Deep Inelastic Scattering (DIS, $e(k)p(P) \rightarrow e(k')X$), in which a high-energy electron scatters off a proton target (Figure 2.2(a)). We can make some heuristic assumptions for this process:

1. The proton is composed of a collection of point-like particles (partons).
2. The partons all move in the same direction, which is longitudinal to the moving direction of the proton. This is, of course, an approximation; we neglect the transverse components because, in the center of mass reference system, the proton is moving at high speed along only one axis.


 (a) Deep Inelastic Scattering $e(k)p(P) \rightarrow e(k')X$


(b) Spatial part of DIS

Figure 2.2: In DIS, a high-energy electron interacts with partons inside a proton, causing the proton to break up. Only the final-state electron is measured.

3. Each parton carries a fraction ξ of the longitudinal momentum of the proton P . The longitudinal momentum of the parton will be $p = \xi P$.
4. The partons are free particles (they do not interact with each other). The interaction occurs between the electron e and one single parton q_i , which reads $e(k) q_i(p) \rightarrow e(k') q_i(p')$.

The last assumption is crucial and is justified by QCD and asymptotic freedom.

When we consider a high-energy interaction, we are in an energy region in which the strong coupling becomes weak, and the electron-parton interaction dominates (*i. e.* the parton interacts with the electron and does not interact with the other partons in the proton). A more intuitive argument can be given to justify the assumption that quarks are free particles [11, 1]: we consider the time scale of the interaction binding two quarks to be $\tau_1 \sim 1 \frac{fm}{c}$ and the time scale of the quark-electron interaction at high energy Q to be $\tau_2 \sim \frac{1}{Qc}$. In the center of mass frame, the time τ_1 dilates and becomes $\tau'_1 = \tau_1(1 - \frac{v^2}{c^2}) \gg \tau_2$ (in the assumption of $Q \gg 1$ GeV). As an example, if we take an event at $Q^2 = 10^4$ GeV² we get $\tau'_1 \sim 100 \frac{fm}{c}$ and $\tau_2 \sim 0.01 \frac{fm}{c}$. It is therefore a reasonable approximation, in the time scale of the parton-electron interaction, to treat the quarks as "frozen" states with a longitudinal momentum of $p = \xi P$.

Using these assumptions, one can reconstruct the elementary cross section,

$$\hat{\sigma}(e(k) q_i(p) \rightarrow e(k') q_i(p')) , \quad (2.8)$$

order by order in perturbative QCD. For example, the tree-level elementary cross section will be the same as the cross section for the leptonic process $e(k) \mu(p) \rightarrow e(k') \mu(p')$, with the only difference being that the considered quark q_i has a different charge than the muon μ .

Deep Inelastic Scattering Kinematics

Before introducing DIS kinematic, it is useful to change coordinate system and work in the light-front coordinates.

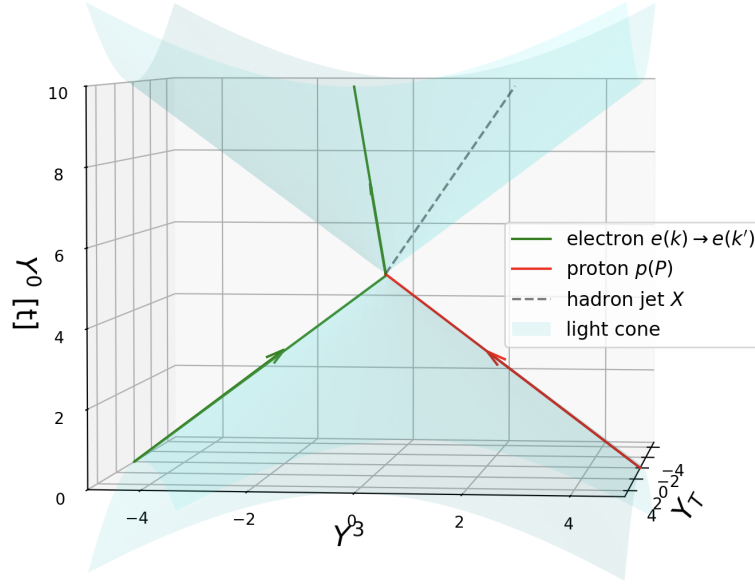
If we define the position four-vector y^μ in space-time coordinate as $y^\mu = (y^0, y^1, y^2, y^3) = (y^0, y^i)$, being y^0 the time component and y^i the spatial components of the vector, we can define the same vector in light-front coordinates as

$$y^\mu = (y^+, y^-, y_T) = \left(\frac{y^0 + y^3}{\sqrt{2}}, \frac{y^0 - y^3}{\sqrt{2}}, y^1, y^2 \right) . \quad (2.9)$$

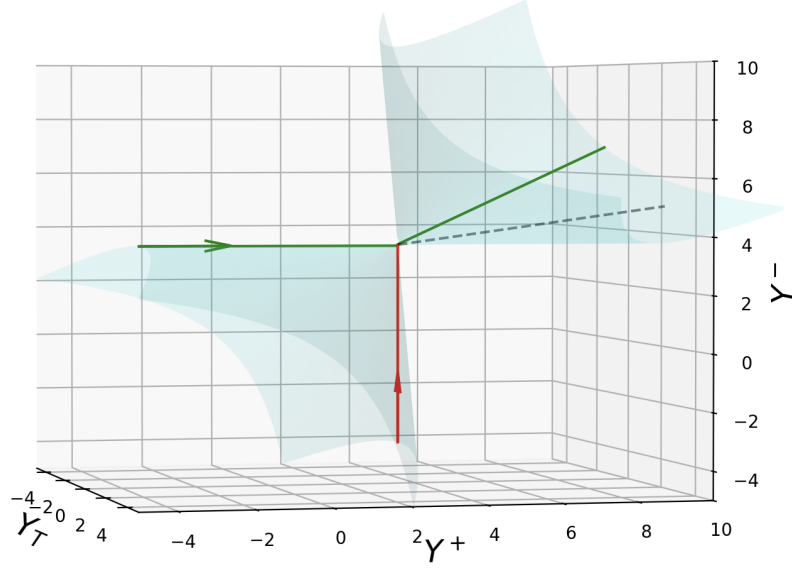
In DIS, y^3 is defined as the direction of motion of the target, and (y^1, y^2) are the transverse directions, which remain unchanged in light-front coordinates: $y_T = (y^1, y^2)$. Figure 2.3 shows a plot of DIS in standard space-time coordinates compared to light-front coordinates. Notably in light-front coordinates the y^+ and y^- axes are parallel to the light cone. More in general, we can define the light-cone basis vector as

$$n_\pm = \frac{1}{\sqrt{2}}(1, 0, 0, \pm 1) \quad (2.10)$$

in which the basis vector $n_\pm^\mu$ can be determined using some physical quantities, as we will see in section 2.4. In this basis, the components n_+ and n_- define a plane, and the transverse components $n_T = (0, 1, 1, 0)$ are transverse to the plane identified by n_+ and n_- .



(a) DIS in space-time coordinates



(b) DIS in light-front coordinates

Figure 2.3: In light-front coordinates, the axes Y^+ and Y^- are parallel to the light cone. In these plots, the particles are approximated to be massless.

To define the kinematics of the process in figure 2.2(a) using light-front coordinates, consider the electron-proton collision

$$e(k)p(P) \rightarrow e(k')X, \quad (2.11)$$

in which X represents a hadronic final state that is not measured. We redefine ξ as the fraction of momentum carried by the parton in the *plus* direction, $\xi = \frac{p^+}{P^+}$, and we can define two Lorentz-invariant quantities: $Q^2 = -t = (-q)^2$ and $P \cdot q$, where P is

the momentum of the initial-state proton, and $q = (k - k')$ is the space-like exchanged momentum. Two useful combinations of these quantities are

$$x = \frac{Q^2}{2P \cdot q} \quad (2.12)$$

$$y = \frac{q \cdot P}{k \cdot P} \quad (2.13)$$

Here, x is the Bjorken variable and, at leading order, represents the fraction of the plus component of momentum carried by the struck parton: $\xi = \frac{p^+}{P^+} = x + \mathcal{O}(\frac{\Lambda^2}{Q^2})$. The variable y represents the fraction of energy lost by the electron during the interaction, $\frac{E-E'}{E}$. Both variables lie between 0 and 1 and can be expressed in terms of the measurable momenta k , k' , and P . We now have all the ingredients to construct the hadronic cross section for deep inelastic scattering.

DIS Cross Section and Parton Distribution Functions

Using perturbative QCD, we are able to compute the elementary cross section

$$\hat{\sigma}(e(k)q_i(p) \rightarrow e(k')q_i(p')) \quad (2.14)$$

with high precision, but we cannot directly measure this short-range interaction. This means that we do not know which parton the electron interacts with, nor do we know the internal momentum distribution of the hadron. To address this problem, we introduce the parton distribution functions (PDFs), $f_i(x)$. These functions represent the probability that the i -th parton carries a fraction x of the *plus* momentum component. Equation (2.15) shows how the total cross section depends on the partonic cross section and the PDFs.

$$\frac{d\sigma}{dt}(e(k)p(P) \rightarrow e(k')X) = \sum_i \int dx f_i(x) \frac{d\hat{\sigma}}{dt}(e(k)q_i(p) \rightarrow e(k')q_i(p')) \quad (2.15)$$

The PDF introduced in equation (2.15) is the link between the total cross section (what we measure) and the partonic cross section (what we can predict in perturbation theory). This is the simplest possible PDF and depends only on the fraction of momentum x ; for this reason, it is named the *collinear* PDF¹. These functions have a physical interpretation as the number density of partons that are in a momentum state $p^+ = xP^+$, so they can also be referred to as parton densities[1].

PDFs represent the non-perturbative part of the process; they can only in part be theoretically predicted, and depend on non-perturbative contributions that must be inferred from experimental data. For illustration, at leading order, the differential cross section in x and Q^2 will be

¹More general PDFs, which we will consider later, can be introduced by accounting for other variables, such as the transverse components of the quark momentum k_T .

$$\frac{d\sigma^{DIS}}{dQ^2 dx} = \sum_i f_i(x) e_i^2 \frac{2\pi\alpha^2}{Q^4} \left[1 + \left(1 - \frac{Q^2}{xs} \right)^2 \right], \quad (2.16)$$

where e_i is the charge of the quark with flavor i , α is the electromagnetic coupling, and s is the Mandelstam variable $(P + k)^2$. From a fit of DIS measurements, we cannot separate the flavor dependence of the PDFs; we can only extract the combination $\sum_i e_i^2 f_i(x)$. However, some theoretical constraints can be applied to characterize the flavor dependence:

- Flavor conservation in the strong interaction, using the valence quark numbers for the proton we obtain:

$$\begin{aligned} \int_0^1 dx [f_u - f_{\bar{u}}] &= 2 \\ \int_0^1 dx [f_d - f_{\bar{d}}] &= 1 \\ \int_0^1 dx [f_s - f_{\bar{s}}] &= 0. \end{aligned} \quad (2.17)$$

- Conservation of the total longitudinal momentum:

$$\int_0^1 x [f_u + f_{\bar{u}} + f_d + f_{\bar{d}} + f_s + f_{\bar{s}} + f_g] dx = 1. \quad (2.18)$$

We have included only the light quarks and gluon densities. Additional information on parton densities can be obtained by considering other processes[8], such as for example electron-neutron scattering ($en \rightarrow eX$), neutrino DIS (e.g., $\nu p \rightarrow \mu^- X$) and Drell-Yan ($pp \rightarrow e^+ e^- X$). To achieve a high-precision global fit of the PDFs, the parton model alone is insufficient; higher-order corrections from perturbative QCD must be taken into account.

Relation Between Static Quark Model and PDFs

Before discussing the QCD corrections, it is useful to examine the relationship between PDFs and the static quark model². Figure 2.4, shows the functional form of the PDFs at $Q = 3.2$ GeV derived by the NNPDF collaboration [12]. In the static quark model, the proton is composed of three valence quarks: two up and one down. If the non-interacting quarks approximation of the parton model were exact, one would expect the PDFs $f_u(x)$ and $f_d(x)$ to be two delta functions peaked at $x = 1/3$, meaning the total longitudinal momentum of the proton is evenly split among the three quarks. In the dynamical picture of the proton captured by the PDFs, one can see a confirmation

²For "static quark model" we refer to the picture in which the hadrons are composed only by the valence quarks. With "static" we mean that the quark number in the proton is not changing. In literature, this picture is also referred to as the "naïve quark model"

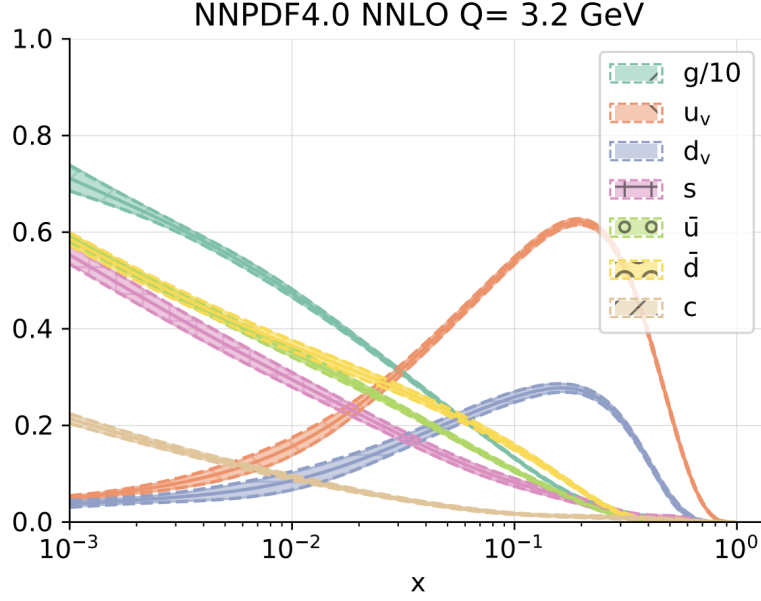


Figure 2.4: PDF set for the proton extracted from a global dataset using NNPDF4.0

of the valence quark idea, where the up and down PDFs are peaked at $x \approx 1/3$. However, the delta functions are smeared out due to quark interactions. We also observe other parton densities that become significant at small x . These are gluon densities, associated with the fact that the quarks are interacting by exchanging gluons, and sea quark densities, due to the radiation of quark-antiquark pairs from valence quarks or excited from the vacuum.

The picture of the proton structure provided by the PDFs is much more detailed, rich, and precise than the static quark model. This is why the study of parton densities is of immense interest and utility for a precise understanding of the dynamical properties of hadrons and more in general for precision high-energy physics.

Collinear and TMD PDFs

Much work has been done in the accurate and precise determination of collinear PDFs (e.g. see the collaborations CTEQ-TEA [13], NNPDF [12], MSHT [14], H1 and ZEUS at HERA [15], and CJ (CTEQ-JLab) [16]). The current challenge is to go beyond the 1D momentum structure of the proton and to capture the transverse-momentum³ partonic distributions. This can be done introducing a new class of parton densities: the transverse-momentum-dependent (TMD) PDFs (also referred to as TMDs, *i.e.* Transverse Momentum Densities). TMDs have an explicit dependence on the transverse momentum components of the quark (k_T) in the light cone basis,

$$f_i(x, k_T^2) , \quad (2.19)$$

³for now on, we will refer to the transverse momentum of the quark in light cone coordinates as k_T

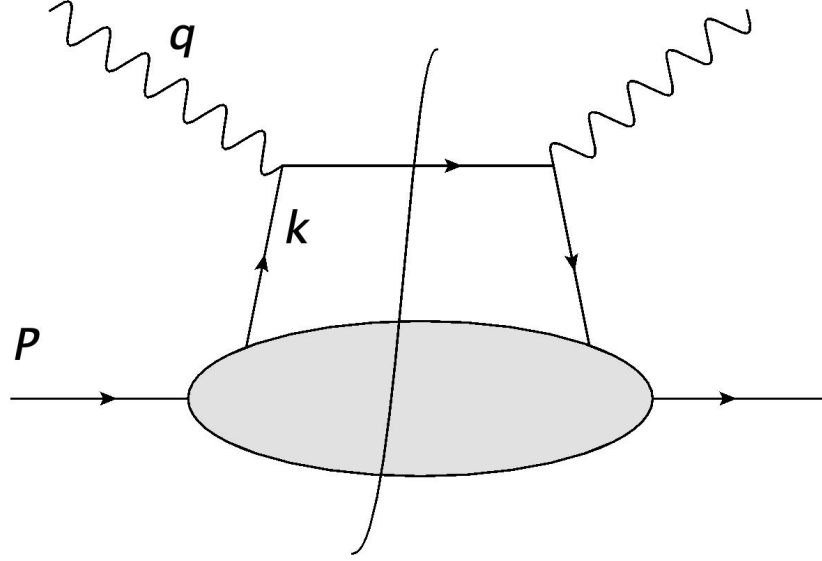


Figure 2.5: The handbag diagram represent the hadronic part of the interaction, that is the scattering between a virtual photon and a single constituent of the hadron.

and characterize the 3D momentum structure of the proton. If one considers the bare (not renormalized) quantities, there is an integral relation between collinear and TMD PDFs [1]:

$$f_i(x) = \int f_i(x, k_T^2) dk_T^2 . \quad (2.20)$$

However, the relationship between the two is not straightforward when considering the renormalized quantities. The reason is that, when QCD perturbative corrections are introduced, divergencies arise and the PDFs acquire a dependence on auxiliary parameters. In the case of collinear PDFs there is a UV divergence that introduces a dependence on the energy scale μ (*i. e.* $f_i(x) \rightarrow f_i(x; \mu)$). For TMDs, a different divergence occurs (discussed in section 2.3), leading the parton densities to depend on the rapidity scale ζ (*i. e.* $f_i(x, k_T^2) \rightarrow f_i(x, k_T^2; \mu, \zeta)$). In the next section, we explore the operator definition of TMDs, their divergences, and their dependence on auxiliary parameters [1, 17, 18].

2.3 Operator definition of TMDs

The definition of physical observables in terms of parton densities relies on factorization theorems [19, 2], which ensure that we can separate the perturbative and non-perturbative parts in the theoretical prediction for the hadronic cross section of certain processes.

To implement factorization theorems, we need a correct and working definition of TMDs that includes field theory. If one assumes that the dominant part of the process

has the form of the handbag diagram (figure 2.5), where the interacting quark has limited transverse momentum and virtuality [1], a simple definition of the partonic density in 3D momentum space is

$$f_i(x, k_T) = \frac{1}{(2\pi)^3 2x \langle p|p \rangle} \langle p| a_i^\dagger a_i(xp^+, k_T, \lambda) |p \rangle \quad (2.21)$$

where $a_i(k^+, k_T, \lambda)$, $a_i^\dagger(k^+, k_T, \lambda)$ are the creation and annihilation operators of a parton with flavor i and helicity λ in *light-front* coordinates. In light-front coordinates, one can express the parton density of equation (2.21) in terms of a quark correlation function:

$$f_i(x, k_T) = \int \frac{dy^- d^2 y_T}{16\pi^3} e^{-ip^+ y^- + ik_T y_T} \langle p| \bar{\psi}_i(0, y^-, y_T) \gamma^+ \psi_i(0) |p \rangle, \quad (2.22)$$

where the fields combination $\bar{\psi}\psi$ is related to the creation/annihilation operators $a^\dagger a$ through second quantization, and the γ^+ matrix is used to project the state on the *unpolarized* parton density $f_i(x, k_T)$. For a review which includes other classes of parton densities, that we do not consider in this thesis, see Refs. [20, 21].

Divergences of TMD PDFs

Equations (2.21) and (2.22) have the correct structure; however, they are not gauge invariant. To give a working definition of transverse momentum densities, we can work in a fixed gauge, specifically the *light-cone gauge*, where we define $A^+ = 0$. In this gauge, a divergence arises: the light-cone gauge divergence, which is associated with the $1/k^+$ singularities of the gluon propagator and has a physical interpretation as the region where the gluon is moving infinitely fast in the direction of the quark jet (the region in which the gluon rapidity goes to minus infinity). Collins and Soper have shown that the light-cone gauge divergence disappears in integrated parton densities (collinear PDFs). However, the integrated PDFs suffer from a UV divergence, associated with the divergence of the integral over k_T at large transverse momentum.

Different approaches have been proposed to cancel the light-cone gauge divergency from TMDs. We will use the definition in [1], which involves inserting non-light-like Wilson lines in the quark correlation function. Equation (2.22) then becomes dependent from both renormalization scale μ and rapidity scale $\zeta = p^2 \cosh^2(\Delta y)$. The auxiliary parameter ζ arises from a cutoff on gluon rapidity, Δy represents the difference in rapidity between the target hadron and the gluon rapidity cutoff. The renormalized TMDs are defined as follows:

$$f_i(x, k_T; \mu, \zeta) = \int \frac{dy^- d^2 y_T}{16\pi^3} e^{-ip^+ y^- + ik_T y_T} \tilde{P}_i(y^-, y_T; \mu, \zeta) \quad (2.23)$$

$$\tilde{P}_i(y^-, y_T; \mu, \zeta) = \langle p| \bar{\psi}_i(0, y^-, y_T) W_y(u)^\dagger I_{u;y,0} \gamma^+ W_0(u) \psi_i(0) |p \rangle_R \quad (2.24)$$

$$W_y(u) = P \left[-ig_{(0)} \int_0^\infty d\lambda u^\mu A_\mu^{(0)}(y + \lambda u) \right] \quad (2.25)$$

This definition of TMDs is gauge-invariant due to the inclusion of Wilson lines. The line in $W_y(u)$ extends from the position y^μ to infinity in the direction u^μ . The line $I_{u;y,0}$ connects $W_y(u)$ and $W_0(u)$ at infinity.

TMDs defined in equation (2.23) are applicable in factorization theorems and can be used to factorize the processes that we want to study. Since this definition depends on two auxiliary parameters, μ and ζ , it may appear to conflict with the independence of TMDs from the kinematic region in which we consider them. This independence is restored through the use of evolution equations for TMDs, developed by Collins, Soper, and Sterman, which allow for the reconstruction of parton densities at any scale (μ, ζ) given the TMDs at an initial scale (μ_0, ζ_0) , this is not a strict independence by the kinematic region, but the dependence is known exactly, and the information about TMDs in a single kinematic point is, in principle, sufficient to reconstruct the TMDs at any other point. These evolution equations differ from the DGLAP equations, which describe the energy dependence of the collinear PDFs. This is one of the reasons why the correspondence between integrated (collinear) and non-integrated (TMD) PDFs is not straightforward.

The evolution equations for TMDs stem from the factorization of the processes we wish to study. Since the fitting frameworks that will be tested in the next chapters are based on Drell-Yan and Semi-Inclusive Deep Inelastic Scattering (SIDIS) cross sections, the next section presents the factorization and evolution of these two processes in terms of TMDs, following the definitions used in the two analyzed frameworks: PV19 [3] and MAPTMD22 [4].

2.4 SIDIS and Drell-Yan Factorization

Three processes can be described in terms of Transverse Momentum Dependent Parton Distribution Functions (TMDs): Semi-Inclusive Deep Inelastic Scattering (SIDIS), Drell-Yan, and lepton-antilepton annihilation with two measured hadrons in the final state ($\ell^+ \ell^- \rightarrow h_1 h_2 X$). Since experimental data for the last process are not available yet, the MAPTMD22 and PV19 frameworks work only with SIDIS and Drell-Yan data, and we will comment on the factorization formulas for the first two processes only. The proofs of the factorization theorems for SIDIS and Drell-Yan can be found in [2].

Semi Inclusive Deep Inelastic scattering (SIDIS)

In SIDIS, a high-energy electron $e^-(l)$ scatters off a proton $p(P)$. The difference from DIS is that in SIDIS we measure the final state electron $e^-(l')$ and a final state hadron $h(P_h)$.

$$e^-(l) + p(P) \rightarrow e^-(l') + h(P_h) + X, \quad (2.26)$$

where l , P , and P_h denote the four momenta of the electron, the proton and the final state hadron, respectively. In this work, we consider only electron-proton SIDIS, but in general, SIDIS refers to the collision between a lepton and a hadron. The process

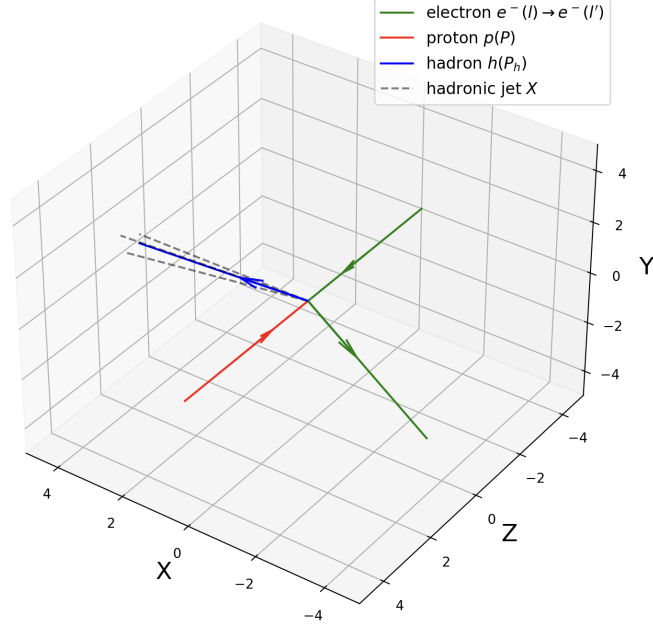


Figure 2.6: Initial and final particles involved in SIDIS in 3D space

is semi-inclusive because only the events with the hadron $h(P_h)$ in the final state are considered. The 3D spatial plot⁴ of SIDIS is represented in figure 2.6.

The kinematic variables of SIDIS are similar to those of DIS, as defined in section 2.2,

$$x = \frac{Q^2}{2P \cdot q} \quad y = \frac{P \cdot q}{P \cdot l} . \quad (2.27)$$

The four-momentum of the exchanged photon is $q = (l - l')$, with $(Q^2 = -q^2)$. In SIDIS, we also measure a final state hadron, which allows us to define a new variable,

$$z = \frac{P \cdot P_h}{P \cdot q} . \quad (2.28)$$

The variables x and z have a similar physical interpretation in light-front coordinates, being $x = \frac{k^+}{P^+}$ the fraction of momentum of the proton in the n_+ axis that is carried by the struck parton, and $z = \frac{P_h^-}{k^-}$ the fraction of the quark momentum along the n_- axis that is transferred to the measured hadron. We can define two other useful quantities, q_T and P_{hT} , in two different frames as shown in figure 2.7.

First, consider the frame where P and q form a light-cone basis (Breit frame). In this frame, we define P_{hT} as the hadron transverse momentum, perpendicular to the plan identified by the momenta of the photon (q) and of the proton (P). Then, we

⁴An interactive 3D plot of SIDIS can be obtained by running the script `DIS_Kinematic.py` contained in the GitHub repository https://github.com/kamillaurent/thesis_laurent

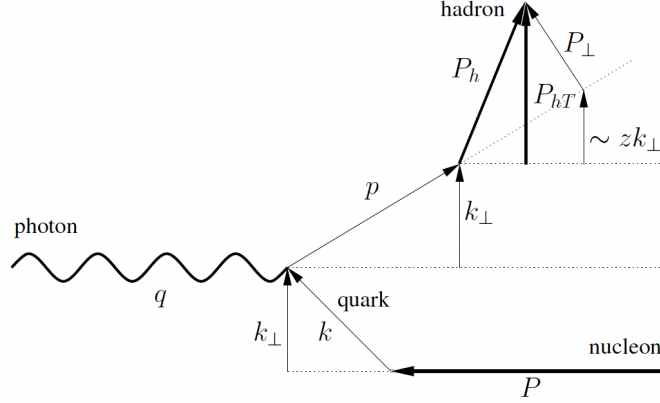


Figure 2.7: Transverse momenta in the hadronic part of SIDIS, image from [4]

define q_T as the photon transverse momentum in the frame where P and P_h form a light-cone basis. The relation between them is

$$q_T^\mu = -\frac{P_{hT}^\mu}{z} - 2x \frac{|q_T|^2}{Q^2} P^\mu. \quad (2.29)$$

In the limit of large photon invariant mass, M^2 , M_h^2 , q_T , $P_{hT}^2 \ll Q^2$, which we will use henceforth, equation (2.29) reduces to:

$$q_T^\mu = -\frac{P_{hT}^\mu}{z} \quad (2.30)$$

In figure 2.7, the parton has transverse momentum $k_\perp$ with respect to the photon direction; $k_\perp$ remains unchanged after the interaction. The final quark with total momentum $p = k + q$ fragments into a hadron with momentum P_h , of which we measure the component P_{hT} .

In Ref. [4], the limit of low q_T ($q_T \ll Q$) is considered. Notably, in SIDIS q_T is not the transverse momentum with respect to the scattering axis, but with respect to the final state hadron axis. In this limit, the differential SIDIS cross section is defined as follows:

$$\frac{d\sigma^{SIDIS}}{dx dz d|q_T| dQ} = \frac{8\pi^2 \alpha^2 z^2 |q_T|}{x Q^3} \left[1 + \left(1 - \frac{Q^2}{xs} \right)^2 \right] F_{UU,T}(x, z, |q_T|, Q) + \text{suppressed factors}, \quad (2.31)$$

where $F_{UU,T}(x, z, |q_T|, Q)$ is the unpolarized SIDIS structure function. It describes the interaction of an unpolarized electron with an unpolarized target through a transversely polarized photon. The structure function depends on the unpolarized TMDs:

$$F_{UU,T}(x, z, |q_T|, Q) = x \mathcal{H}^{SIDIS}(Q, \mu) \sum_a e_a^2 \int d^2 k_\perp \int \frac{d^2 P_\perp}{z^2} f_1^a(x, k_\perp^2; \mu, \zeta) D_1^{a \rightarrow h}(z, P_\perp^2; \mu, \zeta) \delta^{(2)}(k_\perp + P_\perp/z + q_T) . \quad (2.32)$$

Equation (2.32) follows from the SIDIS factorization. $\mathcal{H}^{SIDIS}(Q, \mu)$ represents the hard part of the process and encapsulates all the information related to the short-range interactions, which can be computed perturbatively order by order in QCD. The sum over a includes all active quarks and antiquarks. The variable $k_\perp$ denotes the transverse momentum of the struck parton relative to the proton's axis, while $P_\perp$ denotes the transverse momentum of the hadron relative to the axis of the fragmenting quark. The function $f_1^a(x, k_\perp^2; \mu, \zeta)$ is the unpolarized transverse momentum-dependent (TMD) parton distribution function (PDF) for the parton a within the proton. The TMD evolution equations, along with the explicit dependence on the μ and ζ scales, will be discussed later. Additionally, the SIDIS structure function depends on another type of TMD that has not been defined previously: the unpolarized TMD fragmentation function (FF) $D_1^{a \rightarrow h}(z, P_\perp^2; \mu, \zeta)$. This function describes the dynamical transition between a colored state (the struck parton a) and a colorless hadron $h(P_h)$. Similar to the TMD PDFs, the TMD FF is a transverse-momentum density and depends on both ζ and μ renormalization scales. The condition $\delta^{(2)}(k_\perp + P_\perp/z + q_T)$ ensures momentum conservation in the transverse plane.

It is convenient to express the structure function in terms of the Fourier transforms of the two TMDs involved in the process,

$$\hat{D}_1^{a \rightarrow h}(z, b_T^2; \mu, \zeta) = \int \frac{d^2 P_\perp}{z^2} e^{-ib_T \cdot P_\perp/z} D_1^{a \rightarrow h}(z, P_\perp^2; \mu, \zeta) \quad (2.33)$$

$$\hat{f}_1^a(x, b_T^2; \mu, \zeta) = \int d^2 k_\perp e^{-ib_T \cdot k_\perp} f_1^a(x, k_\perp^2; \mu, \zeta) . \quad (2.34)$$

The structure function in b_T space can then be written as

$$F_{UU,T}(x, z, |b_T|, Q) = \frac{x}{2\pi} \mathcal{H}^{SIDIS}(Q, \mu) \sum_a e_a^2 \int_0^\infty d|b_T| |b_T| J_0(|b_T||q_T|) \hat{f}_1^a(x, b_T^2; \mu, \zeta) \hat{D}_1^{a \rightarrow h}(z, b_T^2; \mu, \zeta) , \quad (2.35)$$

where $J_0(|b_T||q_T|)$ is the Bessel function of the first kind, which can be expressed in integral form as $J_0(x) = \frac{1}{2\pi} \int_0^{2\pi} d\theta e^{ix \cos \theta}$. The formula used in the MAPTMD22 fitting framework for the SIDIS cross section can be obtained by inserting the structure function from equation (2.35) into the cross section defined in equation (2.31). This formula is not used in the PV19 framework, as SIDIS data are not included in that case.

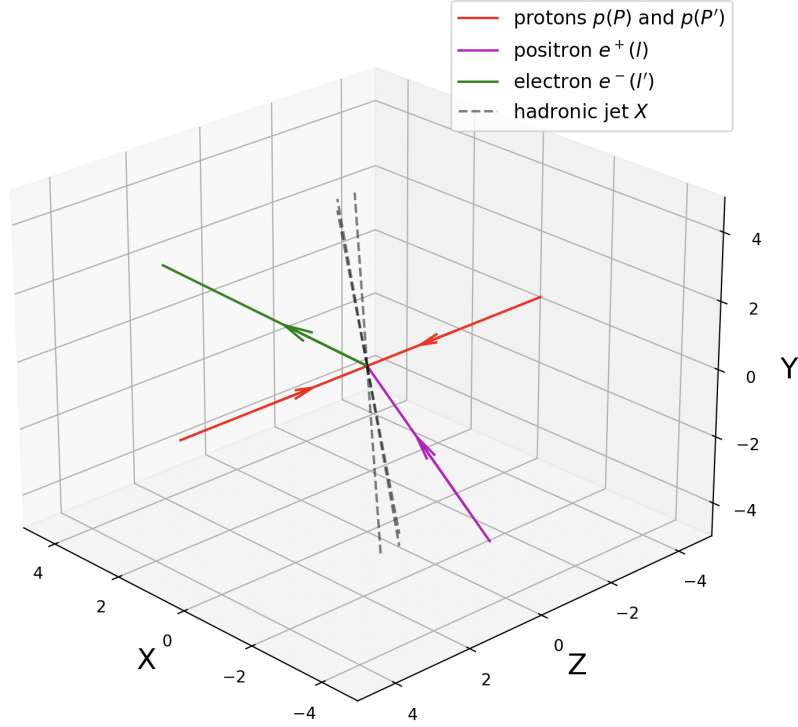


Figure 2.8: Initial and final particles involved in Drell-Yan in 3D space

Drell-Yan

Drell-Yan data are included in both the MAPTMD22 and PV19 frameworks. While the formalism used for factorizing the cross section is the same in both cases, the parametrization of the non-perturbative part of the TMDs differs, which will be discussed in the next chapter.

In the Drell-Yan process, two protons⁵ $p_1(P_A)$ and $p_2(P_B)$ scatter at a high center-of-mass energy $\sqrt{s} = \sqrt{(P_A + P_B)^2}$:

$$p_1(P_A) + p_2(P_B) \rightarrow \gamma^*/Z(q) + X \rightarrow \ell^-(l) + \ell^+(l') + X. \quad (2.36)$$

According to factorization theorems, the short-range interaction occurs between two partons, a quark a and an antiquark $\bar{a}$, which annihilate to produce a neutral vector boson (either a photon or a Z boson) with a large invariant mass $Q = \sqrt{q^2}$, where q is the four-momentum of the exchanged particle. The boson subsequently decays into a lepton-antilepton pair; in this case, we consider a final state consisting of an electron $e^-(l)$ and a positron $e^+(l')$. Conservation of four-momentum applies in the decay, so $q^2 = (l + l')^2$.

⁵More generally, the definition of Drell-Yan includes any two hadrons in the initial state. Here, we specifically consider the hadrons to be protons, as we are considering the TMDs of protons.

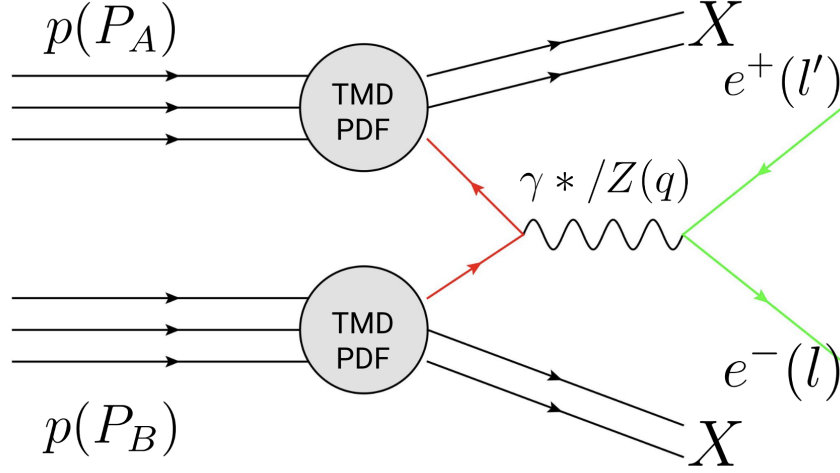


Figure 2.9: Diagrammatic description of the Drell-Yan short-range interaction

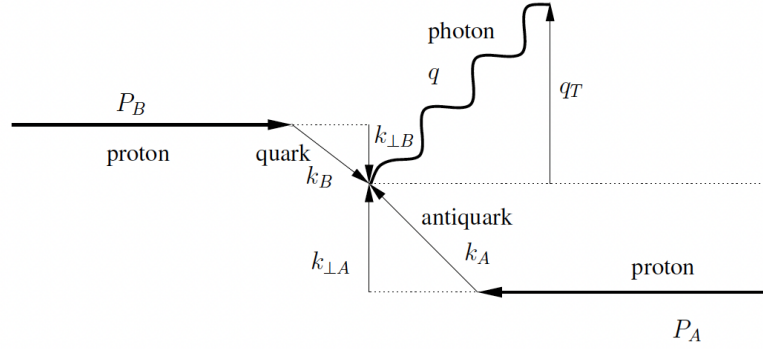


Figure 2.10: Relevant Drell-Yan momenta

A 3D plot of the Drell-Yan process, illustrating the initial and final particles involved, is shown in figure 2.8. The short range interaction is shown in figure 2.9, and the relevant momenta of the hadronic part in the Drell-Yan process are shown in figure 2.10.

We define the rapidity of the vector boson (which is equivalent to the rapidity of the lepton pair) as

$$y = \frac{1}{2} \ln \left(\frac{q_0 + q_z}{q_0 - q_z} \right), \quad (2.37)$$

where q_z is the momentum component along the collision axis z . The transverse momentum of the vector boson is given by:

$$q_T^2 = q_x^2 + q_y^2. \quad (2.38)$$

Similar to the SIDIS process, we consider the limit of low transverse momentum ($q_T \ll$

Q), but here, the transverse momentum is defined relative to the collision axis z . Additionally, we assume $M \ll Q$, where M is the mass of the incoming hadrons. In this limit, the Drell-Yan cross section takes the form:

$$\frac{d\sigma^{DY}}{d|q_T|dydQ} = \frac{16\pi^2\alpha^2|q_T|}{9Q^3} \mathcal{P} F_{UU}^1(x_A, x_B, |q_T|, Q) + \text{suppressed factors}, \quad (2.39)$$

where α is the electromagnetic coupling, and $\mathcal{P}$ is the phase space factor that accounts for cuts in the leptonic kinematics. The structure function $F_{UU}^1(x_A, x_B, |q_T|, Q)$ can be expressed as the convolution of two TMD PDFs:

$$F_{UU}^1(x_A, x_B, |q_T|, Q) = x_A x_B \mathcal{H}^{DY}(Q, \mu) \sum_{a, \bar{a}} C_{a/\bar{a}}(Q^2) \int d^2k_{\perp A} d^2k_{\perp B} f_1^a(x_A, k_{\perp A}^2; \mu, \zeta_A) f_1^{\bar{a}}(x_B, k_{\perp B}^2; \mu, \zeta_B) \delta^{(2)}(k_{\perp A} + k_{\perp B} - q_T), \quad (2.40)$$

where $\mathcal{H}^{DY}(Q, \mu)$ represents the perturbative part of the process, which is computed order by order in the strong coupling α_s . The sum runs over all active quarks and antiquarks ($a, \bar{a}$), and $C_{a\bar{a}}(Q^2)$ are the electroweak charges at the scale Q . The convolution of the two TMDs, f_1^a and $f_1^{\bar{a}}$, describes the quark-antiquark annihilation. In the above approximation, the longitudinal momentum fractions carried by the quarks are:

$$x_A = Q \frac{e^y}{\sqrt{s}} \quad (2.41)$$

$$x_B = Q \frac{e^{-y}}{\sqrt{s}}. \quad (2.42)$$

We can define the structure function using the Fourier transform of the TMD (from equation (2.33)b) as

$$F_{UU}^1(x_A, x_B, |b_T|, Q) = \frac{x_A x_B}{2\pi} \mathcal{H}^{DY}(Q, \mu) \sum_{a, \bar{a}} C_{a/\bar{a}}(Q^2) \int_0^\infty d|b_T| |b_T| J_0(|b_T||q_T|) \hat{f}_1^a(x_A, b_T^2; \mu, \zeta_A) \hat{f}_1^{\bar{a}}(x_B, b_T^2; \mu, \zeta_B). \quad (2.43)$$

By inserting the structure function from equation (2.43) into the definition of the Drell-Yan differential cross section given in equation (2.39), we obtain the expression used in both fitting frameworks for the Drell-Yan cross section.

2.5 TMD Evolution Equations

Up to this point, we have explored the definitions of Transverse Momentum Dependent (TMD) Parton Distribution Functions (PDFs) and Fragmentation Functions (FFs), and

how these parton densities relate to the cross sections of the processes in our dataset (*i.e.*, Drell-Yan and Semi-Inclusive Deep Inelastic Scattering (SIDIS)). To perform a global fit that includes data across different kinematic regions, we need to define the evolution equations for the TMDs.

The dependence of the Fourier transforms of the TMDs, $\hat{f}_1^a$ and $\hat{D}_1^{a \rightarrow h}$, on the scales μ and ζ is governed by two equations [2, 3]:

$$\frac{\partial \ln \hat{f}_1^a}{\partial \ln \mu} = \gamma(\mu, \zeta) \quad (2.44)$$

$$\frac{\partial \ln \hat{f}_1^a}{\partial \ln \sqrt{\zeta}} = K(\mu), \quad (2.45)$$

where γ and K are the anomalous dimensions associated with the renormalization group (RG) evolution in μ and the Collins-Soper evolution in $\sqrt{\zeta}$, respectively. In equations (2.44), we have included only the TMD PDFs, but the same dependence on the auxiliary parameters applies to the TMD FFs. Furthermore, the rapidity anomalous dimension K obeys the RG equation

$$\frac{\partial \ln K}{\partial \ln \mu} = -\gamma_K(\alpha_s(\mu)), \quad (2.46)$$

where γ_K is the cusp anomalous dimension. Since

$$\frac{\partial^2 \ln \hat{f}_1^a}{\partial \ln \sqrt{\zeta} \partial \ln \mu} = \frac{\partial K(\mu)}{\partial \ln \mu}, \quad (2.47)$$

we obtain the equation:

$$\frac{\partial \gamma}{\partial \ln \sqrt{\zeta}} = -\gamma_K(\alpha_s(\mu)) \quad (2.48)$$

which has the solution:

$$\gamma(\mu, \zeta) = \gamma_F(\alpha_s(\mu)) - \gamma_K(\alpha_s(\mu)) \ln \frac{\sqrt{\zeta}}{\mu}. \quad (2.49)$$

In equation (2.49) we used the condition $\zeta = \mu^2$, derived from the condition $\zeta_A \zeta_B = Q^4$. The cross sections for SIDIS and Drell-Yan depend only on the combination $\zeta_A \zeta_B$, allowing us to choose different values for ζ_A and ζ_B . The natural choice is to set $\zeta_A = \zeta_B = Q^2$ for the rapidity scale and $\mu = Q$ for the energy scale, leading to $\zeta = \mu^2$. We also set $\gamma_F(\alpha_s(\mu)) \equiv \gamma(\mu, \mu^2)$.

By solving the RG equations for f_1^a and K , we determine the evolution of the TMDs, as given in equation (2.50); *i.e.*, we can relate the TMD at some final scale (μ_f, ζ_f) to the TMD at an initial scale (μ_i, ζ_i) using

$$\hat{f}_1^a(x, b_T^2; \mu_f, \zeta_f) = \hat{f}_1^a(x, b_T^2; \mu_i, \zeta_i) \exp \left[\int_{\mu_i}^{\mu_f} \frac{d\mu}{\mu} \gamma(\mu, \zeta_f) \right] \left(\frac{\zeta_f}{\zeta_i} \right)^{\frac{K}{2}(|b_T|, \mu_i)}. \quad (2.50)$$

In the limit where $|b_T| \rightarrow 0$, we can match the unpolarized TMDs with the unpolarized collinear PDFs and FFs. In particular, we can express the TMD as a convolution of the collinear PDF and a matching coefficient function C over the fraction of longitudinal momentum x' ,

$$\hat{f}_1^a(x, b_T^2; \mu_i, \zeta_i) = \sum_b \int_x^1 \frac{dx'}{x'} C^{ab}(x', b_T^2; \mu_i, \zeta_i) f_1^b\left(\frac{x}{x'}; \mu_i\right) \equiv [C \otimes f_1](x, b_T^2; \mu_i, \zeta_i), \quad (2.51)$$

where the matching coefficients C are calculated as perturbative expansions in α_s . The terms γ and K in equation (2.50) can also be expanded in powers of α_s , making it convenient to explicitly separate the perturbative parts from the non-perturbative content in the TMDs.

TMD Perturbative and Non-Perturbative Content

Equation (2.50) describes the dependence of the TMDs on the renormalization and rapidity scales. To define the TMDs properly, we must fix the initial and final scales, (μ_i, ζ_i) and (μ_f, ζ_f) , so that each perturbative term in the evolution equation can be computed accordingly. Choosing the scale $\mu_b \sim \frac{1}{b_T}$ and $\zeta = Q^2$ leads to a logarithmic behavior, $\ln\left(\frac{Q}{\mu_b}\right)$ in the perturbative terms of the TMDs [22]. To set the initial scale, we observe that, for any fixed value of b_T , there exists a scale μ_b at which the rapidity anomalous dimension K and the matching coefficient C can be expressed as a pure perturbative expansion in α_s , without the logarithmic dependence. This scale reads

$$\mu_b(b_T) = 2 \frac{e^{-\gamma_E}}{b_T} \quad (2.52)$$

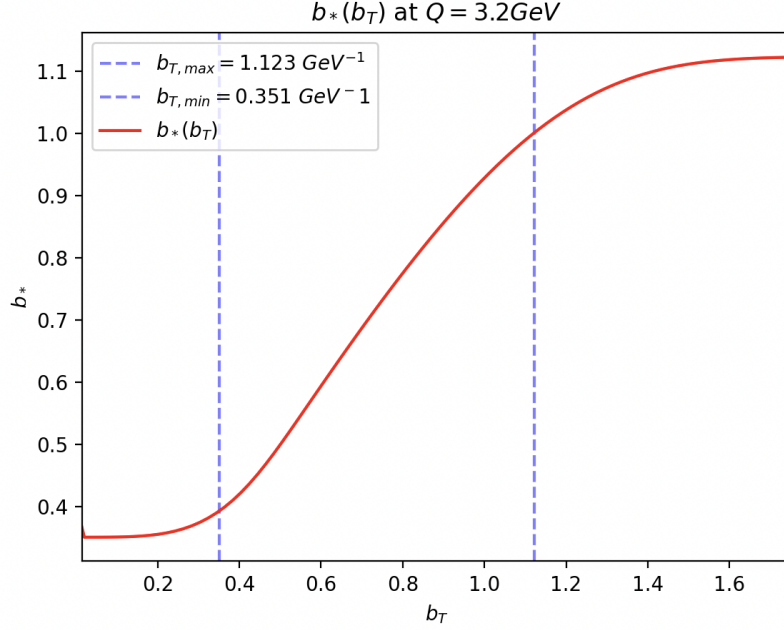
where γ_E is the Euler-Mascheroni constant. We choose μ_b as the initial scale for the evolution equation, setting $\mu_i = \sqrt{\zeta_i} = \mu_b$.

For the final scale, we use the energy scale of the process, setting $\mu_f = \sqrt{\zeta_f} = Q$. Once the initial and final scales for the evolution are set, we can compute the different perturbative parts of the TMDs order by order in the strong coupling constant α_s . The perturbative expansions of the matching function C^{ab} , and of the anomalous dimensions K , γ_K , and γ_F , are reported in ref [3].

When computing the perturbative part of the TMDs, as b_T increases, the scale μ_b decreases (as seen in equation (2.52)), eventually reaching the Landau pole at Λ_{QCD} . As μ_b approaches Λ_{QCD} , $\alpha_s(\mu_b)$ becomes larger and larger, making it impossible to perform a perturbative expansion of C and K . To avoid this problem, we set a maximum value for b_T beyond which the perturbative expansions of K and C are no longer valid:

$$b_{T,max} = 2e^{-\gamma_E} \approx 1.123 \text{ GeV}^{-1}, \quad (2.53)$$

note that this is the choice made in MAPTMD22 and PV19, but other choices are available. To ensure that μ_b does not exceed the energy scale of the process Q , we also set a minimum value:


 Figure 2.11: Functional form of b_* at $Q = 3.2 \text{ GeV}$

$$b_{T,min} = \frac{2e^{-\gamma_E}}{Q} . \quad (2.54)$$

In order to avoid the large and small b_T regions, we then replace the parameter b_T in equation (2.50) with a new parameter $b_*(b_T)$ (figure 2.11), which is constrained between $b_{T,min}$ and $b_{T,max}$ by construction:

$$b_*(b_T) = b_{T,max} \left(\frac{1 - e^{-\left(\frac{b_T}{b_{T,max}}\right)^4}}{1 - e^{-\left(\frac{b_T}{b_{T,min}}\right)^4}} \right)^{\frac{1}{4}} . \quad (2.55)$$

Figure 2.11 shows the asymptotic behavior of b_* , which approaches $b_{T,max}$ as $b_T \rightarrow \infty$ and $b_{T,min}$ as $b_T \rightarrow 0$.

Using b_* instead of b_T , we obtain a TMD Fourier transform $\hat{f}_1^a(x, b_*; \mu_f, \zeta_f)$ that is confined in the perturbative region of b_T . We can then rewrite the TMD in terms of this perturbative-TMD, obtaining:

$$\hat{f}_1^a(x, b_T^2; \mu_f, \zeta_f) = \left[\frac{\hat{f}_1^a(x, b_T^2; \mu_f, \zeta_f)}{\hat{f}_1^a(x, b_*; \mu_f, \zeta_f)} \right] \hat{f}_1^a(x, b_*; \mu_f, \zeta_f) \equiv f_{1NP}(x, b_T^2; \zeta_f) \hat{f}_1^a(x, b_*; \mu_f, \zeta_f) . \quad (2.56)$$

In this way, we isolate the non-perturbative component f_{1NP} , which dominates at large values of b_T and can be determined through fits to experimental data. By inserting

equation (2.56) into equation (2.51), we obtain the non-perturbative corrections to the TMD,

$$\hat{f}_1^a(x, b_T^2; \mu_f, \zeta_f) = [C \otimes f_1](x, b_*; \mu_{b_*}, \mu_{b_*}^2) \exp \left[\int_{\mu_{b_*}}^{\mu_f} \frac{d\mu}{\mu} \gamma(\mu, \zeta_f) \right] \left(\frac{\zeta_f}{\mu_{b_*}^2} \right)^{\frac{K}{2}(b_*, \mu_{b_*})} f_{1NP}(x, b_T^2; \zeta_f, Q_0) . \quad (2.57)$$

In equation (2.57), the Collins-Soper kernel K is also modified to account for non-perturbative corrections:

$$K(|b_T|, \mu_{b_T}) = K(b_*, \mu_{b_*}) + g_K(b_T^2) . \quad (2.58)$$

The functional forms of the corrections f_{1NP} and $g_K(b_T^2)$ are arbitrary and will be discussed in the next chapter. A similar reasoning applies to the Fragmentation Function $D_1^{a \rightarrow h}$, where a non-perturbative correction $D_{1NP}^{a \rightarrow h}(z, b_T^2; \zeta_f, Q_0)$ can be introduced.

The goal of the fitting frameworks that we will analyze is to capture the non-perturbative parts of the TMD PDFs and FFs. In the next chapter, we will discuss how $f_{1NP}(x, b_T^2; \zeta_f, Q_0)$ and $D_{1NP}^{a \rightarrow h}(z, b_T^2; \zeta_f, Q_0)$ are parameterized, and examine the structures of the MAPTMD22 and PV19 frameworks and how they are used to infer TMDs from experimental data.

Chapter 3

TMD Extraction from Experimental Data

This chapter discusses the methodology used by the frameworks MAPTMD22 and PV19 to extract the non-perturbative part of the TMDs from experimental measurements of Drell-Yan and SIDIS cross sections. We will begin by defining inverse problems in general and then explaining how the extraction of TMDs from experimental data relates to these problems.

3.1 The Inverse Problem

Determining TMDs from experimental data is an example of an *inverse problem*: the challenge of finding the input to a given model based on a set of observations [5]. In our case, the model is the factorized cross section¹, the observations are the experimental measurements, and the inputs are the non-perturbative TMDs. Let's now define the inverse problem.

Suppose we have a known function G that takes an unknown input $u \in X$ and produces a measurable output $r \in R$, such that

$$r = G(u) . \quad (3.1)$$

Here, G is a forward map, X is the space of inputs and R is the space of results:

$$G : X \rightarrow R . \quad (3.2)$$

In a practical case, we have a finite set of data points $y \in \mathbb{R}^{N_{data}}$ sampled from r , with added experimental noise η :

$$y = \mathcal{G}(u) + \eta \quad (3.3)$$

We assume η is Gaussian-distributed noise that causes the measured data points y to fluctuate around the true (but unknown) values $\mathcal{G}(u)$. Here, $\mathcal{G}(u)$ differs from

¹In literature one can find the "model" defined as the parametrization we use to describe the non-perturbative TMDs. In this work we follow Ref. [5], where the parametrization is defined as the "input".

$G(u) = r \in R$, mapping the continuous model input u into a finite set of experimental values $y_0 = \mathcal{G}(u) \in \mathbb{R}^{N_{data}}$. To find the underlying input u , we construct a parametrization that can capture the functional form of u using a finite set of parameters, composed by a number N_{model} of parameters. In this context, $u \in \mathbb{R}^{N_{model}}$.

To summarize, in a practical scenario, we have three key components:

- A dataset $\mathcal{D} \in \mathbb{R}^{N_{data}}$, which can be represented as a point in a real space of dimension N_{data} . The dataset is measured and thus subject to experimental uncertainty.
- An input u , which is parametrized using a set of N_{model} parameters. Once parametrized, each possible input can be identified with a point in the space $\mathbb{R}^{N_{model}}$.
- A known model $\mathcal{G} : \mathbb{R}^{N_{model}} \rightarrow \mathbb{R}^{N_{data}}$ that maps the input u to the dataset $\mathcal{D}$.

The inverse problem can be solved finding the best possible input, given the model, the dataset, and the parametrization. The best input, denoted as u_* , is the one that maximizes the likelihood of obtaining the dataset $\mathcal{D} = \{y_i\}$, given the model $\mathcal{G}(u)$. Since the maximum likelihood corresponds to the minimum χ^2 between the experimental data $\mathcal{D}$ and the predictions $\mathcal{G}(u)$, we use an optimization algorithm to find the minimum χ^2 in the parameter space. The algorithm searches for the point in $\mathbb{R}^{N_{model}}$ that minimizes χ^2 , identifying the set of parameters for the input u , that define the best model u_* :

$$u_* = \arg \min_{u \in X} \chi^2[\mathcal{G}(u), \mathcal{D}] . \quad (3.4)$$

We will now explore the three components (data, model, and input) in the two fitting frameworks under consideration, MAPTMD22 and PV19. The main differences between these frameworks lie in the datasets and in the parametrization of the non-perturbative part of the TMDs.

3.2 Experimental Datasets

The datasets included in the two frameworks for determining the TMDs differ in the number of data points (with MAPTMD22 having a significantly larger dataset) and the processes analyzed. MAPTMD22 uses both Drell-Yan and SIDIS experiments, with a total of 2032 data points (485 for Drell-Yan and 1547 for SIDIS); PV19, on the other hand, includes only Drell-Yan measurements, with a total of 353 data points. Tables 3.1, 3.2, and 3.3 show the experimental measurements included in PV19 and MAPTMD22. The dataset used in PV19 is a subset of the MAPTMD22 dataset. Specifically, the MAPTMD22 dataset includes the entire PV19 dataset and adds Drell-Yan data from the experiments E772, PHENIX, and ATLAS 13 TeV, along with SIDIS data from the HERMES and COMPASS experiments.

The kinematic region covered by the MAPTMD22 dataset in the (x, Q^2) space is shown in figure 3.1.

CHAPTER 3. TMD EXTRACTION FROM EXPERIMENTAL DATA

Experiment	N_{dat}	Observable	$\sqrt{s}$ [GeV]	Q [GeV]	y or x_F	Lepton cuts	Ref.
E605	50	$Ed^3\sigma/d^3q$	38.8	7 - 18	$x_F = 0.1$	-	[23]
E288 200 GeV	30	$Ed^3\sigma/d^3q$	19.4	4 - 9	$y = 0.40$	-	[24]
E288 300 GeV	39	$Ed^3\sigma/d^3q$	23.8	4 - 12	$y = 0.21$	-	[24]
E288 400 GeV	61	$Ed^3\sigma/d^3q$	27.4	5 - 14	$y = 0.03$	-	[24]
STAR 510	7	$d\sigma/dq_T$	510	73 - 114	$ y < 1$	$p_{T\ell} > 25 \text{ GeV}$ $ \eta_\ell < 1$	-
CDF Run I	25	$d\sigma/dq_T$	1800	66 - 116	Inclusive	-	[25]
CDF Run II	26	$d\sigma/dq_T$	1960	66 - 116	Inclusive	-	[26]
D0 Run I	12	$d\sigma/dq_T$	1800	75 - 105	Inclusive	-	[27]
D0 Run II	5	$(1/\sigma)d\sigma/dq_T$	1960	70 - 110	Inclusive	-	[28]
D0 Run II (μ)	3	$(1/\sigma)d\sigma/dq_T$	1960	65 - 115	$ y < 1.7$	$p_{T\ell} > 15 \text{ GeV}$ $ \eta_\ell < 1.7$	[29]
LHCb 7 TeV	7	$d\sigma/dq_T$	7000	60 - 120	$2 < y < 4.5$	$p_{T\ell} > 20 \text{ GeV}$ $2 < \eta_\ell < 4.5$	[30]
LHCb 8 TeV	7	$d\sigma/dq_T$	8000	60 - 120	$2 < y < 4.5$	$p_{T\ell} > 20 \text{ GeV}$ $2 < \eta_\ell < 4.5$	[31]
LHCb 13 TeV	7	$d\sigma/dq_T$	13000	60 - 120	$2 < y < 4.5$	$p_{T\ell} > 20 \text{ GeV}$ $2 < \eta_\ell < 4.5$	[32]
CMS 7 TeV	4	$(1/\sigma)d\sigma/dq_T$	7000	60 - 120	$ y < 2.1$	$p_{T\ell} > 20 \text{ GeV}$ $ \eta_\ell < 2.1$	[33]
CMS 8 TeV	4	$(1/\sigma)d\sigma/dq_T$	8000	60 - 120	$ y < 2.1$	$p_{T\ell} > 15 \text{ GeV}$ $ \eta_\ell < 2.1$	[34]
ATLAS 7 TeV	6 6 6	$(1/\sigma)d\sigma/dq_T$	7000	66 - 116	$ y < 1$ $1 < y < 2$ $2 < y < 2.4$	$p_{T\ell} > 20 \text{ GeV}$ $ \eta_\ell < 2.4$	[35]
ATLAS 8 TeV on-peak	6 6 6 6 6 6	$(1/\sigma)d\sigma/dq_T$	8000	66 - 116	$ y < 0.4$ $0.4 < y < 0.8$ $0.8 < y < 1.2$ $1.2 < y < 1.6$ $1.6 < y < 2$ $2 < y < 2.4$	$p_{T\ell} > 20 \text{ GeV}$ $ \eta_\ell < 2.4$	[36]
ATLAS 8 TeV off-peak	4 8	$(1/\sigma)d\sigma/dq_T$	8000	46 - 66 116 - 150	$ y < 2.4$	$p_{T\ell} > 20 \text{ GeV}$ $ \eta_\ell < 2.4$	[36]
Total	353	-	-	-	-	-	-

Table 3.1: Drell-Yan measurements included in PV19, table from [3]

CHAPTER 3. TMD EXTRACTION FROM EXPERIMENTAL DATA

Experiment	N_{dat}	Observable	$\sqrt{s}$ [GeV]	Q [GeV]	y or x_F	Lepton cuts	Ref.
E605	50	$Ed^3\sigma/d^3\mathbf{q}$	38.8	7 - 18	$x_F = 0.1$	-	[23]
E772	53	$Ed^3\sigma/d^3\mathbf{q}$	38.8	5 - 15	$0.1 < x_F < 0.3$	-	[37]
E288 200 GeV	30	$Ed^3\sigma/d^3\mathbf{q}$	19.4	4 - 9	$y = 0.40$	-	[24]
E288 300 GeV	39	$Ed^3\sigma/d^3\mathbf{q}$	23.8	4 - 12	$y = 0.21$	-	[24]
E288 400 GeV	61	$Ed^3\sigma/d^3\mathbf{q}$	27.4	5 - 14	$y = 0.03$	-	[24]
STAR 510	7	$d\sigma/d q_T $	510	73 - 114	$ y < 1$	$p_{T\ell} > 25 \text{ GeV}$ $ \eta_\ell < 1$	-
PHENIX200	2	$d\sigma/d q_T $	200	4.8 - 8.2	$1.2 < y < 2.2$	-	[38]
CDF Run I	25	$d\sigma/d q_T $	1800	66 - 116	Inclusive	-	[25]
CDF Run II	26	$d\sigma/d q_T $	1960	66 - 116	Inclusive	-	[26]
D0 Run I	12	$d\sigma/d q_T $	1800	75 - 105	Inclusive	-	[27]
D0 Run II	5	$(1/\sigma)d\sigma/d q_T $	1960	70 - 110	Inclusive	-	[28]
D0 Run II (μ)	3	$(1/\sigma)d\sigma/d q_T $	1960	65 - 115	$ y < 1.7$	$p_{T\ell} > 15 \text{ GeV}$ $ \eta_\ell < 1.7$	[29]
LHCb 7 TeV	7	$d\sigma/d q_T $	7000	60 - 120	$2 < y < 4.5$	$p_{T\ell} > 20 \text{ GeV}$ $2 < \eta_\ell < 4.5$	[30]
LHCb 8 TeV	7	$d\sigma/d q_T $	8000	60 - 120	$2 < y < 4.5$	$p_{T\ell} > 20 \text{ GeV}$ $2 < \eta_\ell < 4.5$	[31]
LHCb 13 TeV	7	$d\sigma/d q_T $	13000	60 - 120	$2 < y < 4.5$	$p_{T\ell} > 20 \text{ GeV}$ $2 < \eta_\ell < 4.5$	[32]
CMS 7 TeV	4	$(1/\sigma)d\sigma/d q_T $	7000	60 - 120	$ y < 2.1$	$p_{T\ell} > 20 \text{ GeV}$ $ \eta_\ell < 2.1$	[33]
CMS 8 TeV	4	$(1/\sigma)d\sigma/d q_T $	8000	60 - 120	$ y < 2.1$	$p_{T\ell} > 15 \text{ GeV}$ $ \eta_\ell < 2.1$	[34]
CMS 13 TeV	70	$d\sigma/d q_T $	13000	76 - 106	$ y < 0.4$ $0.4 < y < 0.8$ $0.8 < y < 1.2$ $1.2 < y < 1.6$ $1.6 < y < 2.4$	$p_{T\ell} > 25 \text{ GeV}$ $ \eta_\ell < 2.4$	[39]
ATLAS 7 TeV	6 6 6	$(1/\sigma)d\sigma/d q_T $	7000	66 - 116	$ y < 1$ $1 < y < 2$ $2 < y < 2.4$	$p_{T\ell} > 20 \text{ GeV}$ $ \eta_\ell < 2.4$	[35]
ATLAS 8 TeV on-peak	6 6 6 6 6	$(1/\sigma)d\sigma/d q_T $	8000	66 - 116	$ y < 0.4$ $0.4 < y < 0.8$ $0.8 < y < 1.2$ $1.2 < y < 1.6$ $1.6 < y < 2$ $2 < y < 2.4$	$p_{T\ell} > 20 \text{ GeV}$ $ \eta_\ell < 2.4$	[36]
ATLAS 8 TeV off-peak	4 8	$(1/\sigma)d\sigma/d q_T $	8000	46 - 66 116 - 150	$ y < 2.4$	$p_{T\ell} > 20 \text{ GeV}$ $ \eta_\ell < 2.4$	[36]
ATLAS 13 TeV	6	$(1/\sigma)d\sigma/d q_T $	13000	66 - 116	$ y < 2.5$	$p_{T\ell} > 27 \text{ GeV}$ $ \eta_\ell < 2.5$	[40]
Total	484						

Table 3.2: Drell-Yan measurements included in MAPTMD22, table from [4] 30

CHAPTER 3. TMD EXTRACTION FROM EXPERIMENTAL DATA

Experiment	N_{dat}	Observable	Channels	Q [GeV]	x	z	Phase space cuts	Ref.
HERMES	344	$M(x, z, P_{hT} , Q)$	$p \rightarrow \pi^+$	$1 - \sqrt{15}$	$0.023 < x < 0.6$ (6 bins)	$0.1 < z < 1.1$ (8 bins)	$W^2 > 10 \text{ GeV}^2$ $0.1 < y < 0.85$	[41]
			$p \rightarrow \pi^-$					
			$p \rightarrow K^+$					
			$p \rightarrow K^-$					
			$d \rightarrow \pi^+$					
			$d \rightarrow \pi^-$					
			$d \rightarrow K^+$					
COMPASS	1203	$M(x, z, P_{hT}^2, Q)$	$d \rightarrow h^+$	1 - 9 (5 bins)	$0.003 < x < 0.4$ (8 bins)	$0.2 < z < 0.8$ (4 bins)	$W^2 > 25 \text{ GeV}^2$ $0.1 < y < 0.9$	[42]
			$d \rightarrow h^-$					
Total	1547							

Table 3.3: SIDIS measurements included in MAPTMD22, table from [4]

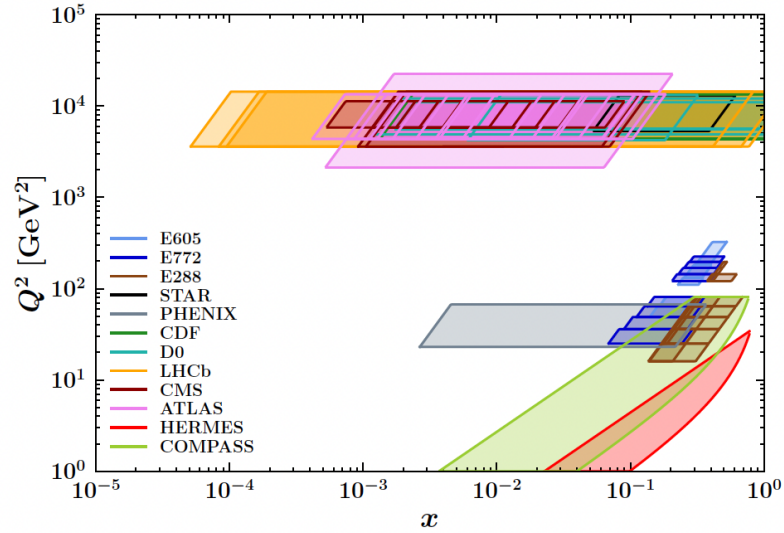


Figure 3.1: Kinematic coverage in (x, Q^2) space of the dataset used in MAPTMD22, image from [4]

When examining tables 3.1, 3.2, and 3.3, it is noticeable that the physical observables measured are not identical for each experiment, nor do they directly correspond to the cross sections for Drell-Yan and SIDIS as defined in equations (2.31) and (2.39). The data preprocessing required to effectively utilize these datasets is discussed in Refs. [3, 4]. Kinematic cuts have been applied to the Drell-Yan datasets to ensure the condition $q_T \ll Q$ is met: in both frameworks, only data points satisfying the condition $q_T < 0.2Q$ are considered. For the SIDIS dataset included in MAPTMD22, a cut on the virtuality Q is applied, imposing $Q > 1.4 \text{ GeV}$ to satisfy the condition $Q \gg \Lambda_{QCD}$.

Additionally, a restriction of $0.2 < z < 0.7$ is imposed to restrict the measurements on SIDIS in the current fragmentation region, along with a cut on $|P_{hT}|$, which is dependent on the values of Q and z , and comes from the implementation of the $q_T \ll Q$ condition. The choice made in Ref. [4] was:

$$|P_{hT}| < \min[\min[0.2 Q, 0.5 z Q] + 0.3 \text{ GeV}, z Q] .$$

The number of data points reported in the tables above accounts only for those points that survived the kinematic cuts. According to the definition of the inverse problem, we thus have a dataset for PV19, $\mathcal{D}_{PV19} \in \mathbb{R}^{353}$, and a dataset for MAPTMD22, $\mathcal{D}_{MAP22} \in \mathbb{R}^{2032}$.

3.3 PV19 and MAPTMD22 Models

In both frameworks, the model $\mathcal{G}$ is represented by the factorized cross sections introduced in equations (2.31) and (2.39), using the structure functions in b_T space described by equations (2.35) and (2.43). Since the experiments cover different values of Q (as shown in figure 3.1), the model must include evolution equations for the TMDs: we use the standard evolution equations (see section 2.5).

Here, we briefly discuss the models used in PV19, $\mathcal{G}_{PV19}$, and MAPTMD22, $\mathcal{G}_{MAP22}$. Notably, the definition of the Drell-Yan cross section is identical in both frameworks, meaning that the MAPTMD22 model, $\mathcal{G}_{MAP22}$, contains the PV19 model (the part of the model used for Drell-Yan data is the same) and adds new information for SIDIS data. We can conveniently divide the MAPTMD22 model into two sub-models:

$$\mathcal{G}_{MAP22} = \begin{cases} \mathcal{G}_{MAP22, \text{SIDIS}} & \text{for SIDIS data} \\ \mathcal{G}_{MAP22, \text{DY}} & \text{for Drell-Yan data,} \end{cases} \quad (3.5)$$

and identify

$$\mathcal{G}_{MAP22, \text{DY}} = \mathcal{G}_{PV19}. \quad (3.6)$$

For the practical solution of the inverse problem, we need to explicitly compute each perturbative contribution to the factorized cross sections up to a specific order. Despite having the same definition for the Drell-Yan cross section, the models $\mathcal{G}_{MAP22, \text{DY}}$ and $\mathcal{G}_{PV19}$ could differ if different orders were used in the perturbative approximation of the Drell-Yan cross section in the two frameworks. However, this is not the case here, as both frameworks use calculations of the same accuracy for each ingredient in the Drell-Yan cross section. Table 3.4 shows the truncation orders in the perturbative expansion of the ingredients of the factorized SIDIS and Drell-Yan cross sections. The accuracy of the model depends on the truncation order of each ingredient. In the PV19 framework, the model $\mathcal{G}_{PV19}$ reaches full N³LL (*next-to-next-to-next-to-leading*) accuracy. For the MAPTMD22 framework, the same accuracy is reached for the Drell-Yan cross section, but the full N³LL accuracy for SIDIS was not achieved due to the absence of extractions of the collinear FF up to NNLO.

The accuracy of both models, $\mathcal{G}_{MAP22}$ and $\mathcal{G}_{PV19}$, is shown in the fourth row of table 3.4. Since MAPTMD22 includes the PV19 dataset and model, its accuracy level

$\mathcal{G}_{\text{MAP22, DY}} (= \mathcal{G}_{\text{PV19}})$	$\mathcal{G}_{\text{MAP22, SIDIS}}$	H and C	K and γ_F	γ_K	PDF and α_s evolution	FF evolution
NLL	NLL	0	1	2	LO	LO
N ² LL	N ² LL	1	2	3	NLO	NLO
N ³ LL	N ³ LL ⁻	2	3	4	NNLO	NLO
N ³ LL	N ³ LL	2	3	4	NNLO	NNLO

Table 3.4: Levels of accuracy in the models, depending on the truncation orders in the perturbative content of the factorized cross sections. The accuracy reached by $\mathcal{G}_{\text{MAP22, DY}}$ and $\mathcal{G}_{\text{PV19}}$ is N³LL; the accuracy of $\mathcal{G}_{\text{MAP22, SIDIS}}$ is N³LL⁻

is referred to as N³LL⁻, indicating that it is fully N³LL for the Drell-Yan part but N³LL⁻ for the SIDIS part.

Up to this point, we have introduced the datasets and the models used in the two frameworks. We mentioned that the model connecting the Drell-Yan cross section with experimental data is the same in both frameworks. What differs between the frameworks is the input parametrization, which we will now discuss.

3.4 Input Parametrization

The input u to the model is the part we aim to infer from experimental data by solving the inverse problem. In our frameworks, the input is represented by the non-perturbative TMDs entering in the factorized cross sections. However, we cannot directly identify the input with the non-perturbative TMD PDF ($f_{1NP}(x, b_T^2; \zeta_f, Q_0)$) and FF ($D_{1NP}^{a \rightarrow h}(z, b_T^2; \zeta_f, Q_0)$), because the factorized cross sections do not depend straightforwardly on these non-perturbative TMDs. Specifically, the Drell-Yan cross section depends on two TMD PDFs (equation (2.39)), and the SIDIS cross section depends on a TMD PDF and a TMD FF (equation (2.31)).

The procedure to define the input to the models $\mathcal{G}_{\text{MAP22}}$ and $\mathcal{G}_{\text{PV19}}$ involves two steps:

- Construct a parametrization that captures the functional form of the non-perturbative TMDs, $f_{1NP}(x, b_T^2; \zeta_f, Q_0)$ and $D_{1NP}^{a \rightarrow h}(z, b_T^2; \zeta_f, Q_0)$.
- Identify the inputs to the models with the set of parameters used to parametrize *all* the non-perturbative TMDs in the models. In PV19, the parametrization involves 9 free parameters, while MAPTMD22 involves a total of 21 free parameters.

We can then define two inputs, one for each framework:

$$u_{\text{MAP22}} \in \mathbb{R}^{21} \quad (3.7)$$

$$u_{\text{PV19}} \in \mathbb{R}^9. \quad (3.8)$$

Before discussing the explicit MAPTMD22 and PV19 parameterizations, let's clarify two aspects of the non-perturbative TMDs. First, the separation between the perturbative and non-perturbative parts of the TMDs is arbitrary and depends on the choice of

b_* ; only the combination in equation (2.57) is meaningful. That is, the separation of the *input* from the *model* depends on arbitrary choices, while the combination of the two, $\mathcal{G}(u)$, has physical meaning. Secondly, since we cannot predict the non-perturbative components using perturbation theory, the parametrization of the non-perturbative TMDs is arbitrary, and different choices for the parametrization could lead to equivalent results in determining the functional form of the TMDs. The physical constraint used to characterize TMDs is that the non-perturbative contributions must approach one as b_T approaches zero, that means

$$f_{1NP}, D_{1NP} \rightarrow 1 + \mathcal{O}\left(\frac{1}{Q^p}\right) \text{ for } b_T \rightarrow 0. \quad (3.9)$$

Table 3.5 summarizes the parameters defining the two inputs, u_{MAP22} and u_{PV19} , with the functional forms of the inputs described in the following sections.

MAPTMD22 Input Parametrization

In the MAPTMD22 framework, both unpolarized TMD PDFs and FFs are parametrized, using a total of 21 free parameters. One parameter is used for the non-perturbative part of the Collins-Soper kernel,

$$g_K(b_T^2) = -g_2^2 \frac{b_T^2}{2}, \quad (3.10)$$

this defines a common factor, S_{NP} , which enters into the parametrization of both non-perturbative TMD PDFs and FFs:

$$S_{\text{NP}}(\zeta, b_T) = \exp\left[-g_2^2 \frac{b_T^2}{4} \log\left(\frac{\zeta}{Q_0^2}\right)\right] = \left[\frac{\zeta}{Q_0^2}\right]^{g_K(b_T^2)/2}. \quad (3.11)$$

Eleven parameters are associated with the non-perturbative TMD PDF (assumed to be flavor-independent):

$$f_{\text{NP}}(x, \zeta, b_T) = S_{\text{NP}}(\zeta, b_T) \frac{g_1(x)e^{-g_1(x)\frac{b_T^2}{4}} + \lambda^2 g_{1B}^2(x)(1 - g_{1B}(x)\frac{b_T^2}{4})e^{-g_{1B}(x)\frac{b_T^2}{4}} + \lambda_2^2 g_{1C}(x)e^{-g_{1C}(x)\frac{b_T^2}{4}}}{g_1(x) + \lambda^2 g_{1B}^2(x) + \lambda_2^2 g_{1C}(x)} \quad (3.12)$$

The functions $g_1(x)$, $g_{1B}(x)$, and $g_{1C}(x)$ describe the dependence of the widths of the distribution on x :

$$g_{1,1B,1C}(x) = N_{1,1B,1C} \frac{x^{\sigma_{1,2,3}}(1-x)^{\alpha_{1,2,3}}}{\hat{x}^{\sigma_{1,2,3}}(1-\hat{x})^{\alpha_{1,2,3}}}. \quad (3.13)$$

The functional form of the non-perturbative TMD FF (also flavor-independent) is described by nine free parameters:

$$D_{\text{NP}}(z, \zeta, b_T) = S_{\text{NP}}(\zeta, b_T) \frac{g_3(z)e^{-g_3(z)\frac{b_T^2}{4z^2}} + \frac{\lambda_F}{z^2} g_{3B}^2(z)(1 - g_{3B}(z)\frac{b_T^2}{4z^2})e^{-g_{3B}(z)\frac{b_T^2}{4z^2}}}{g_3(z) + \frac{\lambda_F}{z^2} g_{3B}^2(z)} \quad (3.14)$$

Framework	Input Name	Parametrized Functions	Parameters	Total Parameters
MAPTMD22	u_{MAP22}	S_{NP}	g_2	21
		f_{NP}	$\alpha_1, \alpha_2, \alpha_3$ $\sigma_1, \sigma_2, \sigma_3, \lambda_2$ $N_1, N_{1B}, N_{1C}, \lambda$	
		D_{NP}	N_3, N_{3B}, β_1 $\beta_2, \delta_1, \delta_2$ $\gamma_1, \gamma_2, \lambda_F$	
PV19	u_{PV19}	f_{NP}	g_2, N_1, α σ, λ, N_{1B} $\alpha_B, \sigma_B, g_{2B}$	9

Table 3.5: Inputs parametrization of the MAPTMD22 and PV19 frameworks.

where $g_3(z)$ and $g_{3B}(z)$ describe the dependence of the width of the FF distribution on z :

$$g_{3,3B}(z) = N_{3,3B} \frac{(z^{\beta_{1,2}} + \delta_{1,2}^2)(1-z)^{\gamma_{1,2}^2}}{(\hat{z}^{\beta_{1,2}} + \delta_{1,2}^2)(1-\hat{z})^{\gamma_{1,2}^2}}. \quad (3.15)$$

In this parametrization, $Q_0^2 = 1; \text{GeV}^2$, $\hat{x} = 0.1$, and $\hat{z} = 0.5$ are fixed.

PV19 Input Parametrization

In PV19, only the non-perturbative TMD PDF is parametrized. In this case, f_{1NP} is also assumed to be flavor-independent and is parametrized using nine free parameters:

$$f_{NP}(x, \zeta, b_T) = \left(\frac{1 - \lambda}{1 + g_1(x)b_T^2/4} + \lambda e^{-g_{1B}(x)b_T^2/4} \right) \exp \left[-g_2 \log \left(\frac{\zeta}{Q_0^2} \right) b_T^2/4 - g_{2B} \log \left(\frac{\zeta}{Q_0^2} \right) b_T^4/4 \right], \quad (3.16)$$

with $g_1(x)$ and $g_{1B}(x)$ functions defined as:

$$g_1(x) = \frac{N_1}{x\sigma} \exp \left[-\frac{\ln^2 \left(\frac{x}{\alpha} \right)}{2\sigma^2} \right] \quad (3.17)$$

$$g_{1B}(x) = \frac{N_{1B}}{x\sigma_B} \exp \left[-\frac{\ln^2 \left(\frac{x}{\alpha_B} \right)}{2\sigma_B^2} \right]. \quad (3.18)$$

Here, the initial energy scale is set to be $Q_0^2 = 1 \text{ GeV}^2$. The second line in equation (3.16), corresponds to the S_{NP} contribution defined for equations (3.12) and (3.14), we do not separated the two terms in the non-perturbative TMD PDF defined in the PV19 framework, following the definition in Ref. [3].

3.5 The Fitting Methodology

We have now defined all the inverse problem ingredients (dataset, model, and input) for the two frameworks. The goal of the analyzed procedures is to find the set of parameters (input) u_* that best fits the experimental data $\mathcal{D}$, given the model $\mathcal{G}$. This task can be done using an optimization algorithm, such as the gradient descent, the Levenberg-Marquardt, and the MIGRAD algorithms [43, 44]. The utilized optimization algorithm in the framework depend on the adopted minimizer, in the MAPTMD22 and PV19 frameworks the choices are Minuit [45] and Ceres [44].

Optimization Process

To explain the workflow of the optimization algorithm, consider the PV19 fitting framework (the optimization process in MAPTMD22 will be the same). The algorithm starts by considering a random input $u_{PV19}^{(0)} \in \mathbb{R}^9$ (i.e., it starts from a random point in the parameter space). Using this random set of parameters, together with the model $\mathcal{G}_{PV19}$, represented by the Drell-Yan factorized cross section computed at N³LL accuracy, the predictions $\mathcal{G}(u^{(0)})$ are computed (from now on, we will drop the label $PV19$ from $u_{PV19}^{(0)}$ for conciseness). This produces a predicted dataset $\mathcal{D}_{pred}^{(0)} = \mathcal{G}(u^{(0)}) \in \mathbb{R}^{N_{data}}$, with $N_{data} = 353$ being the number of data points considered in the PV19 analysis. The predicted dataset $\mathcal{D}_{pred}^{(0)}$ is then compared to the real dataset $\mathcal{D}$ using the $\chi^2[\mathcal{D}_{pred}^{(0)}, \mathcal{D}]$ as a figure of merit. Equation (3.19) represents the χ^2 between a predicted dataset $\mathcal{D}_{pred}^{(n)} = \mathcal{G}(u^{(n)})$ and the real dataset $\mathcal{D}$,

$$\chi^2[\mathcal{D}_{pred}^{(n)}, \mathcal{D}] = \sum_{i,j}^{N_{data}} (\mathcal{D}_i^{(n)} - \mathcal{D}_i) C_{ij}^{-1} (\mathcal{D}_j^{(n)} - \mathcal{D}_j) , \quad (3.19)$$

where $\mathcal{D}_{pred}^{(n)}$ is the dataset predicted in the $(n+1)$ -th step of the optimization algorithm; $\mathcal{D}_i^{(n)}$ and $\mathcal{D}_i$ are the i -th points in the predicted and real datasets, respectively, and C_{ij} is the $N_{data} \times N_{data}$ covariance matrix of the real dataset.

The structure of the covariance matrix and its relationship with the experimental errors of the dataset is discussed in section 4.7.

Once the χ^2 for the first set of parameters, $u^{(0)}$, is computed, the optimization algorithm explores the χ^2 around the point $u^{(0)}$ and takes a step in the parameter space in the direction in which the χ^2 function decreases most rapidly [43], arriving at a new point in the parameter space, $u^{(1)}$. The same process is then repeated, and a new point $u^{(2)}$ with a smaller χ^2 is found, and so on. If the optimization algorithm works properly, at each step the χ^2 decreases²,

$$\chi^2[\mathcal{D}_{pred}^{(n+1)}, \mathcal{D}] < \chi^2[\mathcal{D}_{pred}^{(n)}, \mathcal{D}] , \quad (3.20)$$

meaning the algorithm finds a better set of parameters at each step. This process continues until a stopping criterion is reached. The stopping criterion used in the

²This is not necessarily true for each step, but if we take a large number of steps, we can affirm that the mean behavior is a decreasing χ^2 in the parameter space.

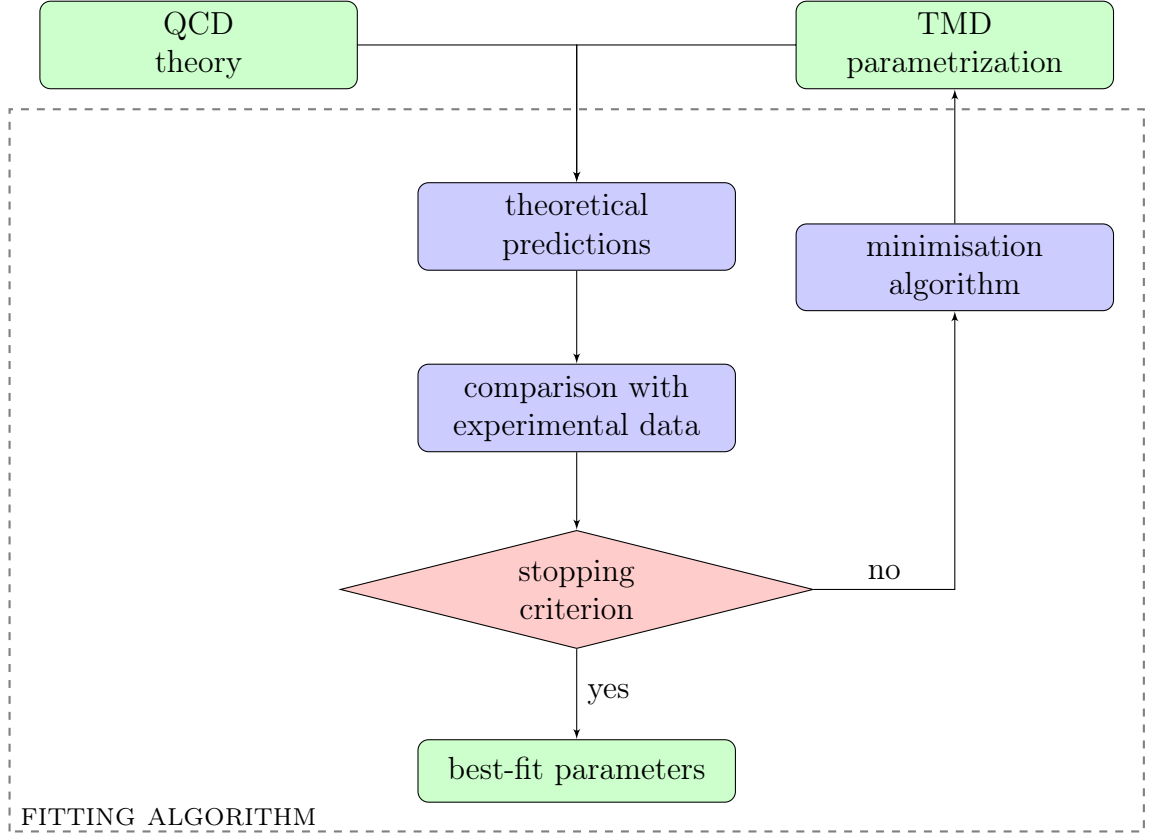


Figure 3.2: Flowchart of the optimization process in the analyzed fitting frameworks. In this chart, the model $\mathcal{G}$ is represented by the label "QCD theory", the input $u^{(n)}$ is represented by the "TMD parametrization", and the predicted dataset $\mathcal{G}(u^{(n)})$ is represented by "theoretical predictions".

MAP frameworks is

$$\frac{\chi^2[\mathcal{D}_{pred}^{(n+1)}, \mathcal{D}] - \chi^2[\mathcal{D}_{pred}^{(n)}, \mathcal{D}]}{\chi^2[\mathcal{D}_{pred}^{(n+1)}, \mathcal{D}]} < 10^{-5} . \quad (3.21)$$

When the criterion in equation (3.21) is reached, the algorithm stops, and we identify the set of parameters $u^{(n+1)}$ as the solution to the inverse problem u_* ,

$$u^{(n+1)} = u_* = \arg \min_{u \in \mathbb{R}^{N_{data}}} \chi^2[\mathcal{G}(u), \mathcal{D}] . \quad (3.22)$$

The flowchart of the optimization process is shown in figure 3.2.

When a solution u_* to the inverse problem is found, we cannot be sure that this is the *correct* solution. Two problems could arise:

- The algorithm stops at a local minimum of the χ^2 , missing the global minimum.
- The algorithm overfits the data, meaning it fits the statistical noise η of the data.

Due to these problems, a small variation in the dataset or in the starting point $u^{(0)}$ can lead to a large variation in the solution. To overcome this issue and quantify the uncertainty in the solution to the inverse problem, the fitting frameworks use the *bootstrap* method.

Bootstrap Approach to Inverse Problem

The bootstrap approach involves resampling the dataset $\mathcal{D}$ multiple times, obtaining a set of Monte Carlo replicas of the dataset, and repeating the workflow described in figure 3.2 for each replica $\mathcal{D}^{(k)}$ (equation (3.25)), starting from a different point in the parameter space each time. The original experimental dataset is composed of a set of N_{data} points. Each point y_i is shifted from the point in the real distribution, $y_i^0 = \mathcal{G}(u_{real})$, by an experimental noise η_i . Here, the notation $y_i = \mathcal{D}_i$ refers to a single data point, and u_{real} is the real, unknown input. The single point reads

$$y_i = y_i^0 + \eta_i . \quad (3.23)$$

It is important to note that we do not know the point y_i^0 and the fluctuation η_i ; we only know their combination y_i . The full experimental dataset is

$$\mathcal{D} = y^0 + \eta , \quad (3.24)$$

where $y^0 \in \mathbb{R}^{N_{data}}$ is the vector containing all the non-fluctuated points, and $\eta \in \mathbb{R}^{N_{data}}$ is the vector containing all the unknown fluctuations.

Starting from the experimental dataset $\mathcal{D}$, we construct a number N_{rep} of Monte Carlo replicas, adding another layer of fluctuation, ϵ . The k -th replica will have the form:

$$\mathcal{D}^{(k)} = y^0 + \eta + \epsilon^{(k)} . \quad (3.25)$$

The replicas are produced using the Cholesky decomposition of the covariance matrix. This method is discussed in detail in section 4.8.

Having a set of N_{rep} replicas, we can solve the inverse problem for each replica. The solution to the k -th replica will be:

$$u_*^{(k)} = \arg \min_{u^{(k)} \in \mathbb{R}^{N_{model}}} \chi^2[\mathcal{G}(u^{(k)}), \mathcal{D}^{(k)}] . \quad (3.26)$$

We thus obtain a set of N_{rep} solutions $u_*^{(k)}$. To determine the value of the non-perturbative TMDs in a kinematic point (x, b_T, ζ) , we average over all the TMDs obtained by inserting the sets of parameters $u_*^{(k)}$ in the parametrization used in the framework (referred to as $f_{NP}(u_*^{(k)})$ and $D_{NP}(u_*^{(k)})$) in that kinematic point:

$$f_{NP,*}(x, b_T, \zeta) = \frac{1}{N_{rep} + 1} \sum_{k=0}^{N_{rep}} f_{NP}(u_*^{(k)}) \quad (3.27)$$

$$D_{NP,*}(x, b_T, \zeta) = \frac{1}{N_{rep} + 1} \sum_{k=0}^{N_{rep}} D_{NP}(u_*^{(k)}) , \quad (3.28)$$

where we consider $u_*^{(0)}$ the solution to the real (non-fluctuated) dataset. We omitted the dependence of the fitted TMDs on x , b_T , and ζ for conciseness. This method removes the dependence of the final solutions $f_{NP,*}$ and $D_{NP,*}$ from the data fluctuation and

from the starting point in the parameter space, providing an uncertainty quantification for the final solution. In particular, to quantify the TMDs total uncertainty, we use each predicted set of parameters $u_*^{(k)}$ to reconstruct a number N_{rep} of non-perturbative TMDs ($f_{NP}(u_*^{(k)})$), with $k = \{1 \dots N_{\text{rep}}\}$, we then compute the standard deviation of the set $\{f_{NP}(u_*^{(k)})\}$,

$$\sigma(x, b_T, \zeta) = \sqrt{\langle (f_{NP,*}^{(k)} - \langle f_{NP,*} \rangle)^2 \rangle} = \sqrt{\frac{1}{N_{\text{rep}}} \sum_{k=0}^{N_{\text{rep}}} (f_{NP,*}^{(k)} - \langle f_{NP,*} \rangle)^2} . \quad (3.29)$$

We assume the TMDs to be Gaussian distributed around the central value $f_{NP,*}(x, b_T, \zeta)$; the same reasoning applies to the uncertainty determination of the TMD FF, $D_{NP,*}(x, b_T, \zeta)$. In chapter 5, we will refer to $f_{NP,*}$ and $D_{NP,*}$ as the predicted TMDs, and the labels will be changed to $f_{NP,\text{pred}}$ and $D_{NP,\text{pred}}$ for consistency with the plots labels.

Weaknesses of the Fitting Frameworks

Despite providing a solution to the inverse problem (equations (3.27), (3.28)) and a quantitative estimation of the uncertainty of this solution (equation (3.29)), there remain some unanswered questions regarding the frameworks we are considering. In particular, there are three questions that the current methodologies are not able to address:

1. We cannot be sure that the methodologies can reproduce the real input (u_{real}) that we are searching for, nor can we quantify how far the predicted TMD $f_{NP,*}$ is from the real solution $f_{NP}(u_{\text{real}})$ (in our context, the real solution is represented by the actual functional form of the TMDs constructed with the real input u_{real}).
2. We cannot estimate the contribution of interpolation/extrapolation, functional, and experimental uncertainties³ to the total uncertainty.
3. We don't know if the uncertainties declared by the methodologies is faithful to the real uncertainties of the TMDs.

The main goal of this thesis is to answer these questions for the MAPTMD22 and PV19 methodologies by implementing the procedure of closure testing in the frameworks, which we will discuss in the next chapter.

³These three types of uncertainty are discussed in the next chapter.

Chapter 4

Closure Testing

In this chapter, we discuss the theory of closure testing and the implementation of this feature within the MAP fitting frameworks. The results obtained for the two analyzed frameworks, MAPTMD22 and PV19, will be presented in the next chapter.

4.1 Closure Tests

As discussed in section 3.5, when the optimization algorithm searches for the best input parameters, it compares the real dataset $\mathcal{D}$ with the predicted dataset $\mathcal{D}_{pred}$ generated using a set of parameters u and a theoretical model $\mathcal{G}$: $\mathcal{D}_{pred} = \mathcal{G}(u)$. Unfortunately, since $\mathcal{D}_{pred}$ does not trivially depend on the optimal input u_* (*i.e.*, the model $\mathcal{G}$ is complex), knowing how close the predictions $\mathcal{D}_{pred}$ are to the real values $\mathcal{D}$ is not enough to determine how close the predicted input u_* is to the true input u_{real} . In other words, comparing the real dataset with the predicted one does not tell us how close the predicted functional form of the TMDs are to the real TMDs.

Closure testing is a statistical method that allows us to directly compare u_* with u_{real} , providing a measure of how well the methodology reproduces the true underlying law that we are trying to uncover [5, 6].

To perform closure testing, we choose an input to act as the *real* input to the model¹. For instance, we can set the predicted parameters from MAPTMD22 as the real input, $u_{*,MAP22} = u_{real}$, but the choice of u_{real} is arbitrary and in principle does not need to be similar to a functional form inferred from experimental data. Using this chosen input u_{real} , we generate a set of pseudo-data,

$$\mathcal{D}_{L0} = \mathcal{G}(u_{real}) , \quad (4.1)$$

where the notation L0 indicates "Level 0", meaning that the dataset $\mathcal{D}_{L0}$ is distributed without any statistical noise. An example of L0 data is shown in figure 4.1, where the data points of a measured physical observable, the cross section $\frac{d\sigma}{d|q_T|}$, are perfectly

¹This is not the actual underlying law, which is unknown, but rather an artificial "real" law represented by a chosen set of parameters.

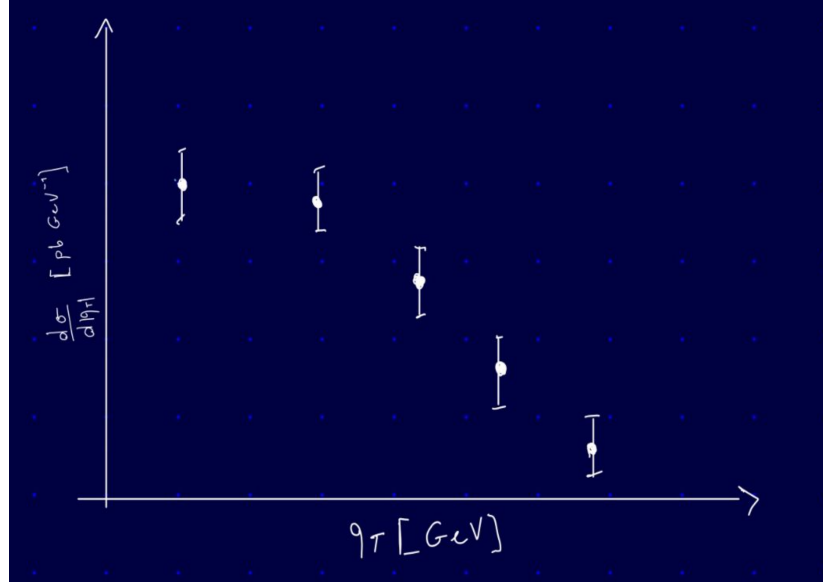


Figure 4.1: L0 data for a physical observable $\frac{d\sigma}{d|q_T|}$. The points are perfectly distributed on the theoretical functional form of the observable.

aligned with the theoretical functional form of the observable, which was predicted using u_{real} .

Unlike an experimental dataset (see equation (3.25)), where data points are subject to fluctuations and we do not know the non-fluctuated values (due to the unknown underlying law), the L0 dataset is noise-free. We can then add controlled fluctuations to the L0 pseudo-data. By introducing Gaussian-distributed fluctuations η , we obtain the so called Level 1 (L1) dataset,

$$\mathcal{D}_{L1} = \mathcal{G}(u_{\text{real}}) + \eta = \mathcal{D}_{L0} + \eta , \quad (4.2)$$

which simulates a real experimental dataset in which each data is fluctuated around a central point. The key difference between a real dataset and the L1 dataset is that in the L1 dataset, the fluctuation is controlled, meaning we know both the central values (that correspond to the non-fluctuated dataset, $\mathcal{D}_{L0}$) and the size of the fluctuations. To define these fluctuations, η , consider the point i in the dataset $\mathcal{D}_{L0}$, denoted as $y_{L0,i}$, which has an associated error σ_i . For simplicity, we consider σ_i as a single uncorrelated uncertainty; the complete structure of the uncertainties associated to each data-point will be discussed in section 4.8. The fluctuation of the i -th point in the L1 dataset is a random number drawn from a normal distribution centered at zero with a standard deviation of σ_i ,

$$\eta_i = \mathcal{N}(0, \sigma_i) . \quad (4.3)$$

In equation (4.2), η represents the vector containing all the fluctuations η_i . Figure 4.2 illustrates the L1 dataset for the physical observable $\frac{d\sigma}{d|q_T|}$. The L1 data points (shown in yellow) are generated by adding Gaussian fluctuations to the corresponding L0 points

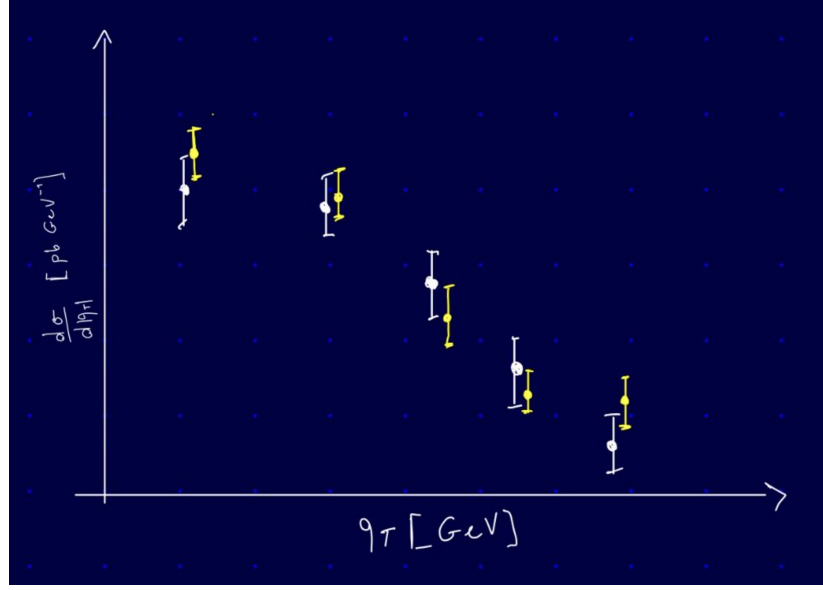


Figure 4.2: L1 data for a physical observable $\frac{d\sigma}{d|q_T|}$ (in yellow). The white points represent the L0 data, the yellow points are the fluctuated data.

(shown in white).

A third layer of Gaussian fluctuations, ϵ , can be added to generate the Level 2 (L2) dataset. Similar to η , ϵ is defined as the vector of fluctuations $\epsilon_i = \mathcal{N}(0, \sigma_i)$. To perform an L2 closure test, a large number N_{rep} of L2 fluctuations are generated. The k -th L2 dataset is constructed by adding the fluctuation $\epsilon^{(k)}$ to the L1 dataset,

$$\mathcal{D}_{L2}^{(k)} = \mathcal{G}(u_{\text{real}}) + \eta + \epsilon^{(k)} = \mathcal{D}_{L1} + \epsilon^{(k)} . \quad (4.4)$$

The L2 fluctuations are analogous to those added to a real dataset during resampling with Monte Carlo replicas, as done in the bootstrap approach to inverse problems, which is used in the MAPTMD22 and PV19 frameworks. The L2 dataset thus consists of an ensemble of N_{rep} fluctuated datasets:

$$\mathcal{D}_{L2} = \{\mathcal{D}_{L2}^{(k)}\} . \quad (4.5)$$

Following the previous example, the L2 dataset (figure 4.3) for the cross section $\frac{d\sigma}{d|q_T|}$ is obtained by generating N_{rep} Monte Carlo replicas of the L1 dataset, yielding a Gaussian-distributed set of L2 data points around each original L1 point. Table 4.1 summarizes the fluctuation levels used in closure testing. The errors associated with the L0, L1, and L2 pseudo-data are the same experimental errors of the real values in the original dataset $\mathcal{D}$.

Once the datasets $\mathcal{D}_{L0}$, $\mathcal{D}_{L1}$, and $\mathcal{D}_{L2}$ are generated, the final step of the closure test is to apply the fitting framework under evaluation to these artificially generated datasets. These three datasets allow us to operate in a controlled environment, where

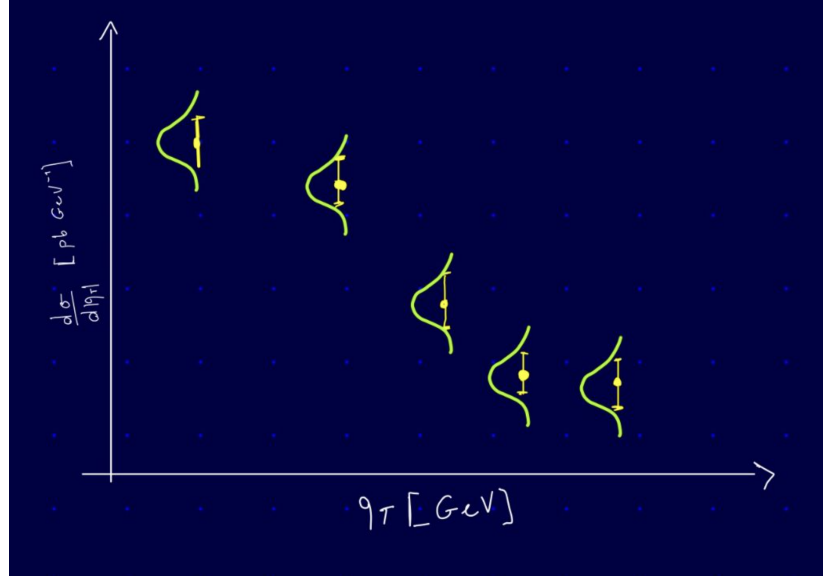


Figure 4.3: L2 data for $\frac{d\sigma}{d|q_T|}$. The L2 dataset is composed of N_{rep} Monte Carlo replicas of each L1 data-point.

Level	Dataset Structure	Corresponding to
Level 0 (L0)	$\mathcal{G}(u_{\text{real}})$	dataset without experimental error
Level 1 (L1)	$\mathcal{G}(u_{\text{real}}) + \eta$	real dataset with noise
Level 2 (L2)	$\mathcal{G}(u_{\text{real}}) + \eta + \epsilon$	Monte Carlo replica of a real dataset

Table 4.1: Levels of data fluctuations in closure testing

both the perfectly distributed central points, derived from the known input u_{real} , and the fluctuations applied to them are fully understood. This setup enables us to carefully examine the robustness of the fitting framework by comparing the fits to the L0, L1, and L2 pseudo-data with the known input used to generate those data.

For each testing level, we generate a large number of replicas of the dataset and fit them, starting from various points in the parameter space. From the fit of the k -th replica, we obtain the predicted input u_*^k , which represents the set of parameters used to reconstruct the functional form of the non-perturbative TMDs. The fitted TMDs are calculated by averaging over the replicas. For instance, the fitted TMD PDF is computed as

$$\langle f_{1NP,*} \rangle = \frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} f_{1NP,*}^k, \quad (4.6)$$

where $f_{1NP,*}^k$ is the predicted TMD PDF from the fit of the k -th replica. The variance of $f_{1NP,*}$ is given by

$$\sigma^2 = \langle (f_{1NP,*}^{(k)} - \langle f_{1NP,*} \rangle)^2 \rangle = \frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} (f_{1NP,*}^{(k)} - \langle f_{1NP,*} \rangle)^2 . \quad (4.7)$$

In equations (4.6) and (4.7), we omitted the dependences on (x, b_T, ζ) for conciseness.

Fitting data with different levels of fluctuation (L0, L1, and L2) provides insights into the flexibility of the methodology and the extent of the uncertainties that it introduces. In the following, we will delve into the details of the information that can be extracted from each fluctuation level.

4.2 Level 0 Test

A Level 0 closure test involves using the fitting framework to fit the L0 dataset, $\mathcal{D}_{L0}$. A complete Level 0 closure test consists of repeating the fit on $\mathcal{D}_{L0}$, multiple times, each time starting the optimization algorithm from a different point in the parameter space. The Level 0 test allows us to evaluate two key aspects of the framework: the flexibility of the methodology and the interpolation/extrapolation uncertainty.

Flexibility of the Methodology

When fitting a non-fluctuated dataset, there exists a perfect solution, *i.e.* a functional form that exactly interpolates each data point in the dataset, given by $\mathcal{G}(u_{\text{real}})$. If the methodology applied to the $\mathcal{D}_{L0}$ dataset is capable of finding this perfect solution, the functional form of the predicted and real input will match,

$$u_* = u_{\text{real}} . \quad (4.8)$$

In this case, the dataset predicted by the methodology,

$$\mathcal{D}_{\text{pred}} = \mathcal{G}(u_*) , \quad (4.9)$$

will coincide with the L0 dataset, $\mathcal{D}_{L0} = \mathcal{G}(u_{\text{real}})$, resulting in

$$\mathcal{D}_{\text{pred}} = \mathcal{D}_{L0} . \quad (4.10)$$

Consequently, the χ^2 between the two datasets (see equation (3.19)) will be zero,

$$\chi^2[\mathcal{D}_{\text{pred}}, \mathcal{D}_{L0}] = 0 . \quad (4.11)$$

The methodology will pass the L0 closure test if the final $\chi^2[\mathcal{D}_{\text{pred}}, \mathcal{D}_{L0}]$ is sufficiently close to zero. In this analysis, the criterion for success in the L0 test is

$$\chi^2[\mathcal{D}_{\text{pred}}, \mathcal{D}_{L0}] < 10^{-10} . \quad (4.12)$$

The L0 test evaluates the flexibility of the methodology. If the methodology is sufficiently flexible, it can reproduce the functional form of the input parametrization used to generate the pseudo-data, u_{real} . Importantly, we compare the functional forms

of the inputs u_{real} and u_* , not the parameter values, since the goal of a parametrization is to capture the functional form of the input. In fact, the same form can be represented by different parameter sets. Furthermore, it is not necessary for u_{real} and u_* to share the same parametrization. We can test a methodology using a parametrization different from the one used to generate the pseudo-data. For instance, if we use the PV19 framework to fit data generated using the input $u_{*,\text{MAP22}}$, and the PV19 methodology is sufficiently flexible, the functional form predicted with $u_{*,\text{PV19}}$ will match the functional form used to generate the data, $u_{*,\text{MAP22}}$ ensuring the success of the closure test.

Interpolation-Extrapolation Uncertainty

Even though the points in the L0 dataset are perfectly distributed and the χ^2 of the perfect solution is zero, there remains an inherent uncertainty due to the finite nature of the dataset: the interpolation/extrapolation uncertainty. An infinite number of functional forms can capture the data distribution equally well, interpolating exactly through all data points and yielding $\chi^2[\mathcal{D}_{\text{pred}}, \mathcal{D}_{L0}] = 0$. This is because the functional form is not constrained in the intervals between the data points and in regions where no experimental measurements exist. If the L0 closure test succeeds, meaning the methodology finds a solution where $\chi^2[\mathcal{D}_{\text{pred}}, \mathcal{D}_{L0}] < 10^{-10}$, and the test is repeated N_{init} times², using the same dataset and starting each fit from a different point in the parameter space, we obtain a set of N_{init} functional forms that meet the condition $\chi^2[\mathcal{D}_{\text{pred}}, \mathcal{D}_{L0}] < 10^{-10}$. This set of solutions, $\{u_*\}$, can exhibit variation in regions where no experimental data are available. The functional spread between data points represents the interpolation uncertainty, while the spread in kinematic regions not covered by the dataset corresponds to the extrapolation uncertainty. Figure 4.4 illustrates the functional spread in regions where no experimental information is available.

Thus, a Level 0 closure test can be used to quantify the interpolation/extrapolation uncertainty, σ_{int} , introduced in the TMDs by the fitting framework, which we define as the standard deviation of the TMDs predicted across the L0 fits (σ_{L0}):

$$\sigma_{L0}(x, b_T, \zeta) = \sqrt{\langle (f_{NP,*}^{(k),L0} - \langle f_{NP,*}^{L0} \rangle)^2 \rangle}, \quad (4.13)$$

where $\langle f_{NP,*}^{L0} \rangle$ is the mean value of the non-perturbative TMD at (x, b_T, ζ) , and $f_{NP,*}^{(k)}$ is the predicted non-perturbative TMD by the fit on the k -th L0 replica. The Level 0 standard deviation corresponds the interpolation/extrapolation uncertainty,

$$\sigma_{L0} = \sigma_{\text{int}}. \quad (4.14)$$

Computing σ_{L0} is only possible with a controlled dataset, as the true underlying values are unknown in a real dataset. By introducing fluctuations into the L0 dataset, we can further study other contributions to the total uncertainty.

²For the number of replicas of the L0 test, N_{init} , the label "init" stays for "initialization"; the replicas in a L0/L1 test are conceptually different by those in a L2 test, because in the L0 and L1 tests we do not fluctuate the replicas, instead we use N_{init} times the same dataset, where the initialization of the optimization process changes each time.

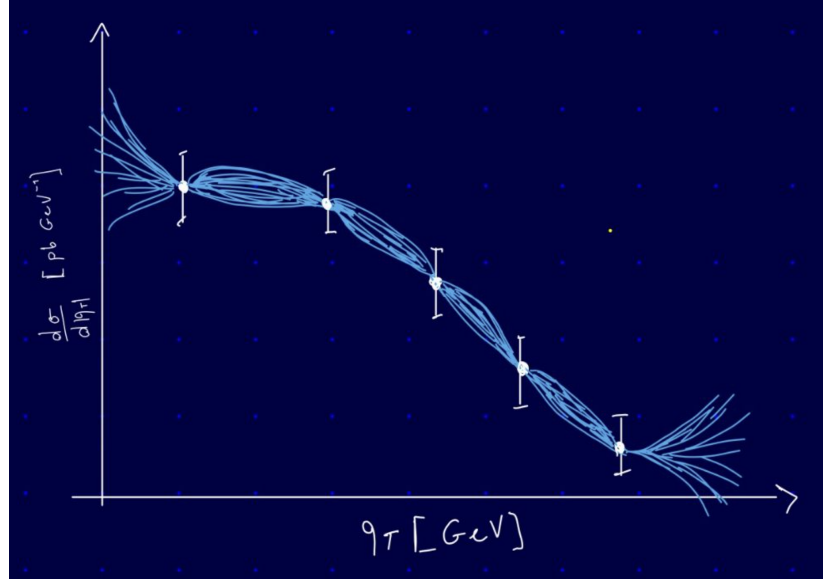


Figure 4.4: Interpolation/extrapolation uncertainty. This uncertainty is associated with the variability of the functional form in the intervals between data points, as well as in kinematic regions not covered by the dataset.

4.3 Level 1 Test

To conduct a Level 1 test, we use the fitting framework to fit the fluctuated dataset $\mathcal{D}_{L1}$ (equation (4.2)). Similar to the L0 test, the fit is repeated a number N_{init} of times, each starting from a different point in the parameter space. Since each L1 data fluctuation is artificially generated following a Gaussian distribution with the standard deviation of each data point, the mean χ^2 between the L0 and L1 datasets approaches 1 as the number of data points increases. For a large number of data points, we have

$$\langle \chi^2[\mathcal{D}_{L1}, \mathcal{D}_{L0}] \rangle = \frac{1}{N_{\text{data}}} \chi^2[\mathcal{D}_{L1}, \mathcal{D}_{L0}] \approx 1. \quad (4.15)$$

Thus, for the L1 dataset, the optimal functional form is not one that perfectly interpolates all data points, but one that generates a predicted dataset $\mathcal{D}_{\text{pred}}$ satisfying

$$\langle \chi^2[\mathcal{D}_{\text{pred}}, \mathcal{D}_{L1}] \rangle = \frac{1}{N_{\text{data}}} \chi^2[\mathcal{D}_{\text{pred}}, \mathcal{D}_{L1}] \approx 1. \quad (4.16)$$

The L1 closure test is considered successful if the χ^2 distribution, over a large number N_{init} of tests on the same dataset $\mathcal{D}_{L1}$ starting from different points in the parameter space, is compatible with 1.

The uncertainty we examine in the L1 test, associated with the choice of functional form, is called functional uncertainty. In this case, there can be an infinite quantity of distributions $\mathcal{G}(u_*)$ that approximate the fluctuated L1 dataset equally well. The functional spread of the set of predicted distributions $\{f_{NP}(u_*)\}$ from a large number of

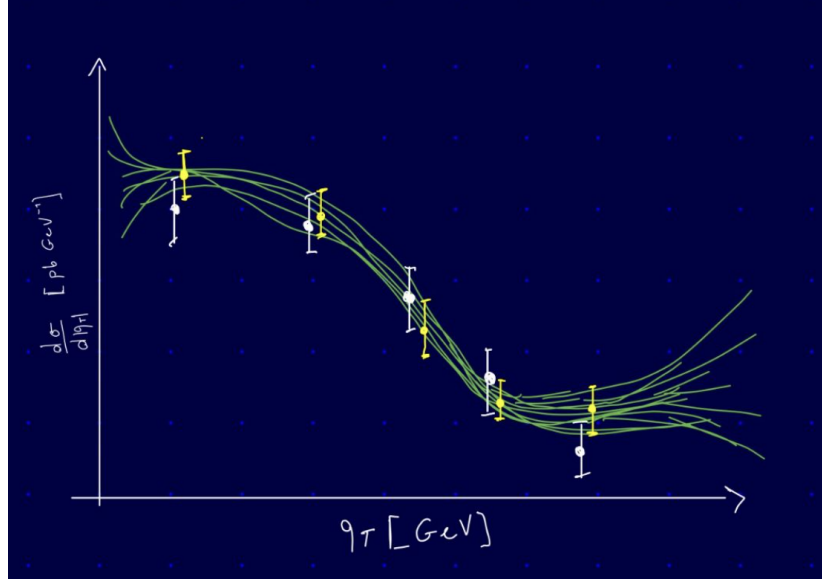


Figure 4.5: Functional uncertainty. This uncertainty is associated to the fluctuations of the dataset and is given by the spread of a large number of predictions. All the predicted functional forms approximate equivalently well the L1 data distribution.

tests represents the combined interpolation/extrapolation and functional uncertainties of the TMDs. The Level 1 uncertainty, σ_1 , defined as

$$\sigma_{L1}(x, b_T, \zeta) = \sqrt{\langle (f_{NP,*}^{(k),L1} - \langle f_{NP,*}^{L1} \rangle)^2 \rangle}, \quad (4.17)$$

is the combination of the interpolation/extrapolation and functional uncertainty,

$$\sigma_{L1} = \sigma_{L0} + \sigma_{\text{func}}, \quad (4.18)$$

where σ_{func} is the functional uncertainty. Figure 4.5 shows the functional spread obtained in a Level 1 closure test.

In the L1 test, we are using the same dataset for each of the N_{init} fits, without generating N_{rep} Monte Carlo replicas of the L1 dataset. As a result, the experimental error is not propagated into the error of the predicted input $f_{NP,*}$ (equation (3.27)). To explore how experimental error affects the predictions, we move to the Level 2 closure test.

4.4 Level 2 Test

The Level 2 test provides a quantitative estimate of the total uncertainty in the predicted TMDs. By comparing the uncertainties from the L1 and L2 tests, one can assess how the experimental uncertainty propagates into the predicted TMDs. The L2 test is performed using a dataset consisting of a large number N_{rep} of Monte Carlo replicas of the L1 dataset (equation (4.4)). Each replica is fitted using the framework under

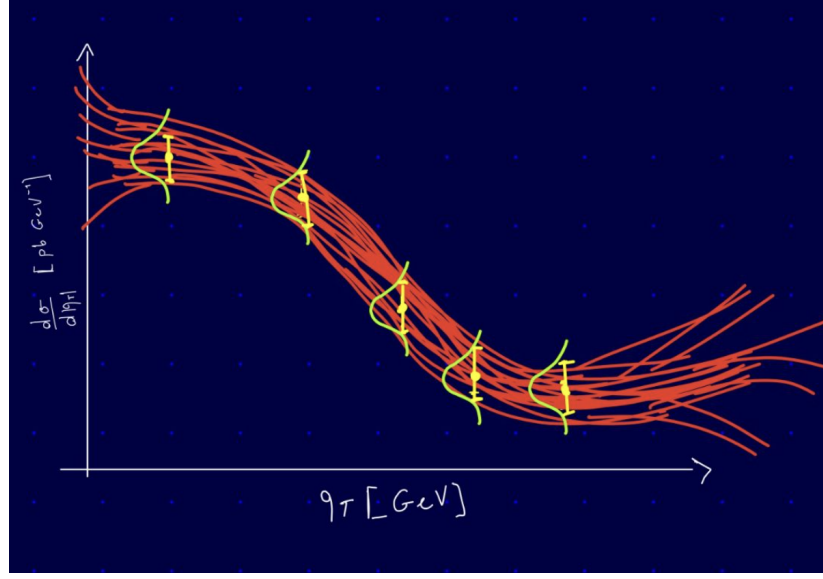


Figure 4.6: The spread of the set of predictions $\{u_*\}$ in a L2 test. This represent the total uncertainty.

analysis, with the optimization starting from a different point in the parameter space. In a L2 test, we both fluctuate the dataset N_{rep} *and* start the optimization process in N_{init} different points; in principle, N_{rep} and N_{init} could differ, but in practice we always set $N_{\text{init}} = N_{\text{rep}}$, such that each fit on a different Monte Carlo replica also starts from a different point in the parameter space. In this case, the spread in the set of predicted distributions $\{f_{NP}(u_*)\}$ is larger and we expect it to estimate the total uncertainty, given by the combination of interpolation/extrapolation, functional, and experimental uncertainty,

$$\sigma_{L2} = \sigma_{L1} + \sigma_{\text{exp}} , \quad (4.19)$$

where σ_{L2} is the Level 2 standard deviation and σ_{exp} is the experimental uncertainty. The L2 standard deviation is defined as

$$\sigma_{L2}(x, b_T, \zeta) = \sqrt{\langle (f_{NP,*}^{(k),L2} - \langle f_{NP,*}^{L2} \rangle)^2 \rangle} . \quad (4.20)$$

The functional spread resulting from the L2 fit is illustrated in figure 4.6.

Table 4.2 summarizes the layers of uncertainty introduced at each testing level. In each level of testing, the starting point of the optimization algorithm changes for each replica, but only in the Level 2 test are Monte Carlo fluctuations applied to generate different replicas.

To summarize, the L0 closure test allows us to assess the flexibility of a methodology, while the combination of L0, L1, and L2 tests helps estimate the contributions of interpolation/extrapolation, functional, and experimental uncertainties to the total uncertainty of the predicted TMDs by examining the spread of the predicted TMDs

Level	Monte Carlo Fluctuations	Uncertainty
Level 0 (L0)	no	interpolation/extrapolation (int/ext)
Level 1 (L1)	no	int/ext + functional
Level 2 (L2)	yes	int/ext + functional + experimental

Table 4.2: Uncertainty layers introduced by each level of fluctuation.

$\{f_{NP}(u_*)\}$ at each level of testing.

If the uncertainties predicted by the methodology are reliable, we expect the real functional form u_{real} (used to generate the pseudo-data) to lie within the one-sigma total error band of the fitted functional forms u_* with a probability of around 68% [6]. The faithfulness of these uncertainties, along with the bias and variance of the predictions, is evaluated using the so-called multi closure test.

4.5 Multi Closure Tests

In the previous sections, we explained how Level 0, Level 1, and Level 2 closure tests allow us to decompose the total TMD uncertainty into contributions from experimental, functional, and interpolation/extrapolation uncertainties. Here, we introduce multi closure tests to estimate the bias and variance of the fitted TMDs, along with two quantitative estimators of the framework's robustness: the bias-variance ratio ($\mathcal{R}_{bv}$) and the quantile estimator ($\xi_{n\sigma}$).

A multi closure test involves performing a large number of fits N_{fit} , where each fit consists of an ensemble of N_{rep} Monte Carlo replicas (L2 fluctuations) of a Level 1 pseudo-dataset. In this process, we start by selecting a "real" input u_{real} to generate the L0 dataset. We then create N_{fit} Monte Carlo replicas of the L0 dataset, producing a set of N_{fit} L1 datasets. For each L1 dataset, we generate N_{rep} Monte Carlo replicas, yielding a set of Level 2 fluctuated datasets.

Model Error

Consider a single fit consisting of N_{rep} replicas of an L1 dataset $\mathcal{D}_{L1}$. The model error $\mathcal{E}^{(k)}$ for a single replica is defined as the χ^2 between the model predictions $\mathcal{G}(u_*^{(k)})$ and the distribution of the L1 data points $\mathcal{D}_{L1}$,

$$\mathcal{E}^{(k)} = \frac{1}{N_{\text{dat}}} (\mathcal{G}(u_*^{(k)}) - \mathcal{D}_{L1})^T C_{L1}^{-1} (\mathcal{G}(u_*^{(k)}) - \mathcal{D}_{L1}) , \quad (4.21)$$

where $(\mathcal{G}(u_*^{(k)}) - \mathcal{D}_{L1})$ is the vector of distances between predictions and their corresponding data points, and C_{L1} is the covariance matrix of the L1 dataset, equivalent to the covariance matrix C of the real dataset $\mathcal{D}$.

We compute the model error of the fit by averaging over all the replicas,

$$\mathcal{E}^{\text{fit}} = \mathbb{E}_{\{u_*\}}[\mathcal{E}^{(k)}] = \frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} \mathcal{E}^{(k)} . \quad (4.22)$$

Equation (4.22) can be used to evaluate the error of a fit on real experimental data by replacing $\mathcal{D}_{L1}$ with $\mathcal{D}$, as both datasets exhibit the same level of fluctuation (L1). However, using the L1 pseudo-data enables further decomposition of equation (4.22). Since we know the L0 values $\mathcal{D}_{L0} = \mathcal{G}(u_{\text{real}})$ from which the L1 dataset is derived, we can express

$$(\mathcal{G}(u_*^{(k)}) - \mathcal{D}_{L1}) = (\mathcal{G}(u_*^{(k)}) - \mathcal{D}_{L0}) + (\mathcal{D}_{L0} - \mathcal{D}_{L1}) , \quad (4.23)$$

resulting in the model error of the fit as

$$\begin{aligned} \mathcal{E}^{\text{fit}} = \frac{1}{N_{\text{dat}}} \{ & \mathbb{E}_{\{u_*\}}[(\mathcal{G}(u_*^{(k)}) - \mathcal{D}_{L0})^T C^{-1} (\mathcal{G}(u_*^{(k)}) - \mathcal{D}_{L0})] + \\ & + \mathbb{E}_{\{u_*\}}[(\mathcal{D}_{L0} - \mathcal{D}_{L1})^T C^{-1} (\mathcal{D}_{L0} - \mathcal{D}_{L1})] + \\ & + 2\mathbb{E}_{\{u_*\}}[(\mathcal{G}(u_*^{(k)}) - \mathcal{D}_{L0})^T C^{-1} (\mathcal{D}_{L0} - \mathcal{D}_{L1})] \} . \quad (4.24) \end{aligned}$$

The second term of this equation represents the shift between the L0 and L1 datasets, which remains constant across replicas. The third term is a cross term, which vanishes when averaged across a large number of replicas. The first term in equation (4.25), $\mathbb{E}_{\{u_*\}}[(\mathcal{G}(u_*^{(k)}) - \mathcal{D}_{L0})^T C^{-1} (\mathcal{G}(u_*^{(k)}) - \mathcal{D}_{L0})]$, can be further decomposed by introducing a vector representing the average of model predictions across replicas, $\mathbb{E}_{\{u_*\}}[\mathcal{G}(u_*^{(k)})]$, leading to

$$\begin{aligned} \mathbb{E}_{\{u_*\}}[(\mathcal{G}(u_*^{(k)}) - \mathcal{D}_{L0})^T C^{-1} (\mathcal{G}(u_*^{(k)}) - \mathcal{D}_{L0})] = \\ = \mathbb{E}_{\{u_*\}} [(\mathcal{G}(u_*^{(k)}) - \mathbb{E}_{\{u_*\}}[\mathcal{G}(u_*^{(k)})])^T C^{-1} (\mathcal{G}(u_*^{(k)}) - \mathbb{E}_{\{u_*\}}[\mathcal{G}(u_*^{(k)})])] + \\ + ((\mathbb{E}_{\{u_*\}}[\mathcal{G}(u_*^{(k)})] - \mathcal{D}_{L0})^T C^{-1} (\mathbb{E}_{\{u_*\}}[\mathcal{G}(u_*^{(k)})] - \mathcal{D}_{L0})) , \quad (4.25) \end{aligned}$$

where we omitted the cross term, which goes to zero when averaging across the replicas. The average across replicas is taken only for the first term in equation (4.25), because the second term is the same for all replicas.

We can finally identify the variance and the bias of a single fit (corresponding to a single L1 dataset) with the first and second term of equation (4.25), respectively. Notably, the model error $\mathcal{E}^{\text{fit}}$, along with the bias and variance of a single fit, are stochastic values, meaning we cannot use them as quantitative estimators for the robustness of a fitting framework. To define a robust estimator, we take a large number of fits (corresponding to a large number of L1 fluctuations) and we average the model error, obtaining

$$\mathbb{E}_{\mathcal{D}_{L1}}[\mathcal{E}^{\text{fit}}] = \mathbb{E}_{\mathcal{D}_{L1}}[\text{bias}] + \mathbb{E}_{\mathcal{D}_{L1}}[\text{variance}] + \mathbb{E}_{\mathcal{D}_{L1}}[\text{noise}] , \quad (4.26)$$

where we defined the average across the fits as

$$\mathbb{E}_{\mathcal{D}_{L1}}[\mathcal{E}^{\text{fit}}] = \frac{1}{N_{\text{fit}}} \sum_{i=1}^{N_{\text{fit}}} \mathcal{E}^{\text{i-th fit}}. \quad (4.27)$$

In equation (4.26), the "noise" term corresponds to the second term of equation (4.25). Since this term is given by the statistical fluctuations of the data and is not dependent by the framework, it will not be used in the construction of the bias-variance ratio estimator, which we define in the next section.

Bias-Variance Ratio Estimator

We can use the bias and variance averaged over a large number of fits to estimate the robustness of a model. Here, we include the input u in what we refer to as a *model*, as model complexity increases with the complexity of the input parametrization. In particular, both bias and variance depend on model complexity, as illustrated in figure 4.7. If the model is not sufficiently complex to capture the distribution of the dataset, we encounter a problem of underfitting. In this case, the predictions $u_*^{(k)}$ will be close to each other (low variance), but the ensemble of predictions $\{u_*^{(k)}\}$ will be far from the true underlying distribution u_{real} that we are trying to approximate³ (high bias). This scenario is represented on the left side of figure 4.7. Conversely, if the model is too complex for the data distribution, we face an overfitting problem, where the parametrization is flexible enough to fit the noise in the dataset. This results in predictions $u_*^{(k)}$ that depend on the fluctuations of the dataset and are far from each other (high variance), but the ensemble of predictions, $\{u_*^{(k)}\}$, is centered in the real law u_{real} that we are searching. This situation is depicted on the right side of the figure. The optimal case is the one represented in the center of figure 4.7, where we are in a minimum of the model error (equation (4.26)), the variance and bias of the fit have the same value, and the distribution of the predictions is narrow and centered in the real law that we are searching.

Following Ref [5], we construct the bias-variance ratio estimator:

$$\mathcal{R}_{vb} = \sqrt{\frac{\mathbb{E}_{\mathcal{D}_{L1}}[\text{bias}]}{\mathbb{E}_{\mathcal{D}_{L1}}[\text{variance}]}} \quad (4.28)$$

where we take the square root of both quantities because bias and variance are mean squared metrics. In an optimized methodology, the values of variance and bias are compatible; thus, the analyzed models will pass the multi closure test if they satisfy the condition

$$\mathcal{R}_{vb} \approx 1, \quad (4.29)$$

where with " ≈ 1 " indicates that 1 lies within the error bar $\sigma_{\mathcal{R}}$ of the bias-variance ratio, which is defined by propagating the errors of the mean bias and variance, $\sigma_{b,m}$ and $\sigma_{v,m}$:

³or, equivalently, the data predictions of each replica will be close to each other but far from the data point of the fitted dataset

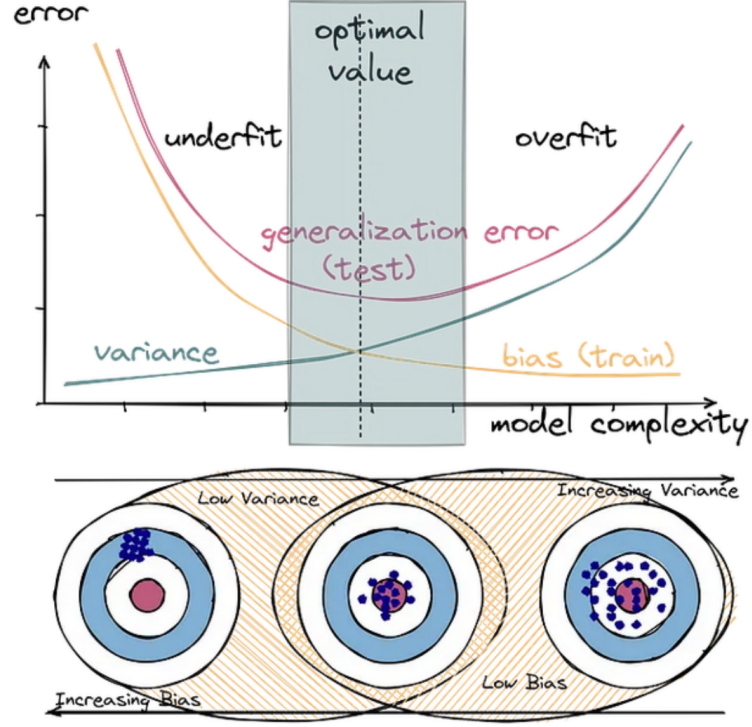


Figure 4.7: The behavior of bias and variance depending on the model complexity and the distribution of the predictions of a data point depending on the values of bias and variance. Image from Ref. [46]

$$\sigma_{\mathcal{R}} = \sqrt{\left(\frac{1}{2} \sqrt{\mathbb{E}_{\mathcal{D}_{L1}}[\text{bias}] \mathbb{E}_{\mathcal{D}_{L1}}[\text{variance}]} \sigma_{b,m}\right)^2 + \left(\frac{1}{2} \sqrt{\frac{\mathbb{E}_{\mathcal{D}_{L1}}[\text{bias}]}{\mathbb{E}_{\mathcal{D}_{L1}}[\text{variance}]^3}} \sigma_{v,m}\right)^2}, \quad (4.30)$$

where

$$\sigma_{v,m} = \frac{\sigma_v}{\sqrt{N_{\text{fit}}}} \quad (4.31)$$

$$\sigma_{b,m} = \frac{\sigma_b}{\sqrt{N_{\text{fit}}}}, \quad (4.32)$$

in which σ_v and σ_b represent the standard deviations of the variance and bias distributions, respectively, across the fits.

Quantile Estimator

A second quantitative estimator can be computed through a multi closure test, the quantile estimator [5]. As discussed in section 4.4, the uncertainties on the TMDs reproduced by the analyzed fitting framework are reliable if the true TMD functional

form lies within the one-sigma band of the TMDs predicted by the framework in approximately 68% of the fits. In a multi-closure test, we generate a large number of fits and know the functional forms of the TMDs used to produce the datasets. This allows us to count how often the true functional form falls within the error band of the predicted functional form. Following Ref [5], we perform this count in data space rather than input space, meaning we count how many points of the L0 dataset fall within the error bars of the predicted dataset for each fit. This count is implemented using an indicator function,

$$I_{[-n\sigma_i, n\sigma_i]} (\mathbb{E}_{\{u_*\}} [\mathcal{G}(u_*^{(k)})]_i^\ell - \mathcal{D}_{L0,i}) , \quad (4.33)$$

where $\mathcal{D}_{L0,i}$ is the i -th non-fluctuated data point, $\mathcal{G}(u_*^{(k)})_i^\ell$ is mean value of the i -th point over the replicas of the ℓ -th fit, and σ_i is the standard deviation associated to the value $\mathbb{E}_{\{u_*\}} [\mathcal{G}(u_*^{(k)})]_i^\ell$. The indicator function is defined to be one if the argument is within the range $[-n\sigma_i, n\sigma_i]$, where $n \in \mathbb{N}$, and zero otherwise,

$$I_{[-n\sigma_i, n\sigma_i]}(x) = \begin{cases} 1 & \text{if } x \in [-n\sigma_i, n\sigma_i] \\ 0 & \text{if } x \notin [-n\sigma_i, n\sigma_i] \end{cases} . \quad (4.34)$$

The quantile estimator $\xi_{n\sigma}$ counts the proportion of L0 data points that lie within the n -sigma band of the predicted dataset,

$$\xi_{n\sigma} = \frac{1}{N_{\text{dat}}} \sum_{i=1}^{N_{\text{dat}}} \frac{1}{N_{\text{fit}}} \sum_{\ell=1}^{N_{\text{fit}}} I_{[-n\sigma_i, n\sigma_i]} (\mathbb{E}_{\{u_*\}} [\mathcal{G}(u_*^{(k)})]_i^\ell - \mathcal{D}_{L0,i}) . \quad (4.35)$$

Our methodology faithfully reproduces the TMD uncertainties if, after computing the quantile estimator for a large number of fits, we find

$$\xi_{1\sigma} \approx 0.683 \quad (4.36)$$

$$\xi_{2\sigma} \approx 0.954 \quad (4.37)$$

$$\xi_{3\sigma} \approx 0.997. \quad (4.38)$$

4.6 L0 Pseudo-Data Generation

In this section, we discuss the generation of the pseudo-datasets used to test the MAPTMD22 and PV19 fitting frameworks. We generated three sets of L0 pseudo-data using three different inputs, u_{real} . Additionally, we fluctuated two of these three pseudo-datasets, obtaining two sets of L1 and L2 fluctuated datasets, generated using two distinct inputs.

Input Parametrizations

To generate the L0 pseudo-data, we chose the input u_{real} to act as the "real" underlying law. We selected three sets of non-perturbative TMDs as the real law:

Input Name	TMD PDF	TMD FF	Input Dimension
$u_{\text{MAP22}}^{\text{real}}$	MAPTMD22	MAPTMD22	$\mathbb{R}^{21}$
$u_{\text{PV19}}^{\text{real}}$	PV19	/	$\mathbb{R}^9$
$u_{\text{MIX24}}^{\text{real}}$	PV19	MAPTMD22	$\mathbb{R}^{19}$

Table 4.3: Input parametrizations used to generate three different set of L0 pseudo-data.

1. **MAP22:** The set of parameters obtained by fitting experimental data using the MAPTMD22 framework. Using these parameters with the MAPTMD22 parametrization described in equations (3.12) and (3.14), we obtain the functional form of the non-perturbative TMD PDF and FF, assumed to be the true TMDs we are searching. This input parametrization is referred to as $u_{\text{MAP22}}^{\text{real}}$. The parameter values used to generate the pseudo-data are listed in the first column of table 4.4.
2. **PV19:** The set of parameters obtained by fitting experimental data using the PV19 framework. Using these parameters with the PV19 parametrization (equation (3.16)), we derive the non-perturbative TMD PDF, assumed to be the real input. This input parametrization is called $u_{\text{PV19}}^{\text{real}}$. The parameter values used to generate the pseudo-data are listed in the second column of table 4.4.
3. **MIX24:** A mixed set of parameters, combining the TMD PDF parametrization from PV19 (equation (3.16)), with the TMD FF from MAPTMD22 (equation (3.14)). This set combines the TMD PDF parameters from PV19 with the TMD FF parameters from MAPTMD22. This input parametrization is referred to as $u_{\text{MIX24}}^{\text{real}}$. The parameter values used to generate the pseudo-data are shown in the third column of table 4.4.

Note that the values of the parameters shown in table 4.4 derive from fits on real data performed with the respective parameterizations, causing the parameters used in MIX24 to have different numerical values with respect to those used in PV19 and MAPTMD22.

Datasets Level 0

Table 4.4 shows the TMD PDF and FF parametrizations used as the real input to generate the three L0 pseudo-datasets. Using the inputs $u_{\text{MAP22}}^{\text{real}}$ and $u_{\text{MIX24}}^{\text{real}}$ we can reproduce the full dataset, including both Drell-Yan and SIDIS cross sections. With the input $u_{\text{PV19}}^{\text{real}}$, only the data derived from Drell-Yan processes are reproduced, as the TMD FF is not included in this input.

The three L0 datasets generated using these inputs are denoted as $\mathcal{D}_{L0}^{\text{MAP22}}$, $\mathcal{D}_{L0}^{\text{PV19}}$, and $\mathcal{D}_{L0}^{\text{MIX24}}$. The MAPTMD22 model, $\mathcal{G}_{\text{MAP22}}$, is used to generate the first and third datasets,

non perturbative TMD	$u_{\text{MAP22}}^{\text{real}}$	$u_{\text{PV19}}^{\text{real}}$	$u_{\text{MIX24}}^{\text{real}}$
S_{NP}	$g_2 = 0.248624$		$g_2 = 0.070105$
f_{1NP}	$\alpha_1 = 1.153911$		
	$\alpha_2 = 4.274540$	$g_2 = 0.037923$	
	$\alpha_3 = 4.375137$	$N_1 = 0.518814$	$N_1 = -0.045526$
	$\sigma_1 = 0.519140$	$\alpha = 0.203142$	$\alpha = 1.365441$
	$\sigma_2 = 0.413861$	$\sigma = 0.373272$	$\sigma = 0.576503$
	$\sigma_3 = 12.625959$	$\lambda = 0.579733$	$\lambda = 4.2780005$
	$\lambda = 1.742679$	$N_{1B} = 0.039554$	$N_{1B} = 0.955875$
	$\lambda_2 = 0.020872$	$\alpha_B = 0.067697$	$\alpha_B = 3.351724$
	$N_1 = 0.318114$	$\sigma_B = 0.363078$	$\sigma_B = 0.632127$
	$N_{1B} = 0.129140$	$g_{2B} = 0.011222$	$g_{2B} = 0.017748$
D_{1NP}	$N_{1C} = 0.015973$		
	$N_3 = 0.004335$		$N_3 = 0.004041$
	$N_{3B} = 0.217415$		$N_{3B} = 0.241838$
	$\beta_1 = 10.845024$		$\beta_1 = 11.391522$
	$\beta_2 = 4.267465$		$\beta_2 = 4.599514$
	$\delta_1 = 0.007126$		$\delta_1 = 0.008340$
	$\delta_2 = 0.161130$		$\delta_2 = 0.159940$
	$\gamma_1 = 1.532915$		$\gamma_1 = 1.506459$
	$\gamma_2 = -0.002678$		$\gamma_2 = -0.009876$
	$\lambda_F = 0.062387$		$\lambda_F = 0.075544$

Table 4.4: Values of the parameters used as input to generate the pseudodata.

$$\mathcal{D}_{L0}^{\text{MAP22}} = \mathcal{G}_{\text{MAP22}}(u_{\text{MAP22}}^{\text{real}}) \quad (4.39)$$

$$\mathcal{D}_{L0}^{\text{MIX24}} = \mathcal{G}_{\text{MAP22}}(u_{\text{MIX24}}^{\text{real}}) , \quad (4.40)$$

while the second dataset is generated using the PV19 model $\mathcal{G}_{\text{MAP22}}$ (which is equivalent to the Drell-Yan part of the MAPTMD22 model),

$$\mathcal{D}_{L0}^{\text{PV19}} = \mathcal{G}_{\text{PV19}}(u_{\text{PV19}}^{\text{real}}). \quad (4.41)$$

The artificially generated datasets consist of the same experimental points included in the original MAP fitting frameworks, as discussed in section 3.2. The pseudo-data contained in $\mathcal{D}_{L0}^{\text{PV19}}$ are listed in table 3.1, those in $\mathcal{D}_{L0}^{\text{MAP22}}$ can be found in tables 3.2 and 3.3, and the experimental points included in $\mathcal{D}_{L0}^{\text{MIX24}}$ are shown in tables 3.1 and 3.3. Table 4.5 summarizes the models and inputs used to generate the three L0 pseudo-datasets.

Since the three datasets are treated equivalently when generating the corresponding L1 and L2 datasets, we will simplify the notation by referring to the L0 datasets generically as $\mathcal{D}_{L0}$. Specific labels will be restored where needed for clarity.

Dataset	Model	Input	Dimension
$\mathcal{D}_{L0}^{\text{MAP22}}$	$\mathcal{G}_{\text{MAP22}}$	$u_{\text{MAP22}}^{\text{real}}$	$\mathbb{R}^{2031}$
$\mathcal{D}_{L0}^{\text{MIX24}}$	$\mathcal{G}_{\text{MAP22}}$	$u_{\text{MIX24}}^{\text{real}}$	$\mathbb{R}^{1900}$
$\mathcal{D}_{L0}^{\text{PV19}}$	$\mathcal{G}_{\text{PV19}}$	$u_{\text{PV19}}^{\text{real}}$	$\mathbb{R}^{353}$

Table 4.5: Models and inputs used to generate the three L0 pseudo-datasets $\mathcal{D}_{L0}^{\text{MAP22}}$, $\mathcal{D}_{L0}^{\text{PV19}}$, and $\mathcal{D}_{L0}^{\text{MIX24}}$.

4.7 Data Uncertainties

When generating a Level 0 dataset, we must select the uncertainties associated with the non-fluctuated points. Although these points have no intrinsic error, we assign them uncertainties identical to those in the original experimental dataset, $\mathcal{D}$. This approach ensures that the covariance matrix and correlations in the original data are preserved in the pseudo-datasets. The same uncertainties are applied when constructing the L1 and L2 datasets, meaning that the covariance matrix remains consistent across all levels,

$$C_{L0} = C_{L1} = C_{L2} = C, \quad (4.42)$$

where C is the covariance matrix of the original dataset, and the equality holds for each of the three datasets discussed in the previous section.

Uncertainty Structure and Covariance Matrix

Understanding the structure of the uncertainties and how the covariance matrix is constructed is essential, as this information will be used to generate the Level 1 and Level 2 fluctuations. Each point in the experimental dataset, $\mathcal{D}_i$ (where i is the index of the data-point) has a total of $N_{\text{unc},i} + N_{\text{add},i} + N_{\text{mult},i}$ associated uncertainties (see table 4.6). These uncertainties are divided as follows:

- **$N_{\text{unc},i}$ uncorrelated uncertainties:** these uncertainties are independent of the uncertainties of other data points, $\mathcal{D}_{j \neq i}$. The k -th uncorrelated uncertainty of the i -th point is denoted as $\sigma_i^{\text{unc},k}$, with $k = 1 \dots N_{\text{unc},i}$.
- **$N_{\text{add},i}$ correlated additive uncertainties:** these are correlated across different data points, $\mathcal{D}_{j \neq i}$. The ℓ -th additive uncertainty of the i -th point is represented as $\sigma_i^{\text{add},\ell}$, with $\ell = 1 \dots N_{\text{add},i}$.
- **$N_{\text{mult},i}$ correlated multiplicative uncertainties:** these are correlated across different data points. The m -th multiplicative uncertainty of the i -th point is denoted as $\sigma_i^{\text{mult},m}$, with $m = 1 \dots N_{\text{mult},i}$.

Uncorrelated uncertainties are expressed as absolute values, while both additive and multiplicative uncertainties are expressed as relative values. To compute the absolute uncertainty from a relative value, we multiply by the data point value, $\mathcal{D}_i$,

Uncertainty Type	Expression	Name	Index
Uncorrelated	Absolute	$\sigma_i^{\text{unc},k}$	$k = 1 \dots N_{\text{unc},i}$
Correlated, Additive	Relative	$\sigma_i^{\text{add},\ell}$	$\ell = 1 \dots N_{\text{add},i}$
Correlated, Multiplicative	Relative	$\sigma_i^{\text{mult},m}$	$m = 1 \dots N_{\text{mult},i}$

Table 4.6: Uncertainties associated to the i -th data point of the experimental dataset $\mathcal{D}$, these uncertainties are transferred to the i -th point of the L0, L1 and L2 pseudo-datasets.

$$\sigma_i^{\text{abs}} = \sigma_i^{\text{rel}} \mathcal{D}_i . \quad (4.43)$$

The dataset's covariance matrix is an $N_{\text{dat}} \times N_{\text{dat}}$ matrix, built from the uncertainties described in table 4.6 and the central values $\mathcal{D}_i$. The diagonal C_{ii} of the matrix is given by

$$C_{ii} = (\sigma_i^{\text{unc}})^2 + \sum_{\ell=1}^{N_{\text{add},i}} \left(\sigma_i^{\text{add},\ell} \mathcal{D}_i \right)^2 + \sum_{m=1}^{N_{\text{mult},i}} \left(\sigma_i^{\text{mult},m} \mathcal{D}_i \right)^2 , \quad (4.44)$$

where $\sigma_i^{\text{unc}} = \sqrt{\sum_{k=1}^{N_{\text{unc}}} (\sigma_i^{\text{unc},k})^2}$ is the total uncorrelated uncertainty for the i -th data point. The off diagonal elements, C_{ij} with $i \neq j$, are constructed using only the correlated uncertainties, as follows:

$$C_{ij} = \sum_{\ell=1}^{N_{\text{add},i}} \left(\sigma_i^{\text{add},\ell} \sigma_j^{\text{add},\ell} \mathcal{D}_i \mathcal{D}_j \right) + \sum_{m=1}^{N_{\text{mult},i}} \left(\sigma_i^{\text{mult},m} \sigma_j^{\text{mult},m} \mathcal{D}_i \mathcal{D}_j \right) . \quad (4.45)$$

The MAPTMD22 and PV19 frameworks use the Cholesky decomposition of the covariance matrix to fluctuate the experimental dataset $\mathcal{D}$ and produce Monte Carlo replicas. We will apply the same method to fluctuate the L0 datasets and generate the L1 pseudo-data.

4.8 Data Fluctuation

We conduct Level 1 and Level 2 closure tests only on the datasets generated with $u_{\text{PV19}}^{\text{real}}$ and $u_{\text{MAP22}}^{\text{real}}$. Therefore, the dataset $\mathcal{D}_{L0}^{\text{MIX24}}$ will not be fluctuated. To obtain a L1 dataset $\mathcal{D}_{L1}$ (*i.e.* a Monte Carlo replica of a L0 dataset), we start from the non-fluctuated dataset $\mathcal{D}_{L0}$ and add a fluctuation to each point, considering the data uncertainties outlined in table 4.6. In Ref [6], the i -th point of the L1 dataset, $\mathcal{D}_{L1,i}$, is calculated using the formula

$$\mathcal{D}_{L1,i} = \prod_{m=1}^{N_{\text{mult},i}} \sqrt{\left(1 + r_m^{\text{mult}} \sigma_i^{\text{mult},m} \right)} \mathcal{D}_{L0,i} \left(1 + \frac{r_i^{\text{unc}} \sigma_i^{\text{unc}}}{\mathcal{D}_{L0,i}} + \sum_{\ell=1}^{N_{\text{add},i}} r_\ell^{\text{add}} \sigma_i^{\text{add},\ell} \right) , \quad (4.46)$$

where the correlated uncertainties are considered relative. The values r_m^{mult} , r_i^{unc} , and r_ℓ^{add} are random weights drawn from a normal distribution $\mathcal{N}(0, 1)$. The weights associated to the additive and multiplicative uncertainties do not have the label i , since this are correlated uncertainties and cause the same shift on each point of the dataset. Equation (4.46) was used in earlier versions of the MAP fitting frameworks to generate Monte Carlo replicas of experimental datasets. It's included here to illustrate how fluctuations depend on different types of uncertainties. However, in this analysis, we employ a different method for fluctuating the L0 dataset. Using the covariance matrix C , we apply the Cholesky decomposition technique, in which C is decomposed into the product of a lower triangular matrix L and its transpose L^T ,

$$C = LL^T, \quad (4.47)$$

which can be performed since the covariance matrix is symmetrical and squared for construction. We then use the lower triangular matrix, L , to generate a fluctuation vector $\eta \in \mathbb{R}^{N_{\text{dat}}}$

$$\eta = rL, \quad (4.48)$$

where $r \in \mathbb{R}^{N_{\text{dat}}}$ is a vector composed of normally distributed random weights $r_i = \mathcal{N}(0, 1)$. The L1 dataset is obtained by adding the fluctuation vector to the L0 dataset,

$$\mathcal{D}_{L1} = \mathcal{D}_{L0} + \eta. \quad (4.49)$$

The same method is then used to fluctuate the L1 dataset, obtaining the k -th replica with a Level 2 fluctuation is

$$\mathcal{D}_{L2}^{(k)} = \mathcal{D}_{L1} + \epsilon^{(k)}, \quad (4.50)$$

where $\epsilon^{(k)} \in \mathbb{R}^{N_{\text{dat}}}$ is the vector of fluctuations generated starting from the vector of random values $p^{(k)} \in \mathbb{R}^{N_{\text{dat}}}$. Each normally distributed random vector $p^{(k)}$ is generated with a different seed and leads to a different Monte Carlo replica of the L1 dataset. Fluctuating a dataset using the decomposed covariance matrix is mathematically equivalent to the method described by equation (4.46).

4.9 Numerical Setup of the Tests

In the previous section, we explained how the L0, L1, and L2 datasets are generated. We begin by generating three L0 datasets: $\mathcal{D}_{L0}^{\text{MAP22}}$, $\mathcal{D}_{L0}^{\text{PV19}}$, and $\mathcal{D}_{L0}^{\text{MIX24}}$. We then fluctuate the first two to create the L1 datasets, $\mathcal{D}_{L1}^{\text{MAP22}}$ and $\mathcal{D}_{L1}^{\text{PV19}}$. Finally, we fluctuate these L1 datasets to produce N_{rep} Monte Carlo replicas (L2 datasets), resulting in $\mathcal{D}_{L2}^{\text{MAP22}}$ and $\mathcal{D}_{L2}^{\text{PV19}}$. In this section, we will describe the setup of the closure tests on the MAPTMD22 and PV19 fitting frameworks.

L0 Tests Setup

We performed tests on both the MAPTMD22 and PV19 frameworks using L0 pseudo-data generated from the three input parameterizations $u_{\text{MAP22}}^{\text{real}}$, $u_{\text{MIX24}}^{\text{real}}$, and $u_{\text{PV19}}^{\text{real}}$. In

Test Name	Fitting Framework	Fitted Dataset	Minimizer	N_{rep}
MAP22oMAP22 _{L0}	MAPTMD22	$\mathcal{D}_{L0}^{\text{MAP22}}$	Ceres	250
MAP22oMAP22 _{L0} ^{Ceres}	MAPTMD22	$\mathcal{D}_{L0}^{\text{MAP22}}$	Ceres	100
MAP22oMAP22 _{L0} ^{Minuit}	MAPTMD22	$\mathcal{D}_{L0}^{\text{MAP22}}$	Minuit	100
PV19oPV19 _{L0} ^{Ceres}	PV19	$\mathcal{D}_{L0}^{\text{PV19}}$	Ceres	100
PV19oPV19 _{L0}	PV19	$\mathcal{D}_{L0}^{\text{PV19}}$	Ceres	250
PV19oPV19 _{L0} ^{Minuit}	PV19	$\mathcal{D}_{L0}^{\text{PV19}}$	Minuit	100
MAP22oPV19 _{L0}	MAPTMD22	$\mathcal{D}_{L0}^{\text{PV19}}$	Ceres	250
PV19oMAP22 _{L0}	PV19	$\mathcal{D}_{L0}^{\text{MAP22}}$	Ceres	250
MAP22oMIX24 _{L0}	MAPTMD22	$\mathcal{D}_{L0}^{\text{MIX24}}$	Ceres	250
PV19oMIX24 _{L0}	PV19	$\mathcal{D}_{L0}^{\text{MIX24}}$	Ceres	250

Table 4.7: Table including all the performed L0 closure tests.

total, we conducted six L0 tests, naming each test with a label that indicates the fitting framework first (MAP22 or PV19) followed by the input parametrization (MAP22, PV19, or MIX24). For example, the label "MAP22oPV19_{L0}" (where "o" stands for "over") refers to a Level 0 test executed using the MAPTMD22 fitting framework over the dataset generated with the PV19 parametrization of the non-perturbative TMDs. Each Level 0 test consists of $N_{\text{rep}} = 250$ identical replicas⁴ of the dataset $\mathcal{D}_{L0}$, but the optimization algorithm starts from a different point in the parameter space for each replica.

The purpose of Level 0 tests is to evaluate the flexibility of the frameworks. Additionally, we conducted two tests where the input and fitting parameterizations remained the same (MAP22oMAP22_{L0} and PV19oPV19_{L0}) but varied the minimizer (the algorithm used to find the minimum χ^2 in parameter space) to assess the effect of different optimization algorithms on the frameworks. The two tested minimizers were Ceres [44] and Minuit [45], and we used the same initializations for the two. The tests conducted for the Ceres-Minuit comparison involved only 100 initializations. A list including all the tests we made with the L0 datasets is shown in table 4.7.

Uncertainty Characterization

For uncertainty characterization, we focused on testing the two frameworks, MAPTMD22 and PV19, with datasets generated from their respective parameterizations, $\mathcal{D}^{\text{MAP22}}$ and $\mathcal{D}^{\text{PV19}}$. As this phase of the closure test aims to quantify the uncertainties rather than the flexibility of the frameworks, we did not perform cross-tests (*i.e.*

⁴Here, we refer to the number of *identical* replica as N_{rep} instead of N_{init} , because the two values coincide in the tests we conducted.

Test Name	Fitting Framework	Fitted Dataset	N_{rep}
MAP22oMAP22 $_{L0}$	MAPTMD22	$\mathcal{D}_{L0}^{\text{MAP22}}$	250
MAP22oMAP22 $_{L1}$	MAPTMD22	$\mathcal{D}_{L1}^{\text{MAP22}}$	250
MAP22oMAP22 $_{L2}$	MAPTMD22	$\mathcal{D}_{L2}^{\text{MAP22}}$	250
PV19oPV19 $_{L0}$	PV19	$\mathcal{D}_{L0}^{\text{PV19}}$	250
PV19oPV19 $_{L1}$	PV19	$\mathcal{D}_{L1}^{\text{PV19}}$	250
PV19oPV19 $_{L2}$	PV19	$\mathcal{D}_{L2}^{\text{PV19}}$	250

Table 4.8: Table including all the performed L0, L1, and L2 closure tests for the uncertainty characterization of MAPTMD22 and PV19.

Test Name	Fitting Framework	Fitted Dataset	N_{fit}	N_{rep}
MAP22oMAP22 $_{\text{multi}}$	MAPTMD22	$\mathcal{D}_{L2}^{\text{MAP22}}$	50	100
PV19oPV19 $_{\text{multi}}$	PV19	$\mathcal{D}_{L2}^{\text{PV19}}$	50	100

Table 4.9: Table including the performed multi closure tests. We performed each test on a total of 5000 L2 dataset (100 L2 fluctuations starting from 50 different L1 datasets).

MAP22oPV19 and PV19oMAP22). We evaluated the relative experimental, functional, and interpolation/extrapolation uncertainties by comparing the TMD functional spreads obtained from L0, L1, and L2 tests. For the L0 and L1 tests, we used $N_{\text{rep}} = 250$ identical replicas of the L0 and L1 datasets, respectively. For the L2 tests, we performed the closure tests using $N_{\text{rep}} = 250$ replicas of the L1 dataset. In these tests, performing an L2 closure is equivalent to performing a fit with the L1 pseudo-dataset instead of the real experimental dataset. Table 4.8 lists the numerical setup for the uncertainty characterization tests, where Ceres was used as the minimizer for all tests.

Multi Closure Tests Setup

For the multi closure tests, we only used datasets where the same parametrization was applied for both the generation and fitting of the pseudo-datasets. The closure estimators defined in equations (4.28) and (4.35) were calculated for both PV19oPV19 and MAP22oMAP22, using a total of $N_{\text{fit}} = 50$ fits. Each fit was performed with $N_{\text{rep}} = 100$ Monte Carlo replicas of the L1 dataset, for a total of 5000 L2 datasets. The number of fits corresponds to the number of distinct L1 datasets, while the number of replicas represents the number of L2 fluctuations generated for each L1 dataset. Table 4.9 lists the setup of the multi closure tests.

In this chapter, we have outlined the setup for various closure tests, *i.e.* L0, L1, L2, and multi closure tests, across the MAPTMD22 and PV19 frameworks. In the

next chapter, we will present and analyze the results of these closure tests for both frameworks.

Chapter 5

Results

In this chapter, we present the results of various closure tests performed on the MAPTMD22 and PV19 fitting procedures. As discussed in Chapter 4, closure testing involves different levels, each designed to test various features of a fitting framework:

1. **framework’s flexibility (Level 0)**: this is tested by comparing the TMD functional forms predicted by the methodology with those used to generate the Level 0 pseudo-dataset;
2. **uncertainty levels**: the contributions of interpolation/extrapolation, functional, and experimental uncertainties to the total TMD uncertainties produced by a fitting procedure are assessed by comparing the results of the fits on the Level 0, Level 1, and Level 2 datasets;
3. **faithfulness of total TMD uncertainties**: this is tested calculating the quantile estimator (equation (4.35)) through a multi closure test;
4. **framework complexity tradeoff**: the bias-variance ratio estimator (equation (4.28)) is calculated through a multi-closure test to assess the tradeoff in framework complexity.

We will begin by presenting the results of the Level 0 closure tests (pseudo-data without fluctuations).

5.1 Choice of the Minimizer

Using the Level 0 datasets $\mathcal{D}_{L0}^{\text{MAP22}}$ and $\mathcal{D}_{L0}^{\text{PV19}}$, we first tested the minimizers employed by the two frameworks: Ceres [44] and Minuit [45]. For this test, we performed 100 fits using the MAPTMD22 framework on the pseudo-data generated by the MAPTMD22 parametrization (see table 4.7), the test is named MAP22oMAP22_{L0}. We also conducted 100 fits using the PV19 framework on the Level 0 pseudo-data generated from the PV19 parametrization, this test is labeled PV19oPV19_{L0}. For each of these 100 fits, we initiated the minimization from a different starting point in the parameter space, the same 100 points were initialized for Ceres and Minuit. We conducted these

Table 5.1: Ceres-Minuit comparison.

Test	Minimiser	χ^2
MAP22oMAP22 _{L0}	Ceres	$\mathcal{O}(10^{-14})$
MAP22oMAP22 _{L0}	Minuit	$\mathcal{O}(10^{-4})$
PV19oPV19 _{L0}	Ceres	$\mathcal{O}(10^{-13})$
PV19oPV19 _{L0}	Minuit	$\mathcal{O}(10^{-5})$
MAP22oPV19 _{L0}	Ceres	$\mathcal{O}(10^{-1})$
MAP22oPV19 _{L0}	Minuit	not convergent

tests with both Ceres and Minuit to determine whether the minimizer could consistently find the correct minimum from a random starting point, or if the identified minimum was dependent on the initial conditions. In both tests we used the final χ^2 as a metric for the accuracy of the fit. In the MAP22oMAP22_{L0} test, we obtain a very good result when we use Ceres as a minimizer: in the majority of the cases the minimizer stops at $\chi^2 \sim \mathcal{O}(10^{-14})$, with only around 10% of the cases failing to reach this level of accuracy. In contrast, when using Minuit as the minimizer, the best results yielded $\chi^2 \sim \mathcal{O}(10^{-4})$. We observed similar results when testing both minimizers with PV19oPV19_{L0}. Additionally, we conducted a test using the MAPTMD22 parametrization on a dataset generated from PV19 (MAP22oPV19_{L0}), again employing both Ceres and Minuit. The results of the Ceres-Minuit comparison are presented in table 5.1. From these results, it is clear that Ceres outperforms Minuit. Consequently, we will proceed with the closure tests using only Ceres as the framework’s minimizer.

5.2 L0 Results

As shown in table 5.1, the closure test of the MAPTMD22 parametrization on a dataset generated by the same parametrization results in a $\chi^2 \sim \mathcal{O}(10^{-14})$. Similarly, the PV19oPV19_{L0} test yields $\chi^2 \sim \mathcal{O}(10^{-13})$. We repeated these tests using 250 replicas of the L0 datasets. Using Ceres, both MAP22oMAP22_{L0} and PV19oPV19_{L0} satisfy the condition defined in equation (4.12). In particular, by computing the mean χ^2 across the 250 replicas, we obtain:

$$\langle \chi^2[\mathcal{D}_{\text{pred}}^{\text{MAP22}}, \mathcal{D}_{L0}^{\text{MAP22}}] \rangle \sim \mathcal{O}(10^{-14}) \quad (5.1)$$

$$\langle \chi^2[\mathcal{D}_{\text{pred}}^{\text{PV19}}, \mathcal{D}_{L0}^{\text{PV19}}] \rangle \sim \mathcal{O}(10^{-13}) . \quad (5.2)$$

In both cases, the two frameworks successfully reproduce the perfect solution that exactly follows the distribution of the L0 data points, and we conclude that the L0 test closes for the two frameworks.

Since both tests involve a certain parametrization (MAPTMD22 or PV19) capturing its own input, we are in a situation where we are *assuming* that our parametrization exactly matches the real parametrization of the TMDs, requiring only the free parameters to be fixed. This initial result demonstrates that, under the assumption that our

parametrization is identical to the real law's parametrization, both fitting procedures are well constructed and capable of reconstructing the real set of parameters, u_{real} , with high precision.

However, a more likely scenario is that we do not know the exact parameterizations of the real TMD PDFs and FFs. Suppose the parameterizations used in our frameworks are approximations of the real law; they may share a similar functional form but are not identical. In this case, can the MAPTMD22 and PV19 fitting methodologies still identify a functional form that matches sufficiently well the real TMDs? This situation is simulated using the cross-tests MAP22oPV19_{L0}, MAP22oMIX24_{L0}, PV19oMAP22_{L0}, and PV19oMIX24_{L0}, where the input u_{real} used to generate the pseudo-data differs from the input used to fit the pseudo-data. In the next section, we will present the results of the cross L0 tests.

MAP22oPV19_{L0} Test Results

In this case, we have a dataset generated using the PV19 parameterization (equation (3.16)), denoted as $\mathcal{D}_{L0}^{\text{PV19}}$. We fit this dataset using the MAPTMD22 framework, which employs the TMD PDF parametrization described in equation (3.12). Since the dataset $\mathcal{D}_{L0}^{\text{PV19}}$ does not contain SIDIS data, the TMD FF parametrized in MAPTMD22 will not be included in the model. Consequently, we expect the minimizer to produce a random set of parameters for the non-perturbative TMD FF, which we will not consider in this test. In this scenario, our parametrization is slightly more complex than that used to generate the pseudo-data.

We conducted 250 fits of the L0 dataset, starting from different points in the parameter space each time. From this fitting procedure, we obtained 220 convergent fits, with a mean χ^2 of

$$\langle \chi^2[\mathcal{D}_{\text{pred}}^{\text{MAP22}}, \mathcal{D}_{L0}^{\text{PV19}}] \rangle \pm \sigma_{\chi^2} = 0.152 \pm 0.093 . \quad (5.3)$$

We conclude that this test does not close, as the condition $\chi^2 < 10^{-10}$ is not met. However, when plotting the real and predicted TMD functional forms $f_{NP}(x, b_T, \zeta)$ —the first generated with the input $u_{\text{real}}^{\text{PV19}}$ and the second with the predicted input u_{*}^{MAP22} —we observe two classes of predicted functions. Figure 5.1(a) shows, in red, the TMD PDF used to generate the fitted dataset, and in blue, the N_{rep} TMD PDFs predicted by the L0 fits. To quantify the similarity between the i -th prediction, which we denote as $f_{NP,i}^{\text{pred}}$, and the real functional form f_{NP}^{real} , we define the Mean Squared Distance (MSD) estimator. The MSD estimator (equation (5.4)) measures the mean squared distance between $f_{NP,i}^{\text{pred}}$ and f_{NP}^{real} over a large number of points in b_T ,

$$\text{MSD} = \frac{1}{N_{\text{points}}} \sum_{p=1}^{N_{\text{points}}} (f_{NP}^{\text{real}} - f_{NP,i}^{\text{pred}})_p^2 , \quad (5.4)$$

where the real and predicted $f_{NP}(x, b_T, \zeta)$ are plotted at $x = 0.1$, $\mu = Q = \sqrt{\zeta} = 10$ GeV, and for $b_T \in [0, 10]$. The number of points used to generate the distributions

within this b_T interval is denoted as N_{points} , with $N_{\text{points}} = 100$ in this case. We can utilize the Mean Squared Distance (MSD) as a metric to assess how closely the predicted TMD aligns with the real TMD we are targeting. We implement a condition to differentiate the two classes of functional forms that can be observed in figure 5.1(a). Specifically, we establish the condition:

$$\text{MSD} \leq 0.0015 . \quad (5.5)$$

Only 109 fits satisfy this condition. The functional forms that meet equation (5.5) are illustrated in figure 5.1(b).

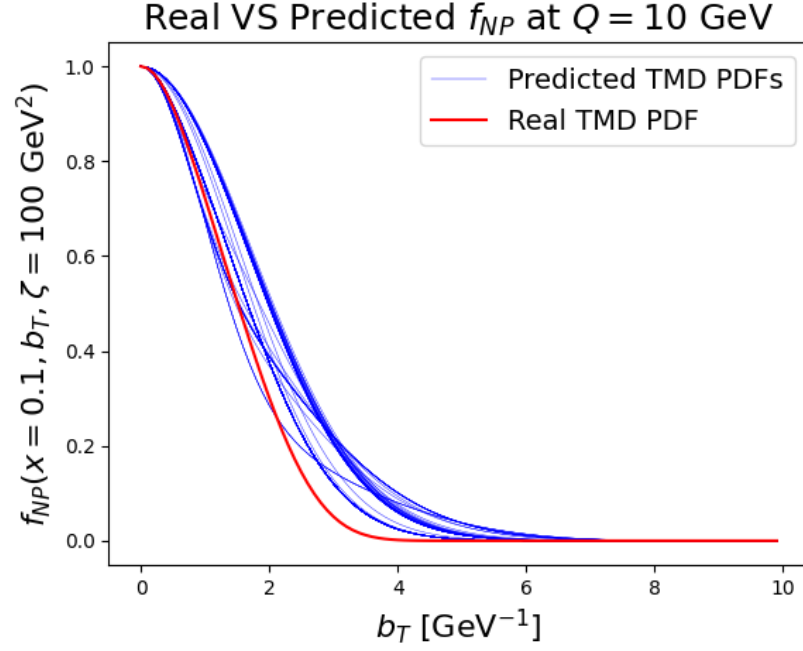
We also plotted the fitted TMD PDF functional forms at $Q = 2$ GeV, to assess whether the result is dependent on the region of Q . We obtained a very similar result compared to the case of $Q = 10$ GeV, with the fitted functional forms splitting in two classes, as shown in figure 5.2, with around 50% of the replicas satisfying the condition $\text{MSD} \leq 0.0015$. The minimizer is still not able to identify the best non-perturbative TMD in all the cases.

We conclude that in the MAP22oPV19 $_{L0}$ test, the optimization algorithm identifies a minimum that is dependent on the initial conditions in the parameter space. Two distinct minima emerge: one corresponds to the real distribution, while the other results in an inaccurate prediction of the TMD PDF. The minimizer employed (Ceres) successfully finds the correct minimum in approximately half of the cases. This leads to a mean χ^2 of 0.146, representing an average over a set of acceptable results (with χ^2 around 0.05) and a set of incorrect results (with χ^2 around 0.25). As shown in figure 5.3, the χ^2 distribution clearly divides into two classes. If we focus only on the best fits (the acceptable results) and neglect the replicas that converge but yield a $\chi^2 > 0.07$, the outcome of the closure test improves. However, even among the best fits, the χ^2 for MAP22oPV19 $_{L0}$ does not satisfy the condition defined in equation (4.12), leading us to conclude that the test does not close.

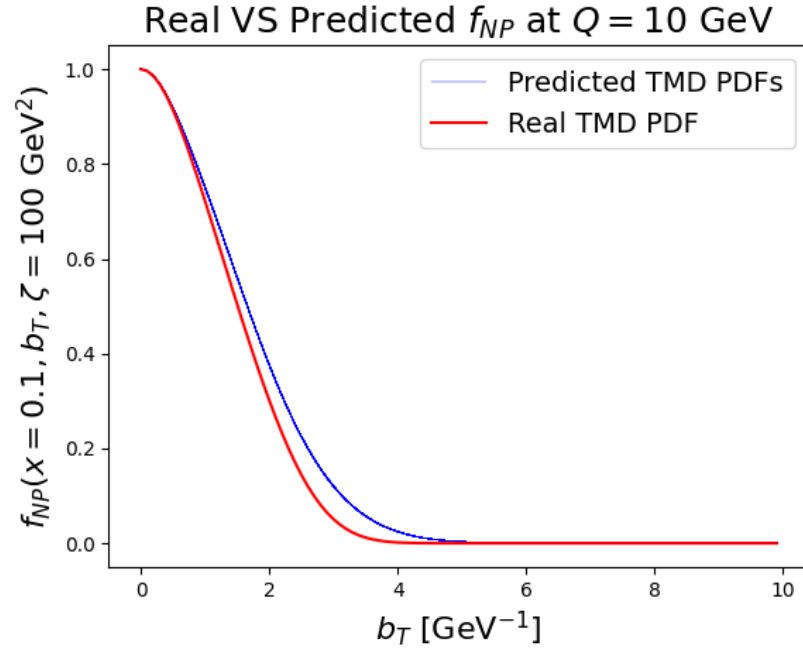
This result indicates that, under the assumption that the input parametrization used for the fit differs from the real parametrization, the MAPTMD22 framework lacks the flexibility necessary to accurately interpolate all the L0 data points and find the perfect solution, which corresponds to the $\chi^2 = 0$ solution.

PV19oMAP22 $_{L0}$ Test Results

We now make the opposite test. In this test, referred to as PV19oMAP22 $_{L0}$, we perform 250 fits using the PV19 parametrization over a dataset generated from the MAPTMD22 parametrization, denoted as $\mathcal{D}_{L0}^{\text{MAP22}}$. For each fit, we initiate the optimization from a different starting point in the parameter space. Similar to the previous test, we only consider Drell-Yan experiments, as the PV19 parametrization does not account for the TMD fragmentation functions (FFs) relevant for SIDIS data. Therefore, the dataset is limited to the points contained in the experimental dataset $\mathcal{D}_{\text{PV19}}$, but generated using the input $u_{\text{real}}^{\text{MAP22}}$. We exclude Drell-Yan data not included in Ref. [3] to replicate the



(a) Complete set of predicted non-perturbative TMD PDFs at $x = 0.1$ and $Q = 10$ GeV.



(b) Set of non-perturbative TMD PDFs at $x = 0.1$ and $Q = 10$ GeV which satisfy $MSD < 0.0015$.

Figure 5.1: Real and predicted $f_{NP}(x = 0.1, b_T, \zeta = 100 \text{ GeV}^2)$ in MAP22oPV19_{L0}.

conditions utilized in the PV19 framework.

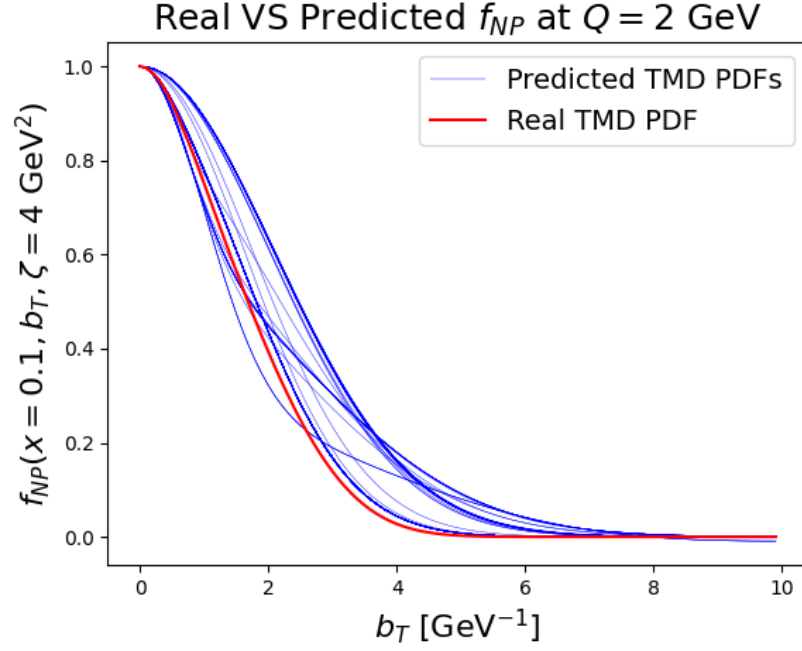


Figure 5.2: Real and predicted $f_{NP}(x = 0.1, b_T, \zeta = 4 \text{ GeV}^2)$ in MAP22oPV19_{L0}.

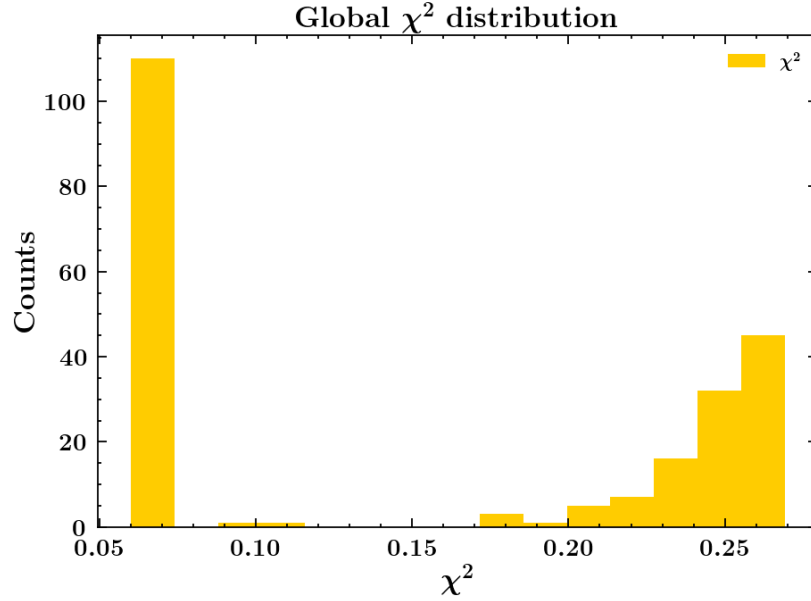
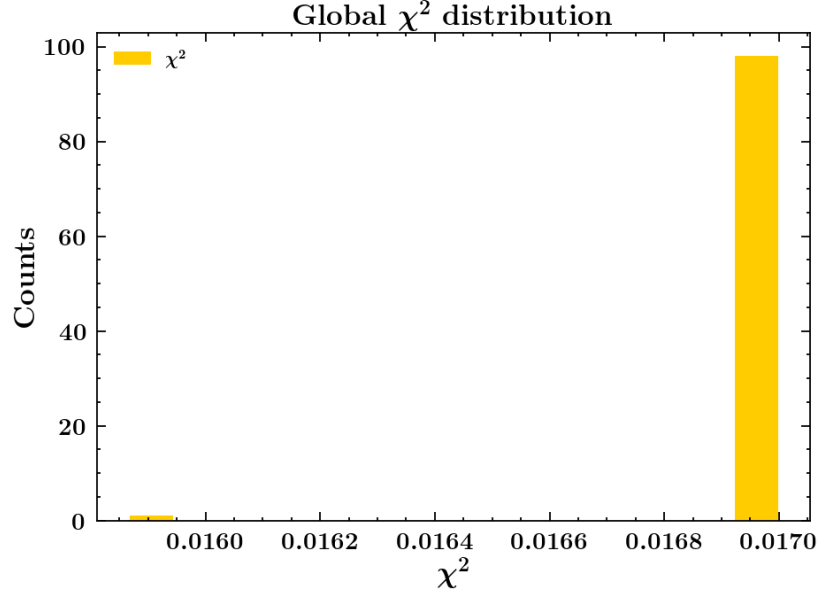


Figure 5.3: χ^2 distribution in MAP22oPV19_{L0}.

The results from this test are better than those from MAP22oPV19_{L0}. The minimizer converged to the same minimum in 100% of the replicas, yielding a mean χ^2 of

$$\langle \chi^2[\mathcal{D}_{\text{pred}}^{\text{PV19}}, \mathcal{D}_{L0}^{\text{MAP22}}] \rangle \pm \sigma_{\chi^2} = 0.017 \pm 0.0001, \quad (5.6)$$

which is an order of magnitude smaller than the mean χ^2 obtained in the previous test. The χ^2 distribution shown in figure 5.4 and the standard deviation from equation (5.6)


 Figure 5.4: χ^2 distribution in the PV19oMAP22 $_{L0}$ test.

indicate that, in this case, the minimizer reliably identifies the correct minimum, independent of the starting point in the parameter space.

Additionally, we assess how closely the i -th predicted functional form, $f_{NP,i}^{\text{pred}}(x, b_T, \zeta)$, aligns with the real non-perturbative function, $f_{NP}^{\text{real}}(x, b_T, \zeta)$, through the MSD estimator (eq 5.4). We find that the previous condition $\text{MSD} \leq 0.0015$ is satisfied in 100% of the fits. To further refine our analysis, we reduce the MSD limit by a factor of 10, leading to a new criterion:

$$\text{MSD} \leq 0.00015 . \quad (5.7)$$

Remarkably, this stricter criterion is also met in 100% of the replicas for the PV19oMAP22 $_{L0}$ test. The distribution of the predicted TMD PDFs at $x = 0.1$ and $Q = 10$ GeV is illustrated in blue in figure 5.5. We observe that all predictions follow a similar functional form but do not perfectly capture the real functional form used to generate the pseudo-data (which is the red curve). A similar result is obtained plotting the real and predicted functional forms at $Q = 2$ GeV.

In conclusion, the results indicate that within the PV19 framework, the minimizer successfully identifies the global minimum of χ^2 in the parameter space, independent of the starting conditions for the optimization process. However, the PV19 parametrization proves insufficiently flexible to accurately represent the real TMD functional form, assuming that the parametrization of the real TMD PDF is different from the PV19 parametrization.

Comparing the MAPTMD22 and PV19 frameworks, we conclude that PV19 is more accurate in capturing the true functional form of the non-perturbative TMD PDF, despite its simpler parametrization. We identified two possible causes that may lead to incorrect predictions of the TMD PDF functional forms in the cross tests:

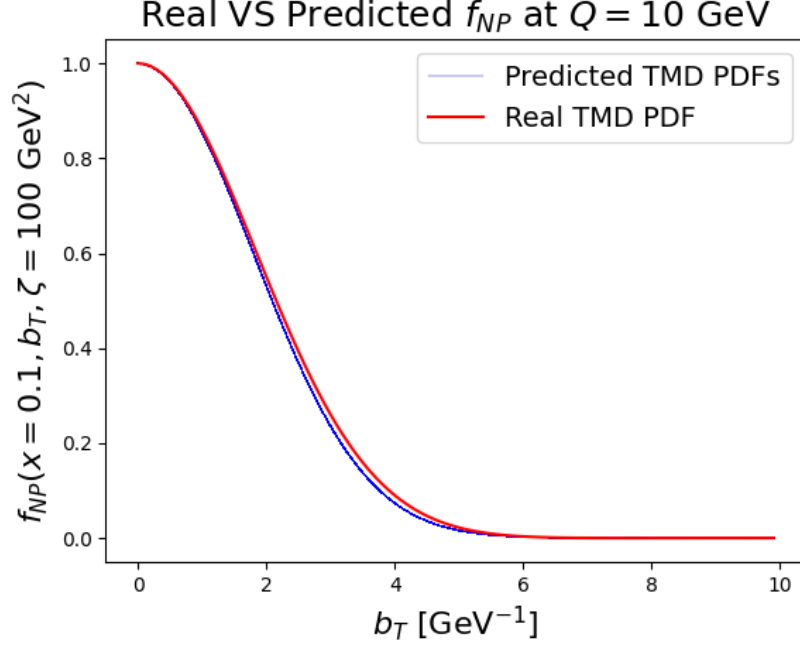


Figure 5.5: Real and predicted $f_{NP}(x = 0.1, b_T, \zeta = 100 \text{ GeV}^2)$ in the PV19oMAP22_{L0} test.

- Interplay between the minimizer and the parametrization:** when Ceres, the minimizer used in both frameworks, searches for the minimum χ^2 , the interaction between the parametrization and the optimization process is crucial for finding the correct minimum. In the PV19 framework, this interplay works well, allowing the procedure to find the global minimum in 100% of the fits. However, in the MAPTMD22 framework, the optimization algorithm implemented in Ceres struggles to distinguish between a local minimum and the global minimum of the χ^2 , finding the parameter set that minimizes χ^2 in only half of the cases. In other words, the task of the minimizer is to find the best possible input u_* to reconstruct the dataset $\mathcal{D}$ given the model $\mathcal{G}$, but whether the best input is found in 100% of the fits depends on the interaction between the optimization process and the parametrization. The minimizer's efficiency depends on the parametrization it is applied to, and this limitation prevents it from exploring all possible solutions for the MAPTMD22 parametrization, missing the best solution in some cases.
- Parametrization flexibility:** the second issue, which is common to both frameworks, is that the best possible input u_* , which describes the data distribution $\mathcal{D} = \mathcal{G}(u_*)$, is not sufficient to capture the true functional form. This reflects a limitation in the parametrization, which lacks the necessary flexibility to accurately describe the real functional form we are trying to capture. As a result, the best approximation of the real input does not coincide with the actual input, as it should in an L0 test.

Another reason why the MAPTMD22 parametrization performs worse in describing

the PV19 dataset, compared to the reverse case, could be that we have not included the SIDIS dataset in the MAP22oPV19_{L0} test. The MAPTMD22 framework is designed to perform a global fit on both Drell-Yan and SIDIS cross sections. Its non-perturbative TMD PDF parametrization has been constructed and optimized not only to fit Drell-Yan data, but also to work in relation with the TMD FF to properly fit the SIDIS part of the dataset. In our test, however, we simplified the information by removing the SIDIS portion. During the optimization process, the minimizer deals with 21 free parameters, of which only 10 are relevant for constructing the TMD PDF, while the remaining 11, used for constructing the TMD FF, are irrelevant when fitting a Drell-Yan dataset. The minimizer might perform better in finding the global minimum if we include the SIDIS part of the dataset in our test. To test this hypothesis, we generated the MIX24 pseudo-dataset and performed the MAP22oMIX24_{L0} test.

MAP22oMIX24_{L0} Test Results

The $\mathcal{D}_{L0}^{\text{MIX24}}$ dataset is an extension of the $\mathcal{D}_{L0}^{\text{PV19}}$ dataset, generated using the non-perturbative TMD PDF from PV19 and the non-perturbative TMD FF from MAPTMD22. However, the numerical set of parameters used as input is different from the ones we used to generate the $\mathcal{D}_{L0}^{\text{PV19}}$ and $\mathcal{D}_{L0}^{\text{MAP22}}$ datasets (see table 4.4). As in the previous test, we fitted 250 identical replicas, each starting from different points in the parameter space. The fit successfully converged in 217 replicas, resulting in a mean χ^2 across the replicas of

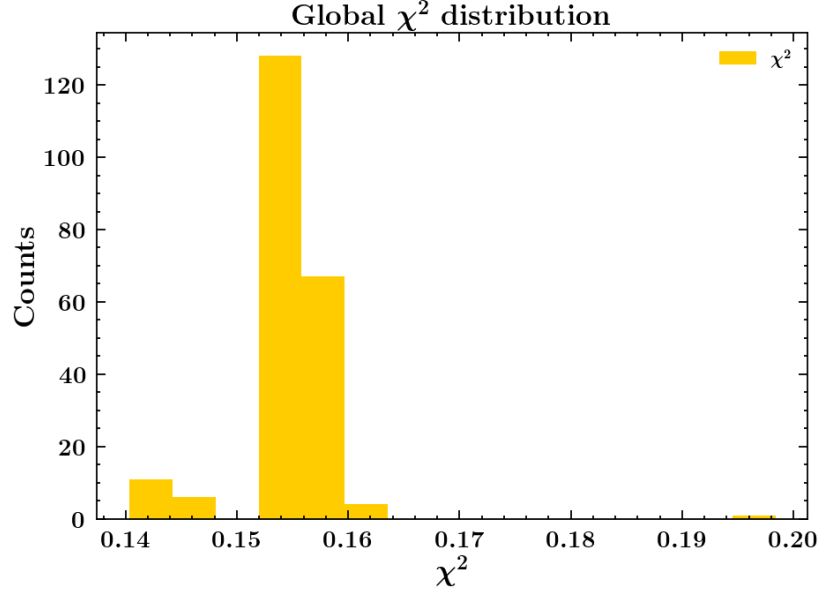
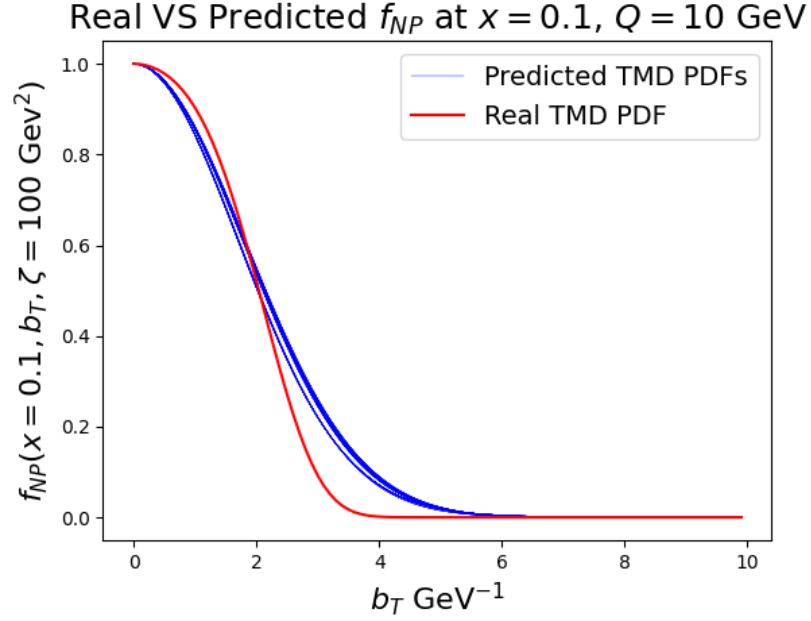
$$\langle \chi^2[\mathcal{D}_{\text{pred}}^{\text{MAP22}}, \mathcal{D}_{L0}^{\text{MIX24}}] \rangle \pm \sigma_{\chi^2} = 0.155 \pm 0.005 . \quad (5.8)$$

Comparing the standard deviations of the mean χ^2 obtained from the MAP22oMIX24_{L0} and MAP22oPV19_{L0} tests (equations (5.8) and (5.3)), we observe that the minimizer performs better when the complete dataset, including the SIDIS points, is considered. However, the χ^2 distribution, shown in figure 5.6, remains broad and split into two categories, with no significant improvement in the mean χ^2 , which stays at a similar level in the two tests.

Figure 5.7 illustrates the split into two classes of predicted functional forms for the non-perturbative TMD PDFs. We conclude that adding the SIDIS portion of the dataset does not resolve the two issues observed in the MAP22oPV19_{L0} test: the predicted set of parameters still depends on the initial starting point of the optimization process, and the MAP22 parametrization of the TMD PDF remains insufficiently flexible to accurately capture the assumed real TMD functional form.

PV19oMIX24_{L0} Test Results

Since the Drell-Yan part of the $\mathcal{D}_{L0}^{\text{MIX24}}$ dataset is generated using a different functional form compared to the $\mathcal{D}_{L0}^{\text{PV19}}$ dataset (*i.e.*, the parametrization is the same, but the parameter values are different, see table 4.4), we test whether the PV19 parametrization can capture this new functional form. This test is entirely analogous to the PV19oPV19_{L0} test, where the dataset is generated with a new set of parameters. We could also name the dataset for this test $\mathcal{D}_{L0}^{\text{PV19}}$. We performed the test on 250 identical


 Figure 5.6: χ^2 distribution in the MAP22oMIX24_{L0} test.

 Figure 5.7: Real and predicted $f_{NP}(x = 0.1, b_T, \zeta = 100 \text{ GeV}^2)$ in the MAP22oMIX24_{L0} test.

replicas, achieving convergence in 100% of the fits, with a mean χ^2 of:

$$\langle \chi^2[\mathcal{D}_{\text{pred}}^{\text{PV19}}, \mathcal{D}_{L0}^{\text{MIX24}}] \rangle \sim \mathcal{O}(10^{-7}) , \quad (5.9)$$

which demonstrates that the PV19 parametrization is flexible enough to capture the TMD PDF functional form in 100% of the fits. Although the χ^2 does not meet the closure condition, the result approaches the expected χ^2 of zero, which aligns with the

Test	Mean χ^2	Best MSD	Closure
MAP22oMAP22 _{L0}	$\mathcal{O}(10^{-14})$	/	yes
PV19oPV19 _{L0}	$\mathcal{O}(10^{-13})$	/	yes
MAP22oPV19 _{L0}	0.152	< 0.0015	no
PV19oMAP22 _{L0}	0.017	< 0.00015	no
MAP22oMIX24 _{L0}	0.155	< 0.004	no
PV19oMIX24 _{L0}	$\mathcal{O}(10^{-7})$	/	no

Table 5.2: Results of the tests performed on the L0 datasets.

outcome expected from an L0 closure test.

Table 5.2 summarizes the results obtained from the tests on the Level 0 datasets. We now proceed with the Level 1 and Level 2 tests only for the combinations where the L0 test successfully closes, *i.e.*, MAP22oMAP22 and PV19oPV19.

5.3 Uncertainty Characterization of MAPTMD22 and PV19

In this section, we present the results of the uncertainty characterization for the MAPTMD22 and PV19 frameworks. As discussed in section 4.4, the total TMD uncertainty—calculated as the standard deviation of the predicted non-perturbative TMDs across the L2 replicas (equation (3.29))—is composed of three sources of uncertainty:

- the interpolation/extrapolation uncertainty σ_{int} , which corresponds to the standard deviation of the TMDs predicted by the L0 replicas:

$$\sigma_{L0} = \sigma_{\text{int}} ; \quad (5.10)$$

- the functional uncertainty σ_{func} , which constitutes, together with the interpolation/extrapolation uncertainty, the standard deviation of the TMDs predicted by the L1 replicas:

$$\sigma_{L1} = \sigma_{\text{int}} + \sigma_{\text{func}} ; \quad (5.11)$$

- the experimental uncertainty σ_{exp} , which contributes only to the standard deviation of the TMDs predicted by the L2 replicas:

$$\sigma_{L2} = \sigma_{\text{int}} + \sigma_{\text{func}} + \sigma_{\text{exp}} . \quad (5.12)$$

Since the uncertainties described above depend on the set of variables variables (x, b_T, ζ) for the TMD PDFs uncertainties, and (z, b_T, ζ) for the TMD FFs, we do not report them as specific numbers. Instead, we plot the TMDs over the range $b_T \in [0, 10] \text{ GeV}^{-1}$, along with their L0, L1, and L2 uncertainties, for fixed values of $\mu = Q = \sqrt{\zeta}$, x and z . We then provide some qualitative interpretations of these results.

non-perturbative function	Q	x, z	Figure
$f_{NP}(x, b_T, \zeta)$	10 GeV	0.1	figure 5.8
		0.01	figure 5.9
		0.001	figure 5.10
$D_{NP}(z, b_T, \zeta)$	10 GeV	0.3	figure 5.11
		0.6	figure 5.12
$f_{NP}(x, b_T, \zeta)$	2 GeV	0.1	figure 5.13
		0.01	figure 5.14
		0.001	figure 5.15
$D_{NP}(z, b_T, \zeta)$	2 GeV	0.3	figure 5.16
		0.6	figure 5.17

Table 5.3: Plots of the uncertainties of the TMD PDFs and FFs predicted by MAPTMD22 in different kinematic regions.

MAP22oMAP22 Uncertainties

For the MAP22oMAP22 test, we plot the L0, L1, and L2 uncertainties of the non-perturbative TMDs at $Q = 10$ GeV and $Q = 2$ GeV. For both values of Q , we plotted the TMD PDF uncertainties at $x = 0.1$, $x = 0.01$, and $x = 0.001$, and the TMD FF uncertainties at $z = 0.3$ and $z = 0.6$. These kinematic points match those used in Ref. [4] to display the complete TMDs (including the perturbative contributions) obtained by the MAPTMD22 fitting framework with their uncertainties, where the TMD PDFs are represented as functions of $|k_T|$ and the TMD FFs in function of $|P_\perp|$. The plots presented in this work represent the non-perturbative part of the TMD PDFs ($f_{NP,\text{pred}}$) and FFs ($D_{NP,\text{pred}}$) as functions of b_T . These are computed using equations (3.27) and (3.28)¹, and can be compared with the functional forms of the TMDs used to generate the pseudo-data, $f_{NP,\text{real}}$ and $D_{NP,\text{real}}$. Table 5.3 summarizes the kinematic points used for the plots and provides references to the figures where the corresponding plots can be found.

For each pair of figures listed in table 5.3, the first figure shows the functional form of the non-perturbative TMD and its L0, L1, and L2 one-sigma spreads. The bands in the second figure represent the standard deviations of the following ratios:

$$R_{L0} = \frac{f_{NP,\text{pred}}^{L0}(x, b_T; \zeta)}{f_{NP,\text{real}}(x, b_T; \zeta)} - 1 \quad (5.13)$$

$$R_{L1} = \frac{f_{NP,\text{pred}}^{L1}(x, b_T; \zeta)}{f_{NP,\text{real}}(x, b_T; \zeta)} - 1 \quad (5.14)$$

$$R_{L2} = \frac{f_{NP,\text{pred}}^{L2}(x, b_T; \zeta)}{f_{NP,\text{real}}(x, b_T; \zeta)} - 1, \quad (5.15)$$

¹Here, we refer to $f_{NP,*}$ and $D_{NP,*}$ as the predicted TMDs, using the labels $f_{NP,\text{pred}}$ and $D_{NP,\text{pred}}$ for consistency with the plots labels.

where $f_{NP,\text{pred}}^{Li}$ is the predicted functional form of the TMD PDF (or FF) from a Level i closure test, and $f_{NP,\text{real}}$ is the functional form used to generate the pseudo-data. We subtract 1 from each ratio to center the quantity around zero. Plotting the one-sigma bands of the ratios provides a clearer visualization of the L0, L1, and L2 uncertainties, offering a representation of the intervals in b_T where the real TMDs lie within the one-sigma bands of the predicted TMDs. Specifically, in the b_T regions where the line $R_{Li} = 0$ lies within the standard deviation band of R_{Li} , the real TMD lies within the standard deviation bar of the predicted TMD from a fit on Level i pseudo-data.

In this analysis, we only plot the non-perturbative component of the TMDs, as these are the components inferred from experimental data. The L0, L1, and L2 uncertainty levels of the non-perturbative TMDs at a given kinematic point maintain their proportionality when considering the complete TMDs. We also observe that the non-perturbative TMD uncertainties decrease in the Q-region where the data are more precise, *i.e.*, at $Q = 10$ GeV. In figure 3.1 are depicted the kinematic regions covered by our dataset. The LHC data are the more precise, and lie at a scale of $Q = 100$ GeV; thus, we expect the non-perturbative TMD PDF at 100 GeV to have a low total uncertainty. The plot of the TMD PDF at 100 GeV (figure 5.18) confirms that the uncertainty decreases in the kinematic regions where the data are more precise.

Another qualitative insight into the uncertainty levels can be obtained by plotting the contributions of the L0 and L1 uncertainties to the total (L2) uncertainty:

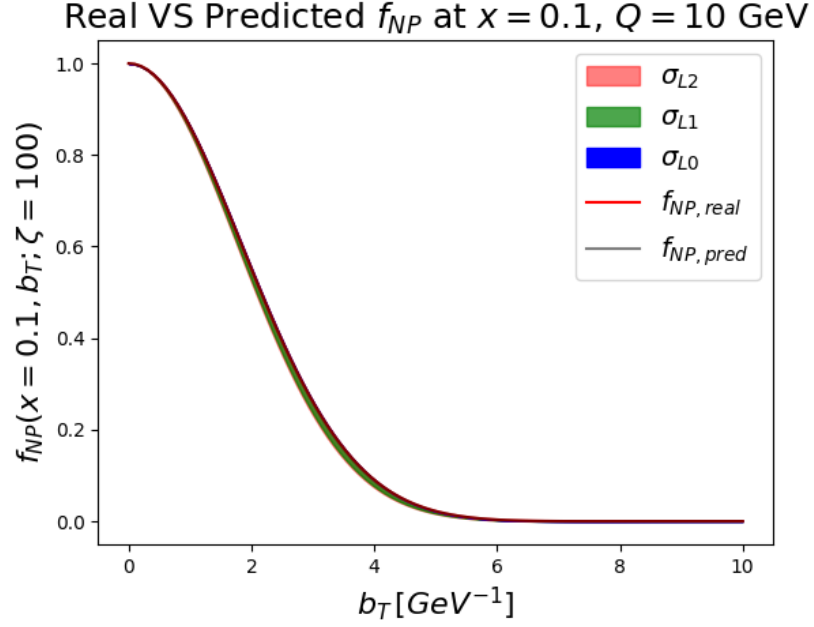
$$\hat{\sigma}_{L0}(x, b_T; \zeta) = \frac{\sigma_{L0}(x, b_T; \zeta)}{\sigma_{L2}(x, b_T; \zeta)} \quad (5.16)$$

$$\hat{\sigma}_{L1}(x, b_T; \zeta) = \frac{\sigma_{L1}(x, b_T; \zeta)}{\sigma_{L2}(x, b_T; \zeta)} \quad (5.17)$$

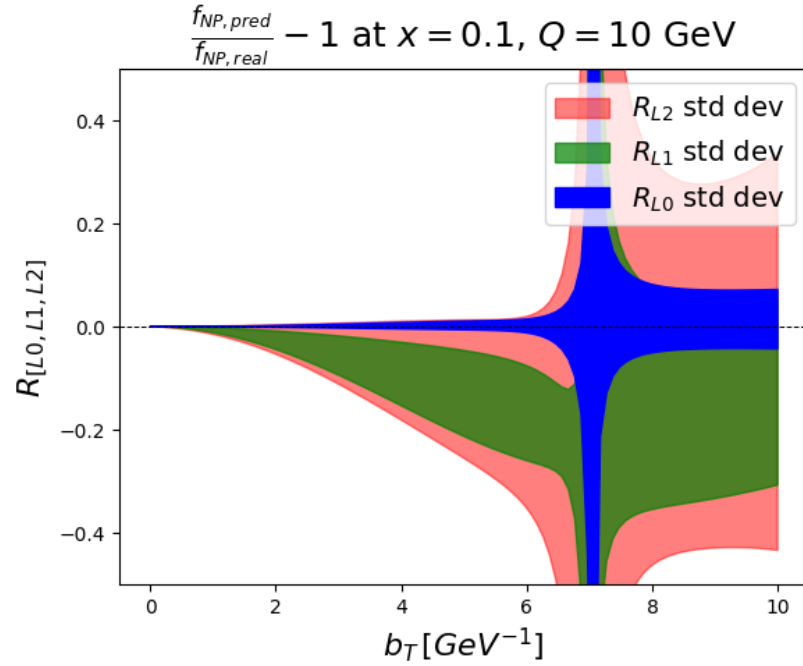
$$\text{total uncertainty} = \frac{\sigma_{L2}(x, b_T; \zeta)}{\sigma_{L2}(x, b_T; \zeta)} = 1 . \quad (5.18)$$

From the plots of $\hat{\sigma}_{L0}$ and $\hat{\sigma}_{L1}$ as functions of b_T at fixed Q , x , and z (where we use the same values for these three variables as in the previous plots, listed in table 5.4), we observe that the dominant contributions to the total uncertainty are the L1 and L2 uncertainties. The L0 uncertainty exceeds 10% of the total uncertainty in rare cases, meaning that *if we want to improve the predictivity of the model, we need to lower the experimental and functional uncertainties. This can be achieved by increasing the precision of the experimental data.* Simply increasing the number of data points for each experiment would not significantly improve the model's predictivity, as the interpolation/extrapolation error of the model already is a subdominant component.

In the figures listed in table 5.3, we observe that, in various kinematic regions, the real TMDs are not included within the one-sigma band of the predicted TMDs at L1, σ_{L1} . This suggests that the L1 error bar might be underestimated (we will quantify the uncertainty reliability in the next section with the multi-closure test on MAP22oMAP22). This underestimation could lead to an underfitting issue. We can

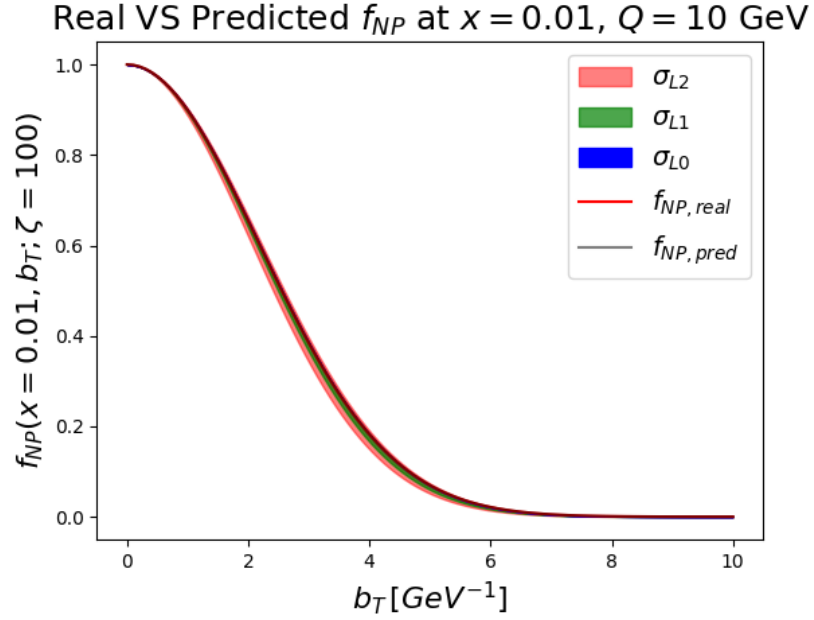


(a) Real and predicted non-perturbative TMD PDFs

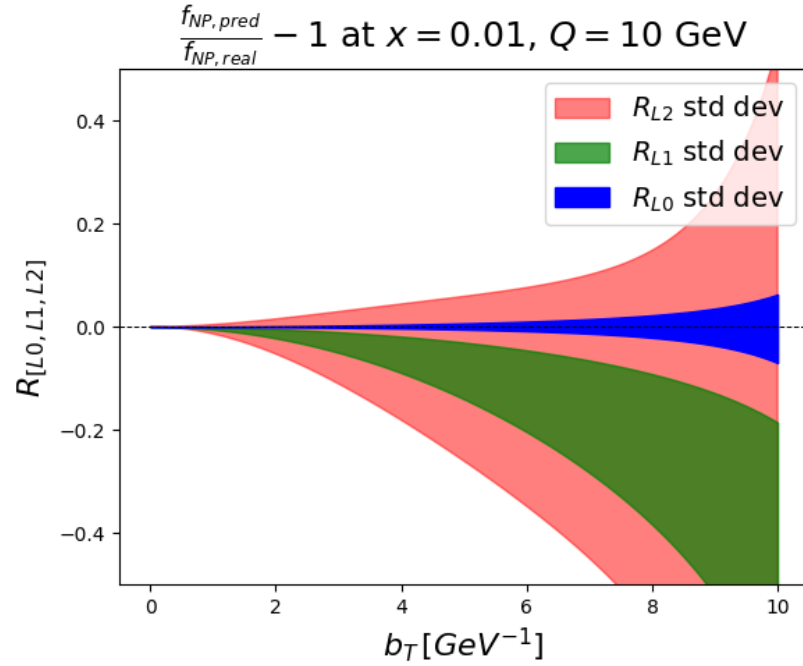


(b) Uncertainty components of the non-perturbative TMD PDF

 Figure 5.8: Plots at $Q = 10$ GeV, $x = 0.1$

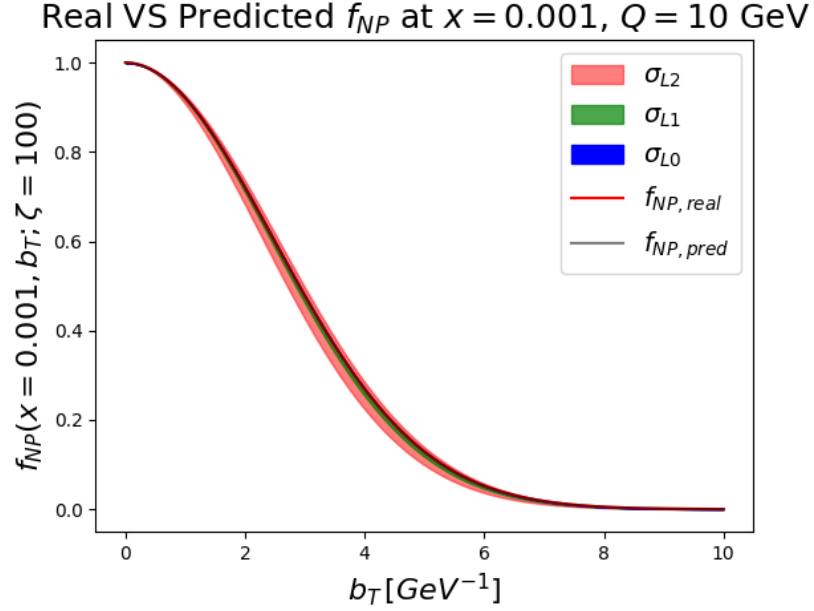


(a) Real and predicted non-perturbative TMD PDFs

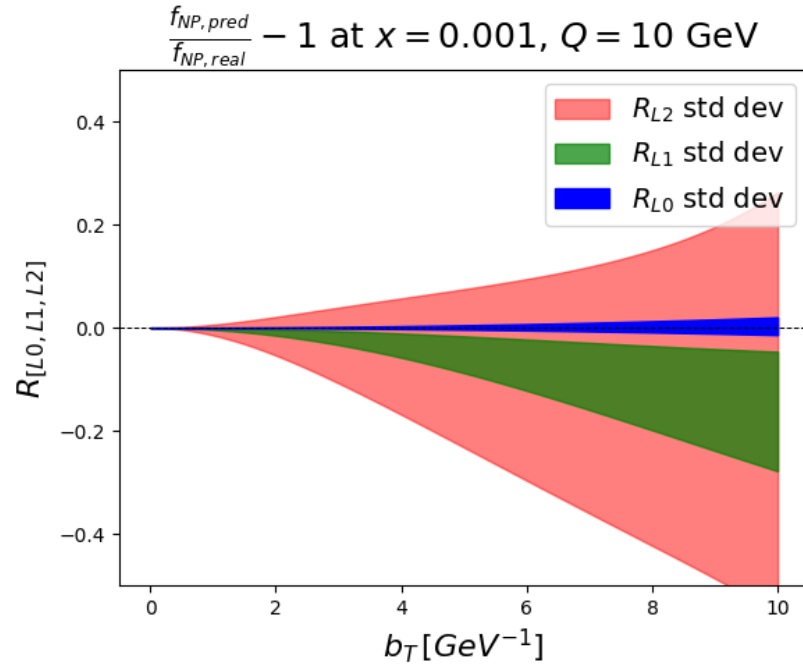


(b) Uncertainty components of the non-perturbative TMD PDF

 Figure 5.9: Plots at $Q = 10$ GeV, $x = 0.01$

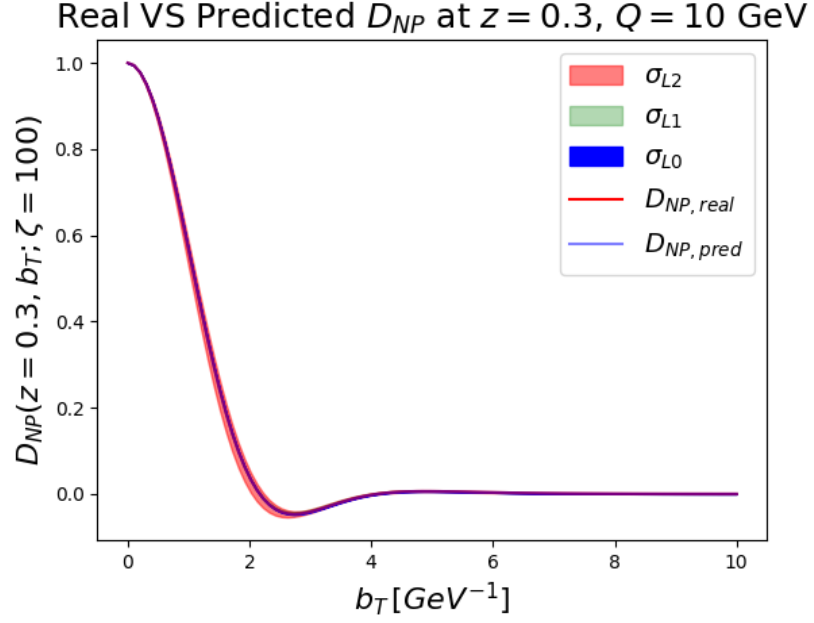


(a) Real and predicted non-perturbative TMD PDFs

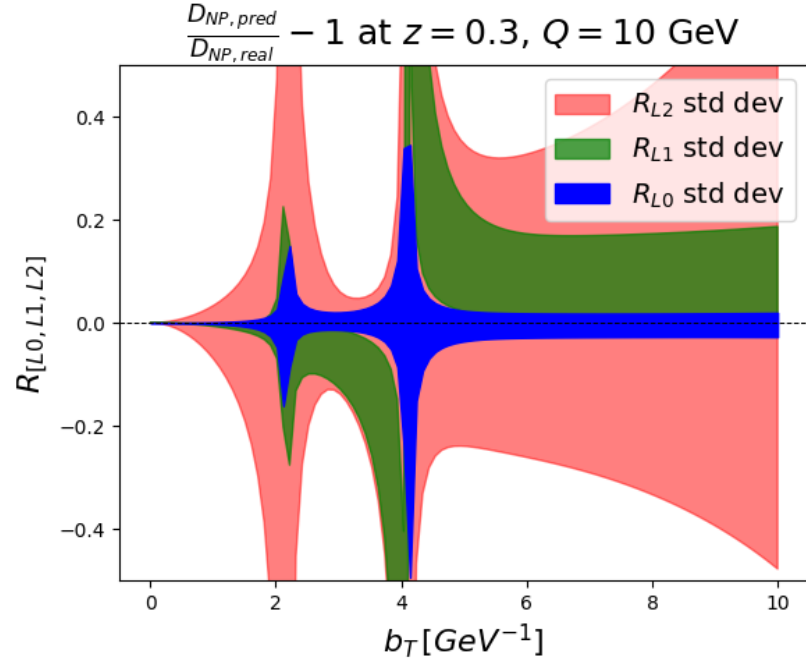


(b) Uncertainty components of the non-perturbative TMD PDF

 Figure 5.10: Plots at $Q = 10$ GeV, $x = 0.001$

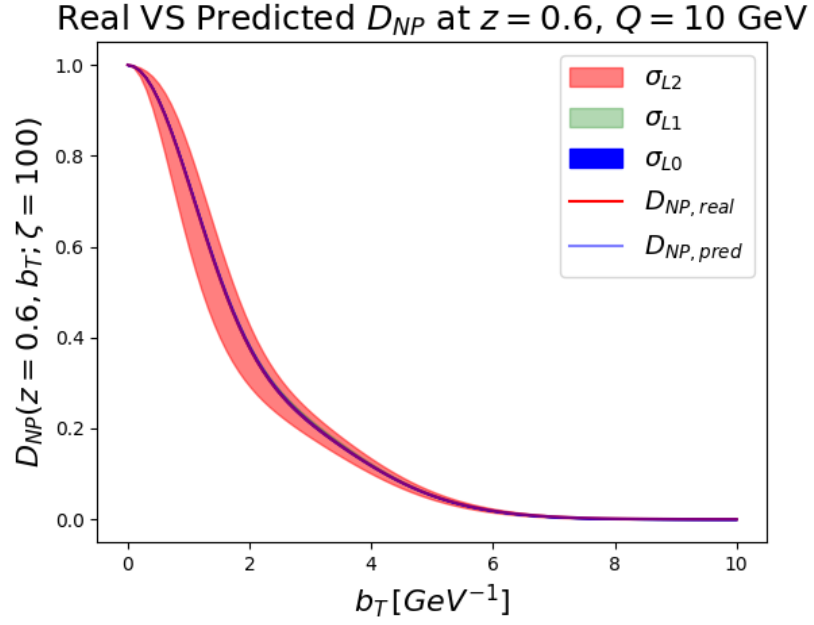


(a) Real and predicted non-perturbative TMD FFs

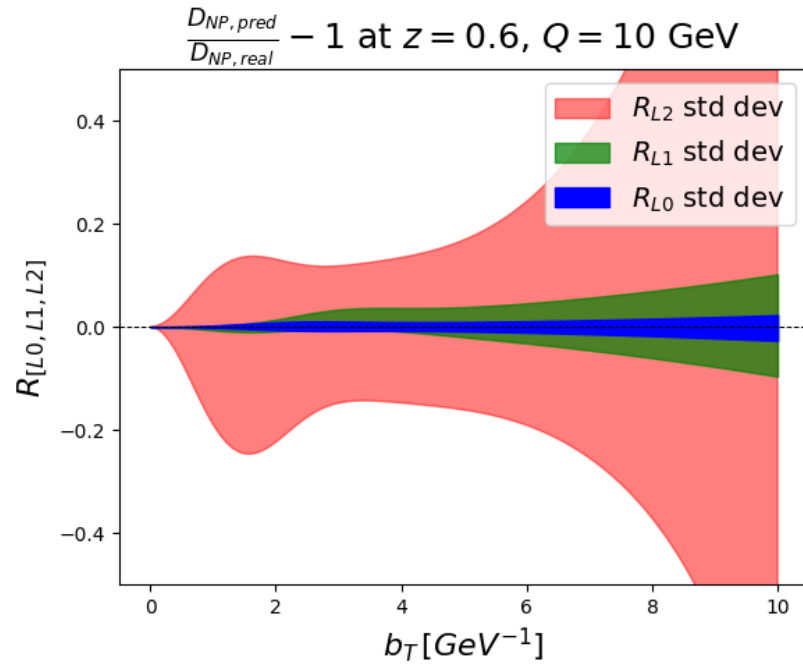


(b) Uncertainty components of the non-perturbative TMD FF

 Figure 5.11: Plots at $Q = 10$ GeV, $z = 0.3$

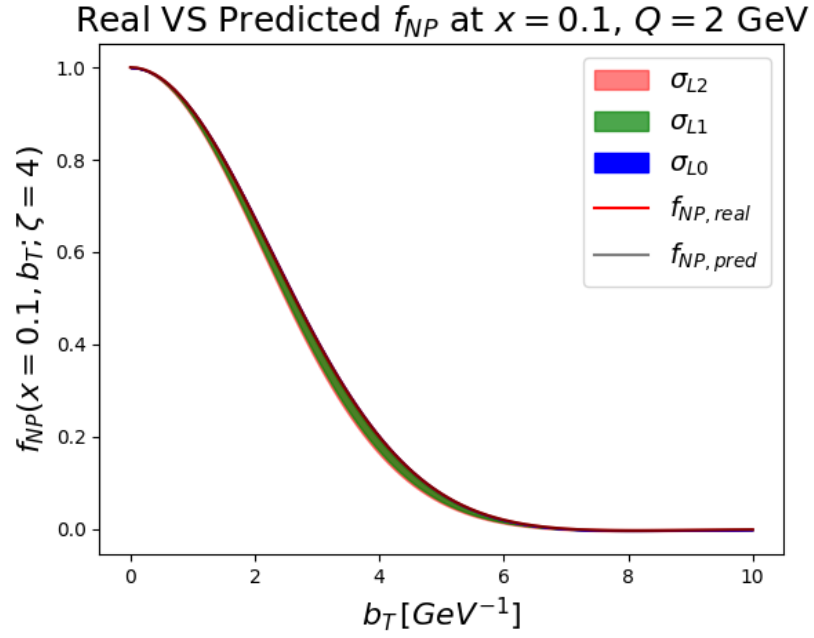


(a) Real and predicted non-perturbative TMD FFs

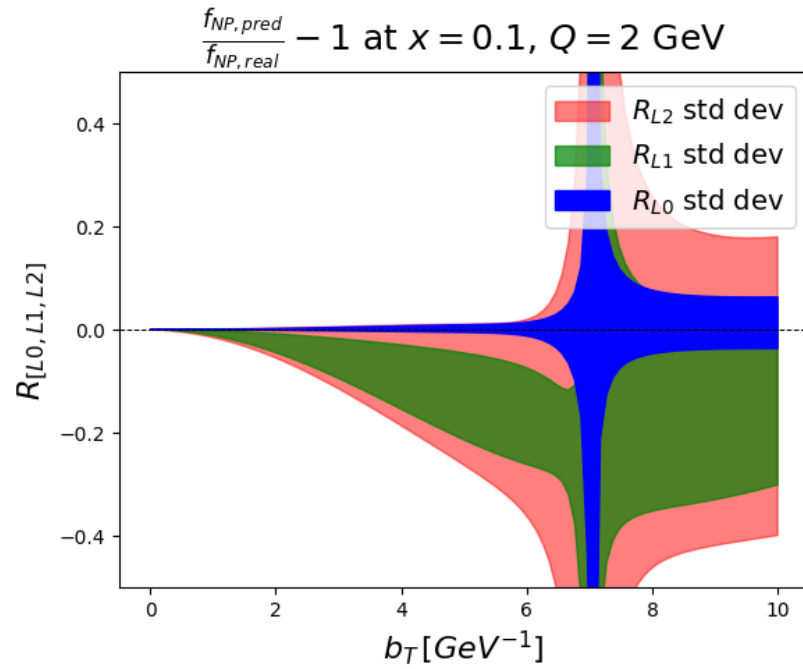


(b) Uncertainty components of the non-perturbative TMD FF

 Figure 5.12: Plots at $Q = 10$ GeV, $z = 0.6$

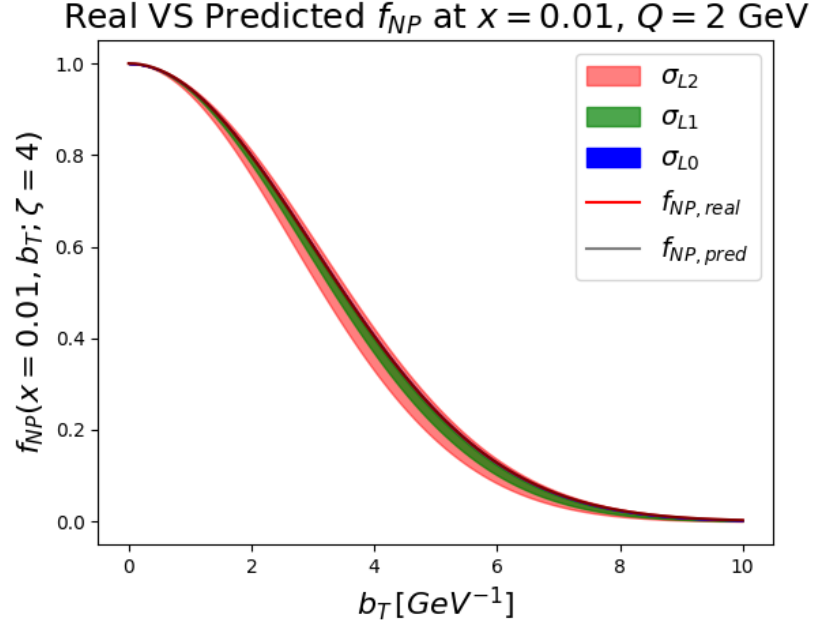


(a) Real and predicted non-perturbative TMD PDFs

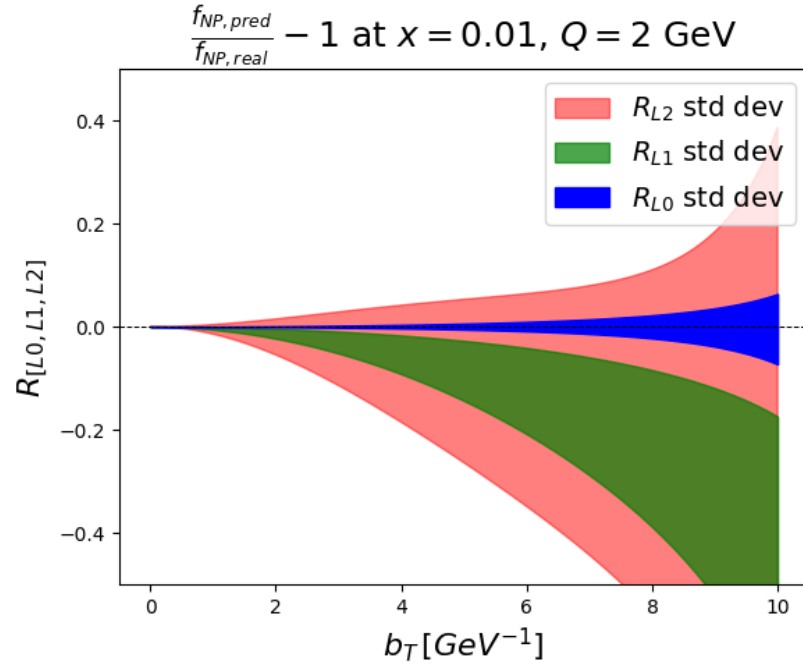


(b) Uncertainty components of the non-perturbative TMD PDF

 Figure 5.13: Plots at $Q = 2$ GeV, $x = 0.1$

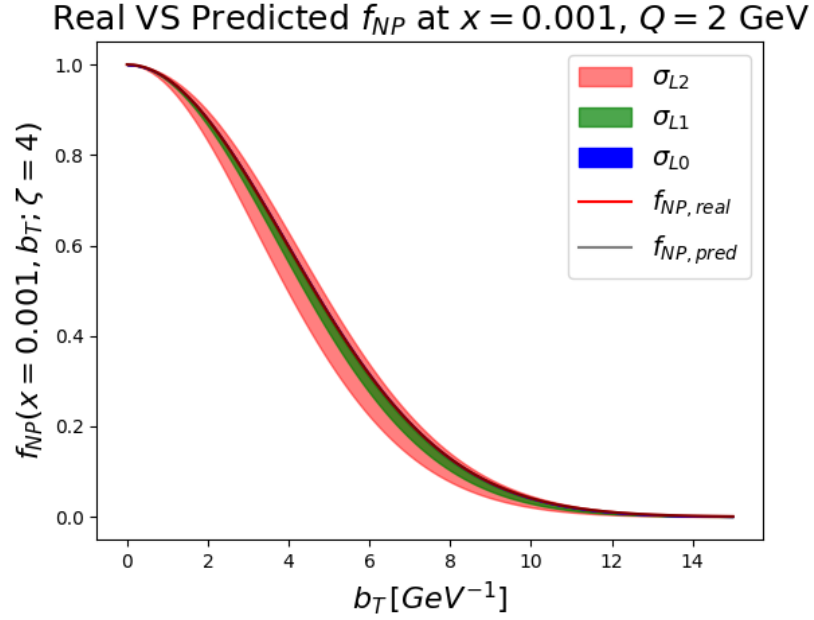


(a) Real and predicted non-perturbative TMD PDFs

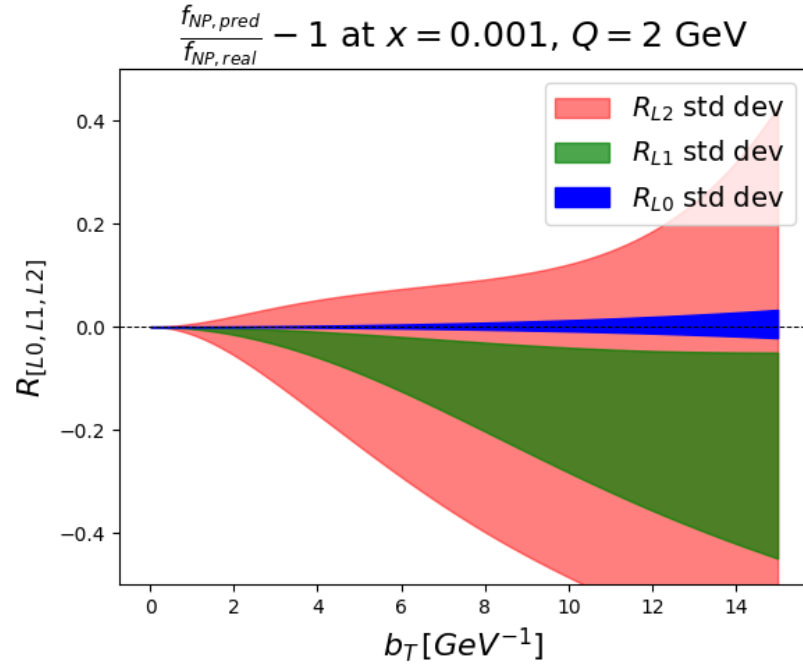


(b) Uncertainty components of the non-perturbative TMD PDF

 Figure 5.14: Plots at $Q = 2$ GeV, $x = 0.01$

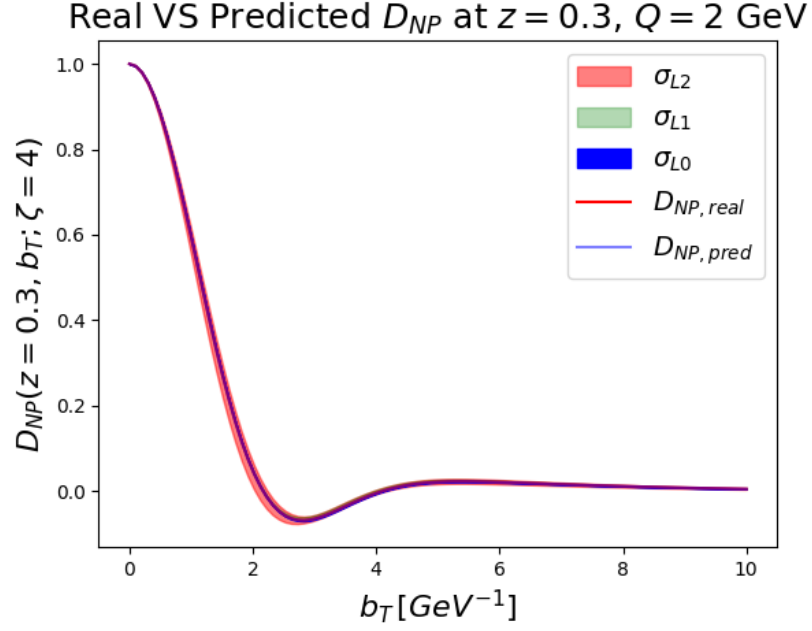


(a) Real and predicted non-perturbative TMD PDFs

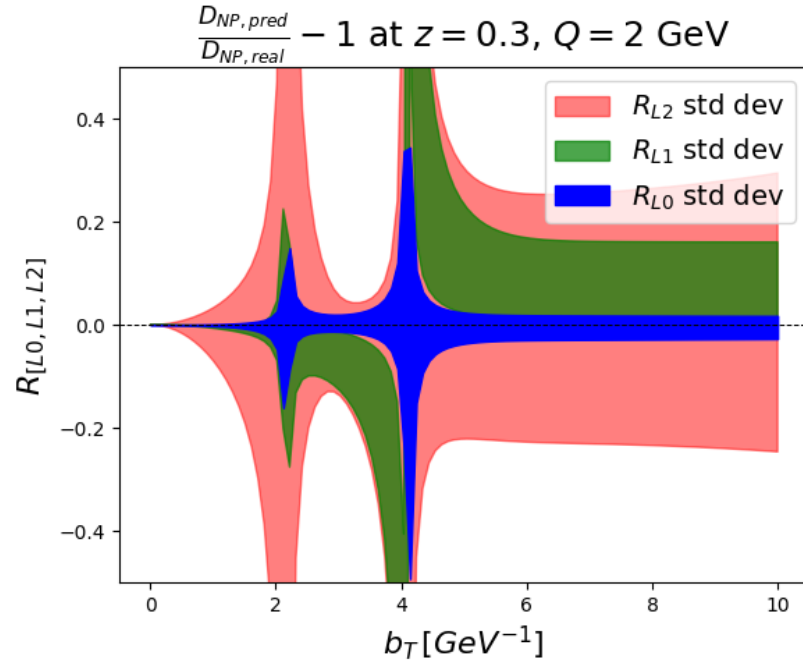


(b) Uncertainty components of the non-perturbative TMD PDF

 Figure 5.15: Plots at $Q = 2$ GeV, $x = 0.001$

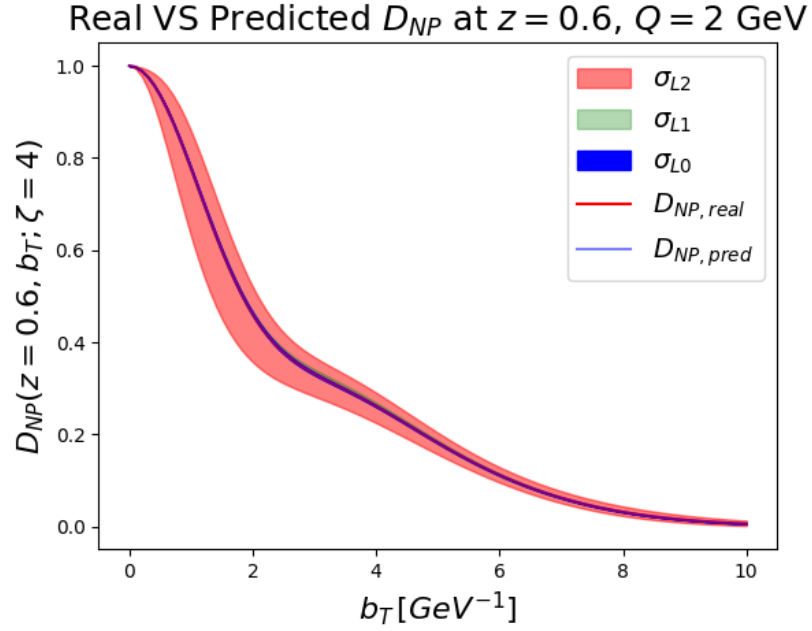


(a) Real and predicted non-perturbative TMD FFs

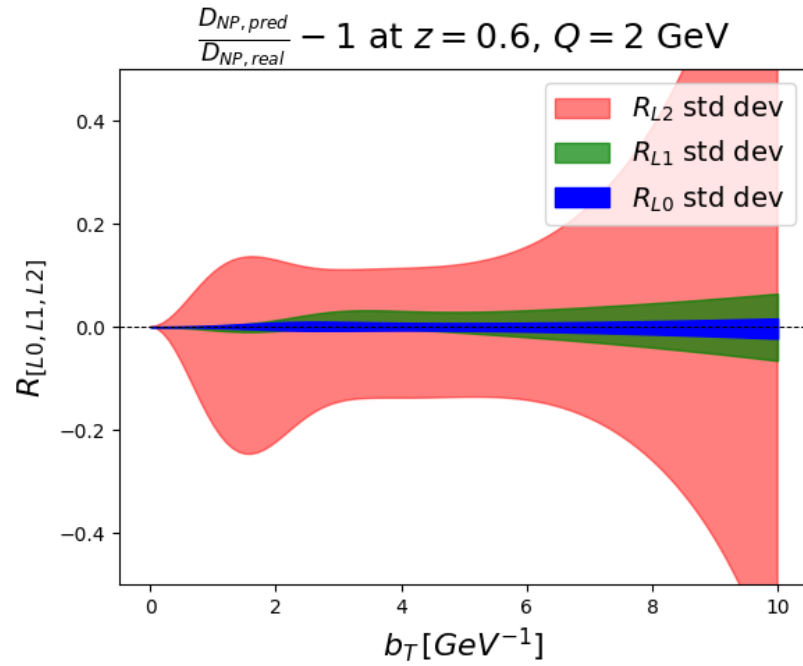


(b) Uncertainty components of the non-perturbative TMD FF

 Figure 5.16: Plots at $Q = 2$ GeV, $z = 0.3$



(a) Real and predicted non-perturbative TMD FFs



(b) Uncertainty components of the non-perturbative TMD FF

 Figure 5.17: Plots at $Q = 10$ GeV, $z = 0.6$

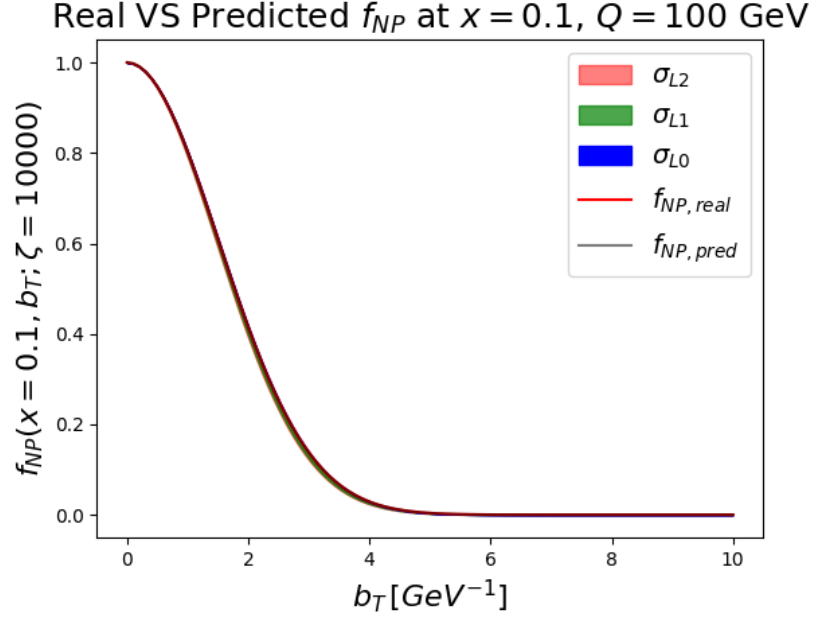


Figure 5.18: Real and predicted $f_{NP}(x = 0.1, b_T, \zeta = 10^4 \text{ GeV}^2)$ in MAPTMD22. This is the Q-region of the LHC data, the fact that the total uncertainty is small compared to the other plots confirms that the uncertainty decreases in the kinematic regions where the data are more precise.

non-perturbative function	Q	x, z	Figure
$f_{NP}(x, b_T, \zeta)$	10 GeV	0.1	figure 5.19
		0.01	
		0.001	
$D_{NP}(z, b_T, \zeta)$	10 GeV	0.3	figure 5.20
		0.6	
$f_{NP}(x, b_T, \zeta)$	2 GeV	0.1	figure 5.21
		0.01	
		0.001	
$D_{NP}(z, b_T, \zeta)$	2 GeV	0.3	figure 5.22
		0.6	

Table 5.4: Plots of the uncertainty ratios $\hat{\sigma}_{L0}$ and $\hat{\sigma}_{L1}$ of the TMD PDFs and FFs predicted by MAPTMD22 in different kinematic regions.

assess whether this is problematic by examining the Level 1 χ^2 . Ideally, the χ^2 between the L1 predictions and the L1 pseudo-dataset, $\chi^2[\mathcal{D}_{\text{pred}}, \mathcal{D}_{L1}]$, should be close to 1 (see section 4.3). From the MAP22oMAP22_{L1} test, we obtain a mean χ^2 of

$$\langle \chi^2[\mathcal{D}_{\text{pred}}, \mathcal{D}_{L1}] \rangle \pm \sigma_{\chi^2} = 1.0055 \pm 0.0159, \quad (5.19)$$

leading us to conclude that the mean displacement between the L1 dataset, $\mathcal{D}_{L1}$, and

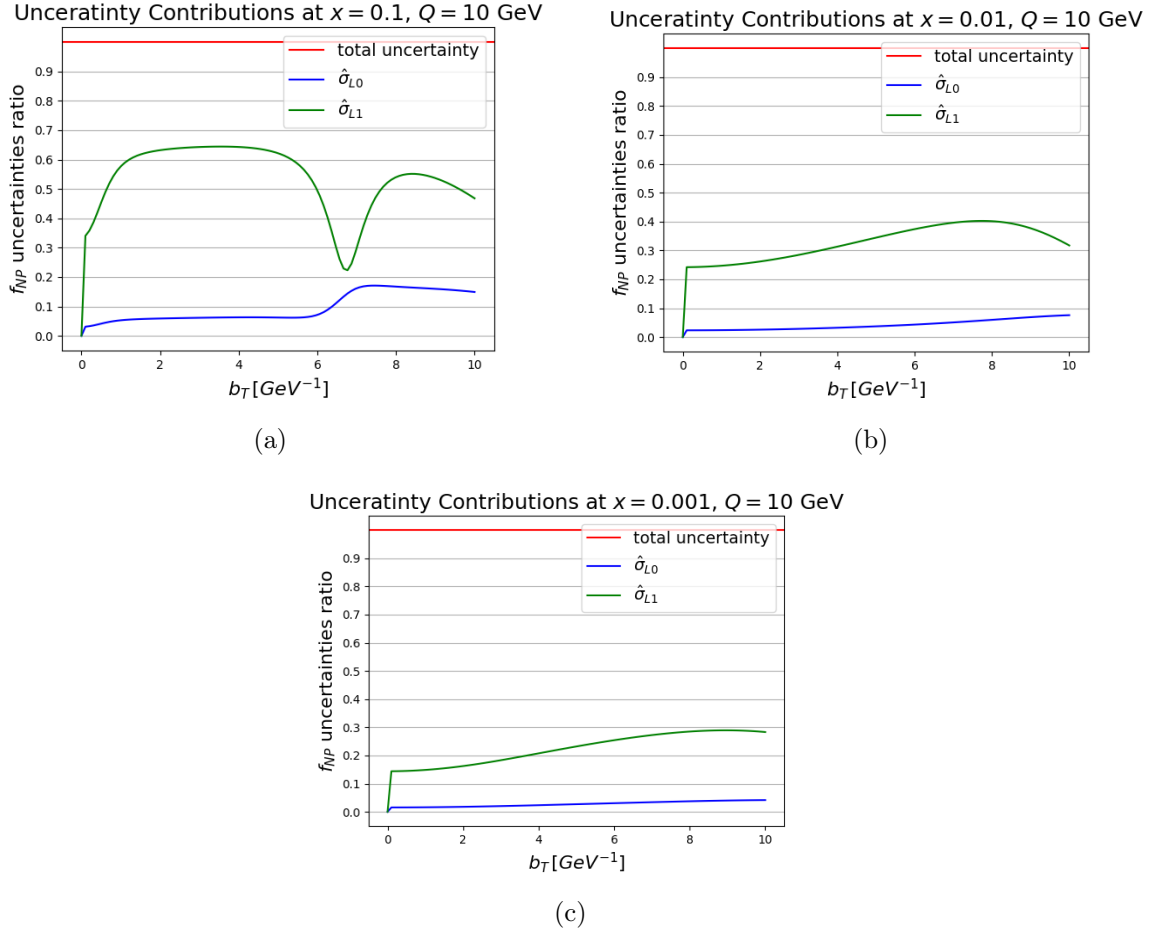


Figure 5.19: Uncertainty ratios $\hat{\sigma}_{L0}$ and $\hat{\sigma}_{L1}$ for the non-perturbative TMD PDF at $Q = 10 \text{ GeV}$

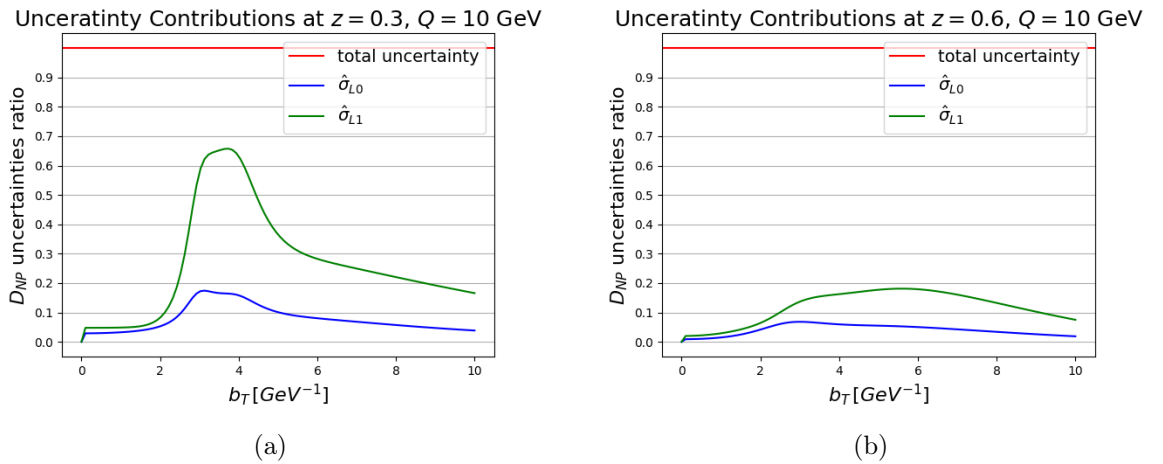


Figure 5.20: Uncertainty ratios $\hat{\sigma}_{L0}$ and $\hat{\sigma}_{L1}$ for the non-perturbative TMD FF at $Q = 10 \text{ GeV}$

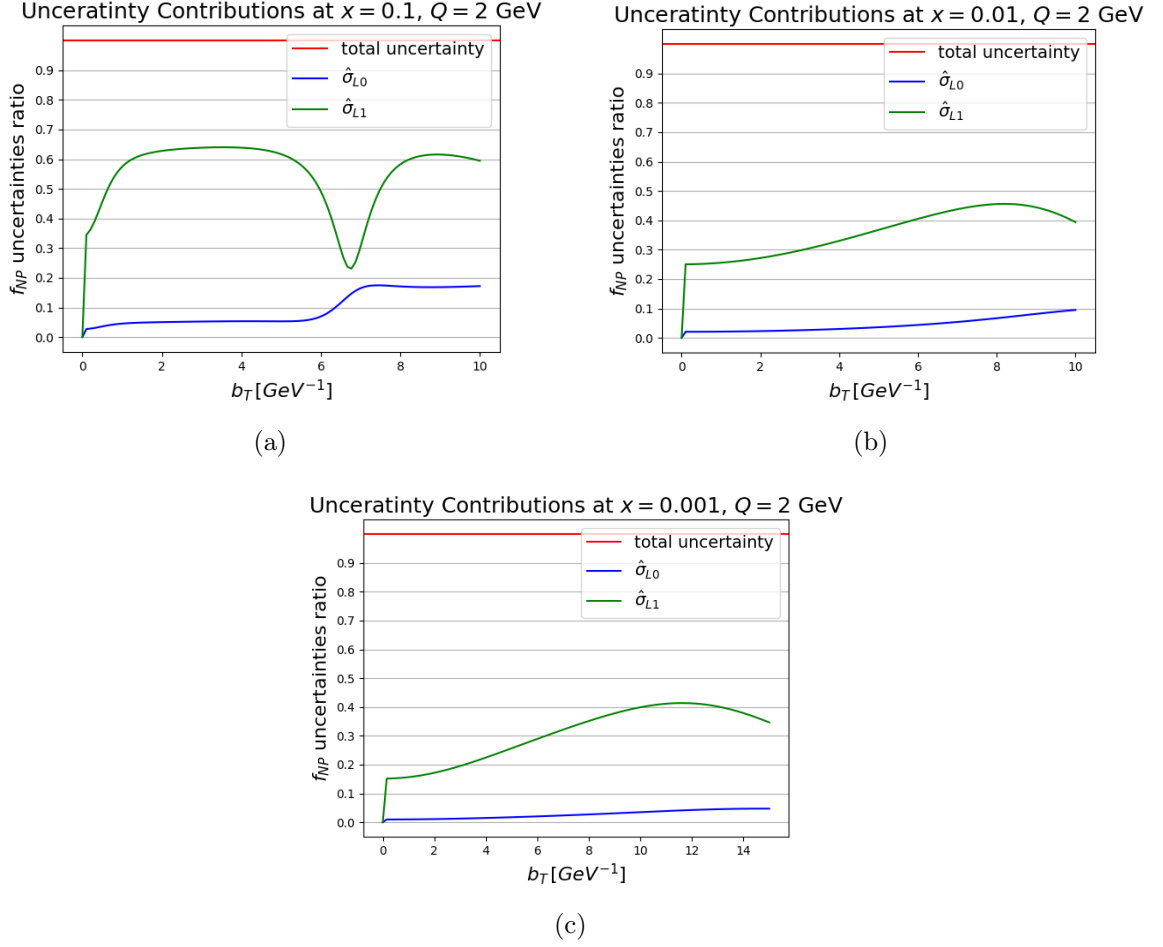


Figure 5.21: Uncertainty ratios $\hat{\sigma}_{L0}$ and $\hat{\sigma}_{L1}$ for the non-perturbative TMD PDF at $Q = 2$ GeV

the predicted dataset $\mathcal{D}_{\text{pred}}$, measured in units of the standard deviation, is 1. An average χ^2 of 1 suggests that the discrepancies between the dataset and the predictions are, on average, consistent with the expected uncertainties. However, this is a qualitative discussion about the properties of the considered framework, and the mean Level 1 χ^2 of a single fit could be subject to fluctuations around the ideal value of 1. A more quantitative estimation of potential underfitting will be addressed in the multi-closure test.

We now proceed to report the uncertainty components of the predicted non-perturbative TMDs using the PV19 framework.

PV19oPV19 Uncertainties

In the PV19oPV19 test, we reproduced the plots of the predicted TMD PDF ($f_{NP,\text{pred}}$) as a function of b_T for fixed x and Q . The obtained plots are listed in table 5.5, which summarizes the kinematic points (x, Q) used and the corresponding figures where the

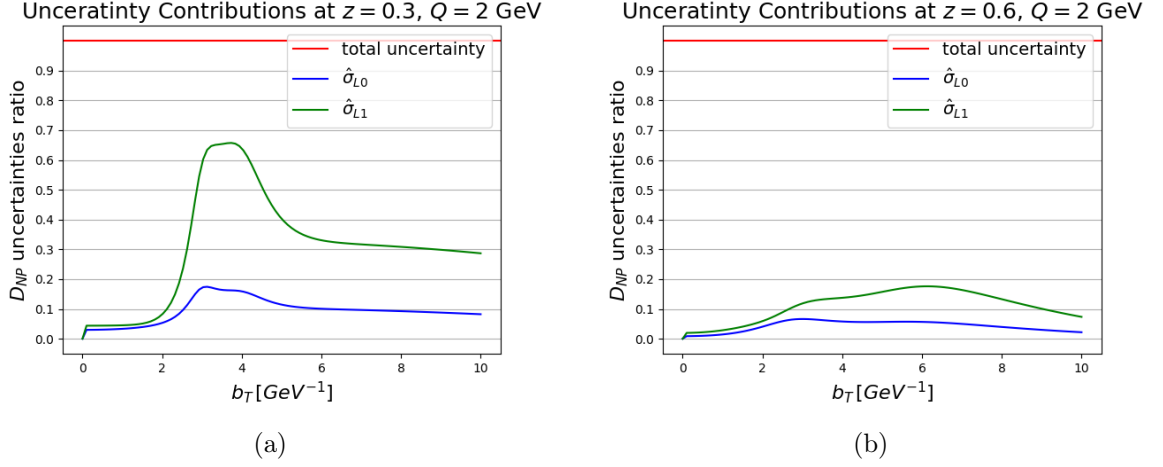


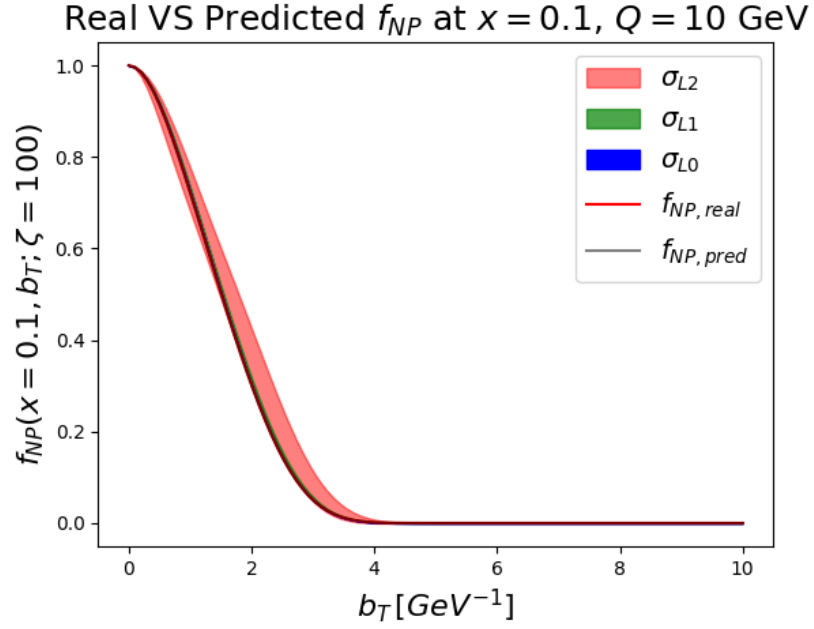
Figure 5.22: Uncertainty ratios $\hat{\sigma}_{L0}$ and $\hat{\sigma}_{L1}$ for the non-perturbative TMD FF at $Q = 2 \text{ GeV}$

plots can be found.

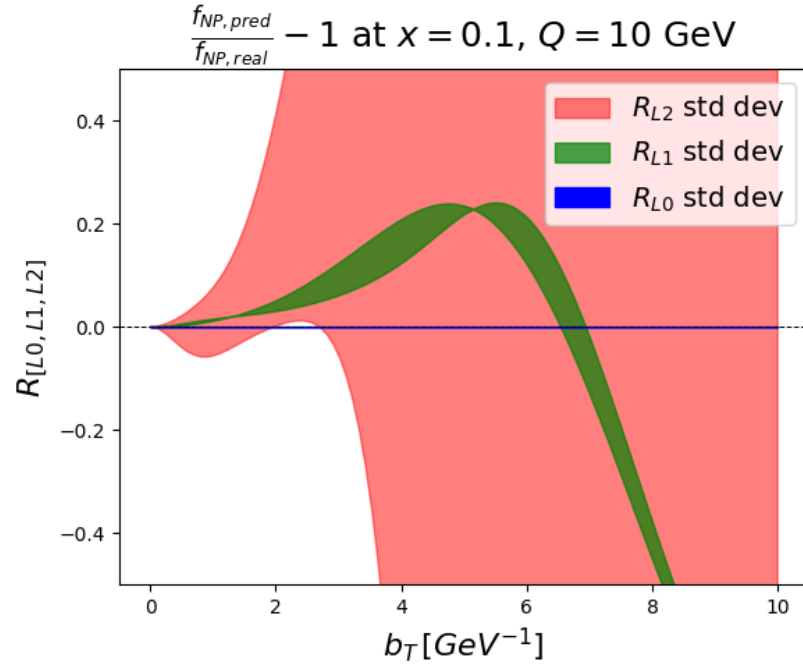
Following the same methodology used in the MAP22oMAP22 test, we also plotted the uncertainty ratios $\hat{\sigma}_{L0}$ and $\hat{\sigma}_{L1}$ (equations (5.16), (5.17)). As before, these plots help us visualize how much of the total uncertainty (σ_{L2}) comes from the interpolation/extrapolation uncertainty (σ_{L0}) and the functional uncertainty (σ_{L1}). The results are summarized in table 5.6, which lists the plots of the ratios for different values of x and Q . From these plots we can observe that the interpolation/extrapolation and functional uncertainties are subdominant in the PV19 framework, meaning that adding more data points would not significantly improve the model's predictivity. Instead, the focus should be on improving the precision of the experimental data, as this would directly reduce the total uncertainty and enhance the predictive power of the PV19 framework. Note that in MAPTMD22 the functional component of the uncertainty was not negligible. The subdominant contribution of the L1 uncertainty could be a symptom of insufficient flexibility of the PV19 parametrization, which produces very similar functional forms across the L1 replicas, leading to a negligible spread of the L1 results.

In fact, the σ_{L1} band is much narrower in PV19 than in MAPTMD22. From the plots listed in table 5.5, we notice that, in most cases, the L1 standard deviation does not encompass the functional form used to generate the pseudo-data. This is another indicator of possible underestimation of the uncertainties. However, we will explore this issue more in detail in the multi closure tests section.

The tests on the uncertainty characterization presented in this section provide a qualitative representation of the contribution of interpolation/extrapolation, functional, and experimental uncertainties to the total TMD uncertainty declared by the MAPTMD22 and PV19 fitting frameworks. In the next section, we test whether the declared TMD uncertainties are reliable through multi closure tests.

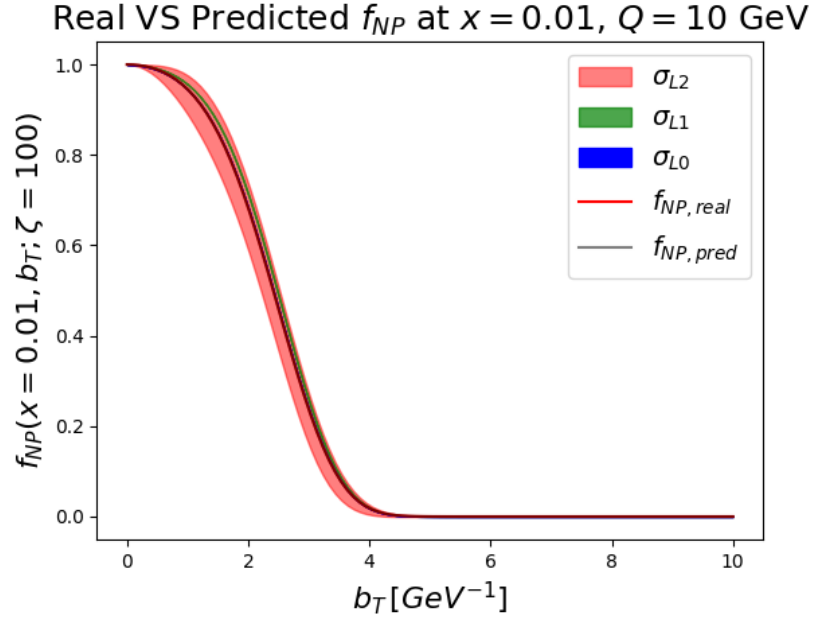


(a) Real and predicted non-perturbative TMD PDFs

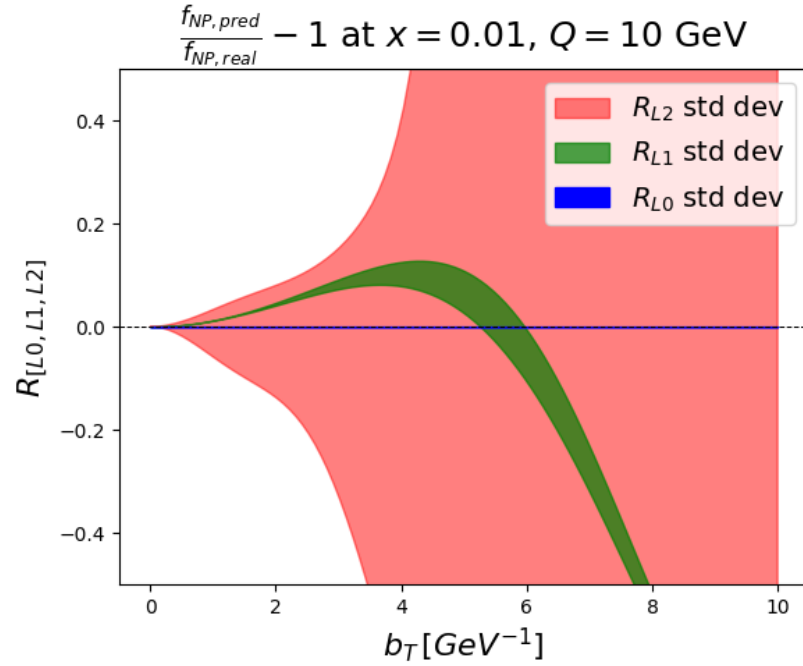


(b) Uncertainty components of the non-perturbative TMD PDF

 Figure 5.23: Plots at $Q = 10$ GeV, $x = 0.1$

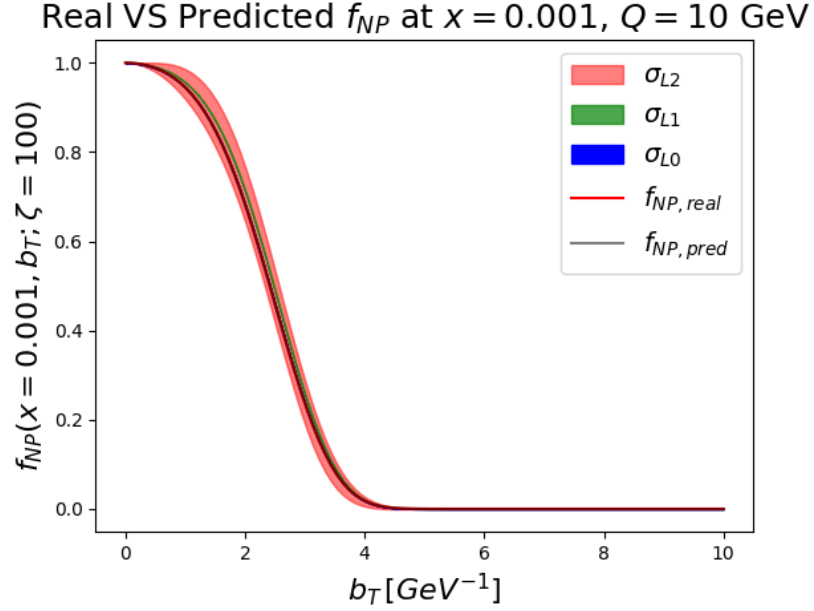


(a) Real and predicted non-perturbative TMD PDFs

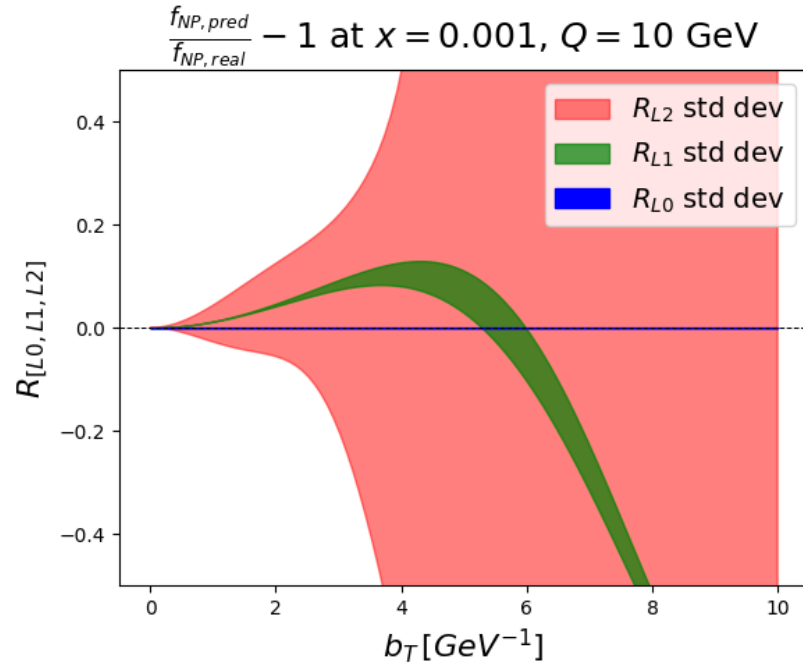


(b) Uncertainty components of the non-perturbative TMD PDF

 Figure 5.24: Plots at $Q = 10$ GeV, $x = 0.01$

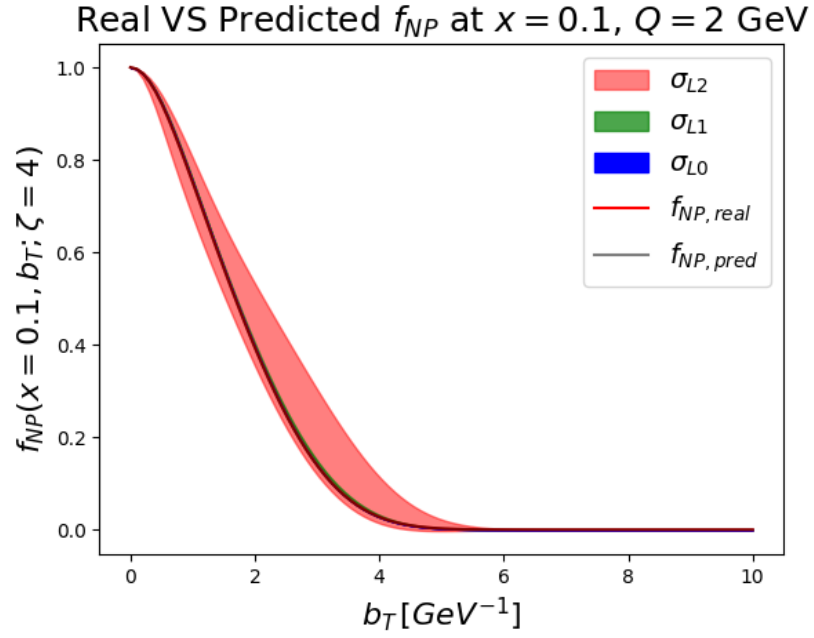


(a) Real and predicted non-perturbative TMD PDFs

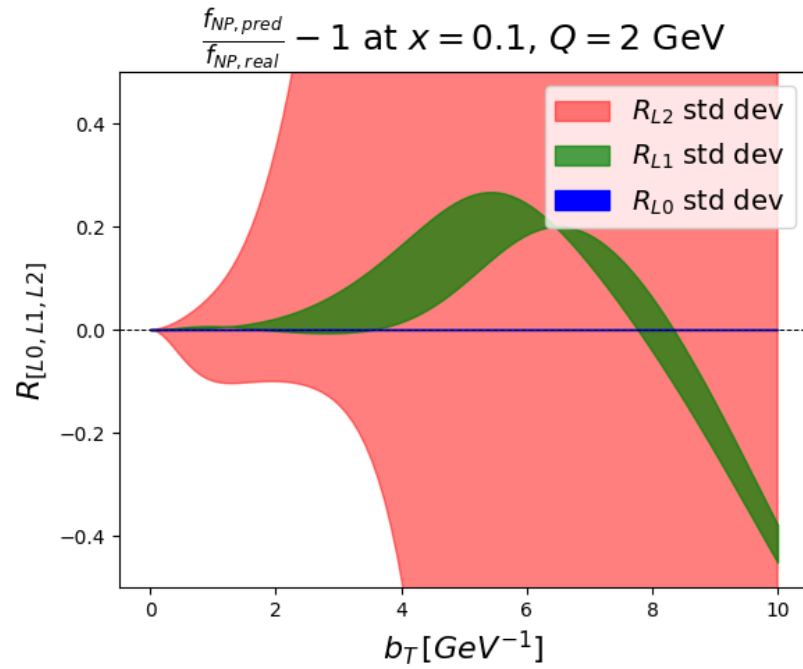


(b) Uncertainty components of the non-perturbative TMD PDF

 Figure 5.25: Plots at $Q = 10$ GeV, $x = 0.001$

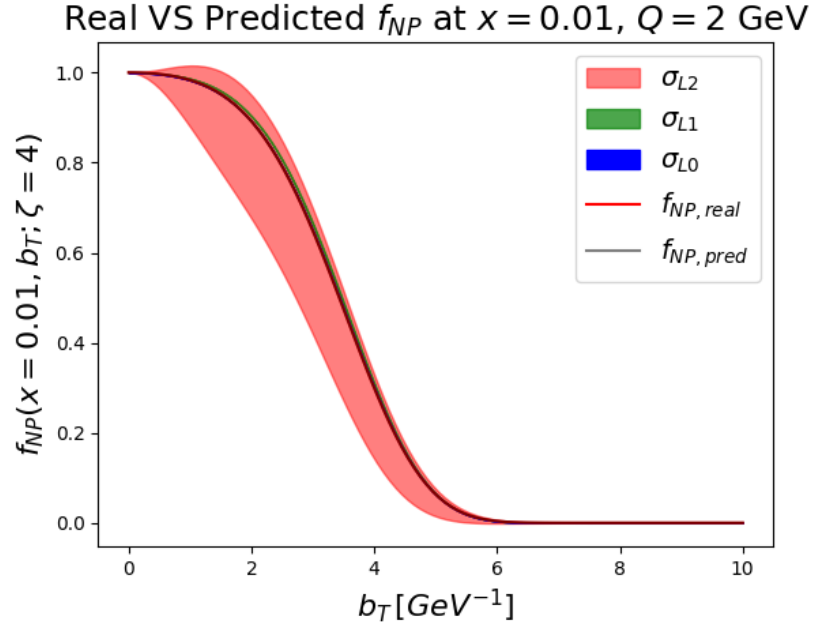


(a) Real and predicted non-perturbative TMD PDFs

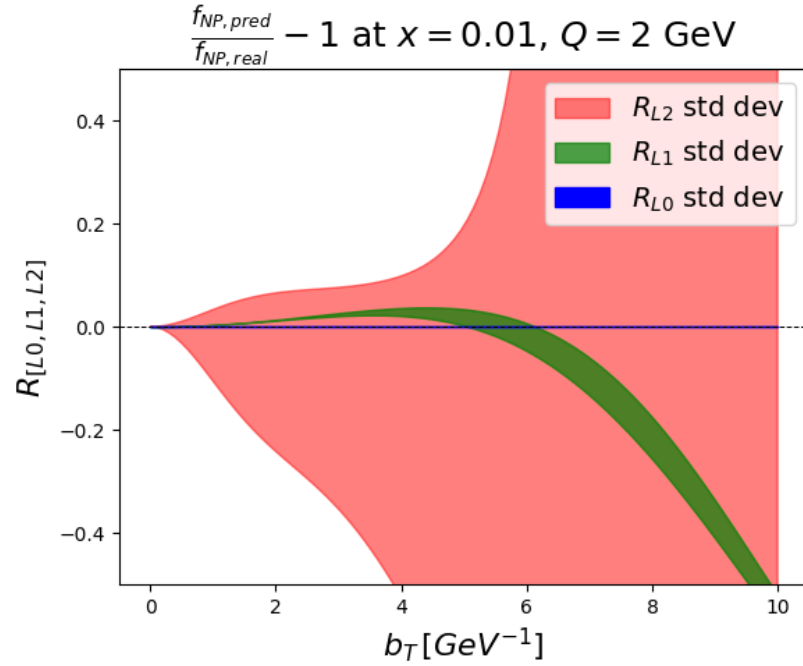


(b) Uncertainty components of the non-perturbative TMD PDF

 Figure 5.26: Plots at $Q = 2$ GeV, $x = 0.1$

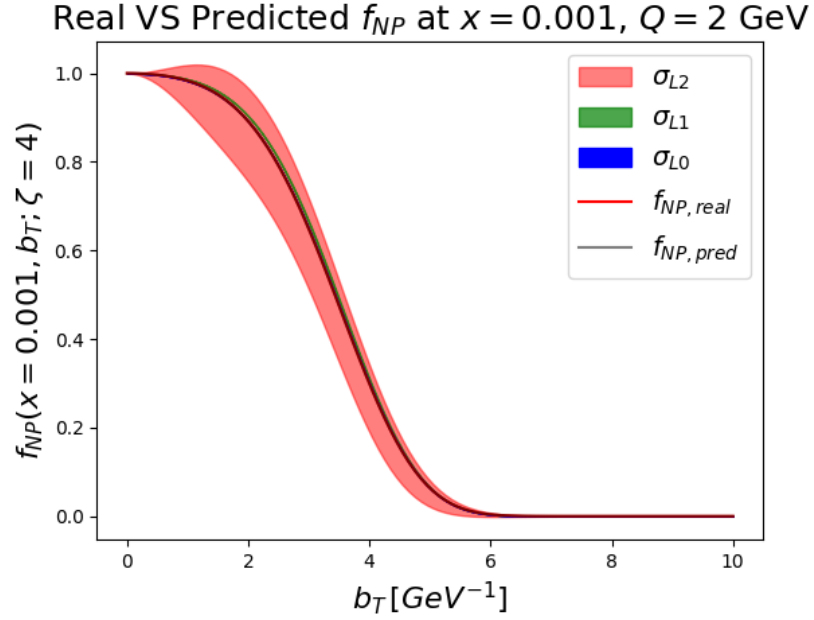


(a) Real and predicted non-perturbative TMD PDFs

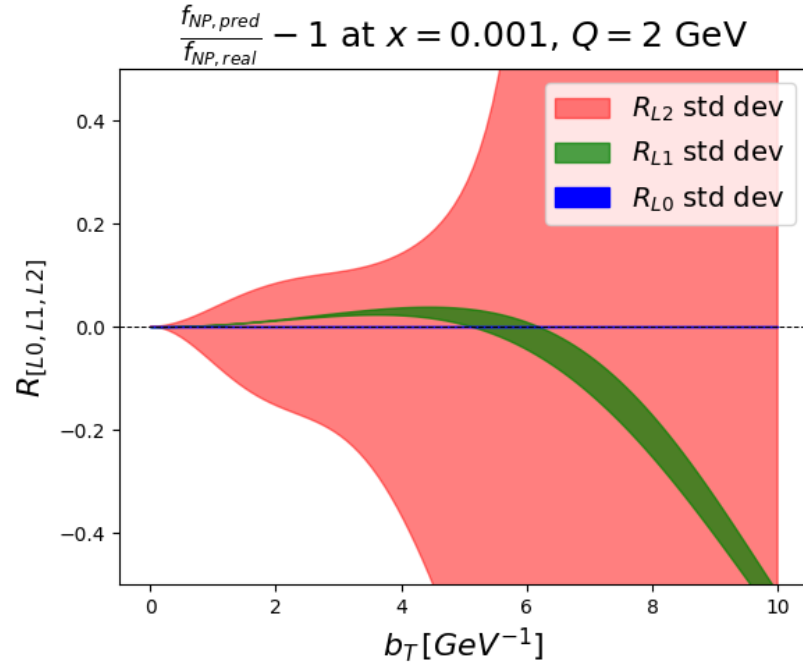


(b) Uncertainty components of the non-perturbative TMD PDF

 Figure 5.27: Plots at $Q = 2$ GeV, $x = 0.01$



(a) Real and predicted non-perturbative TMD PDFs



(b) Uncertainty components of the non-perturbative TMD PDF

 Figure 5.28: Plots at $Q = 2$ GeV, $x = 0.001$

non-perturbative function	Q	x	Figure
$f_{NP}(x, b_T, \zeta)$	10 GeV	0.1	figure 5.23
		0.01	figure 5.24
		0.001	figure 5.25
$f_{NP}(x, b_T, \zeta)$	2 GeV	0.1	figure 5.26
		0.01	figure 5.27
		0.001	figure 5.28

Table 5.5: Plots of the uncertainties of the TMD PDFs predicted by PV19 in different kinematic regions.

non-perturbative function	Q	x	Figure
$f_{NP}(x, b_T, \zeta)$	10 GeV	0.1	figure 5.29
		0.01	
		0.001	
$f_{NP}(x, b_T, \zeta)$	2 GeV	0.1	figure 5.30
		0.01	
		0.001	

Table 5.6: Plots of the uncertainty ratios $\hat{\sigma}_{L0}$ and $\hat{\sigma}_{L1}$ of the TMD PDFs predicted by PV19 in different kinematic regions.

5.4 Bias-Variance Tradeoff and Uncertainty Faithfulness

Through a multi-closure test, we can assess the uncertainty faithfulness of a methodology and determine whether the methodology is optimized in terms of the bias-variance tradeoff (see section 4.5). We introduced two quantitative estimators:

- the bias-variance ratio, $\mathcal{R}_{bv}$ (equation (4.28)), which is compatible with 1 if the bias-variance tradeoff of the model is optimized, (*i.e.* if the average model error across the fits $\mathbb{E}_{\mathcal{D}_{L1}}[\mathcal{E}^{\text{fit}}]$ is at the minimum (see figure 4.7);
- the quantile estimator, $\xi_{1\sigma}$ (equation 4.35), which is compatible with 0.683 if the uncertainties predicted by the model are faithful to the real TMD uncertainties.

For both MAPTMD22 and PV19, we performed fits on a total of 5000 replicas (50 fits, each composed of 100 replicas) to compute the two values $\mathcal{R}_{bv}$ and $\xi_{1\sigma}$. A framework is optimized and has faithful uncertainties if we obtain:

$$\mathcal{R}_{bv} \approx 1 \tag{5.20}$$

$$\xi_{1\sigma} \approx 0.683 . \tag{5.21}$$

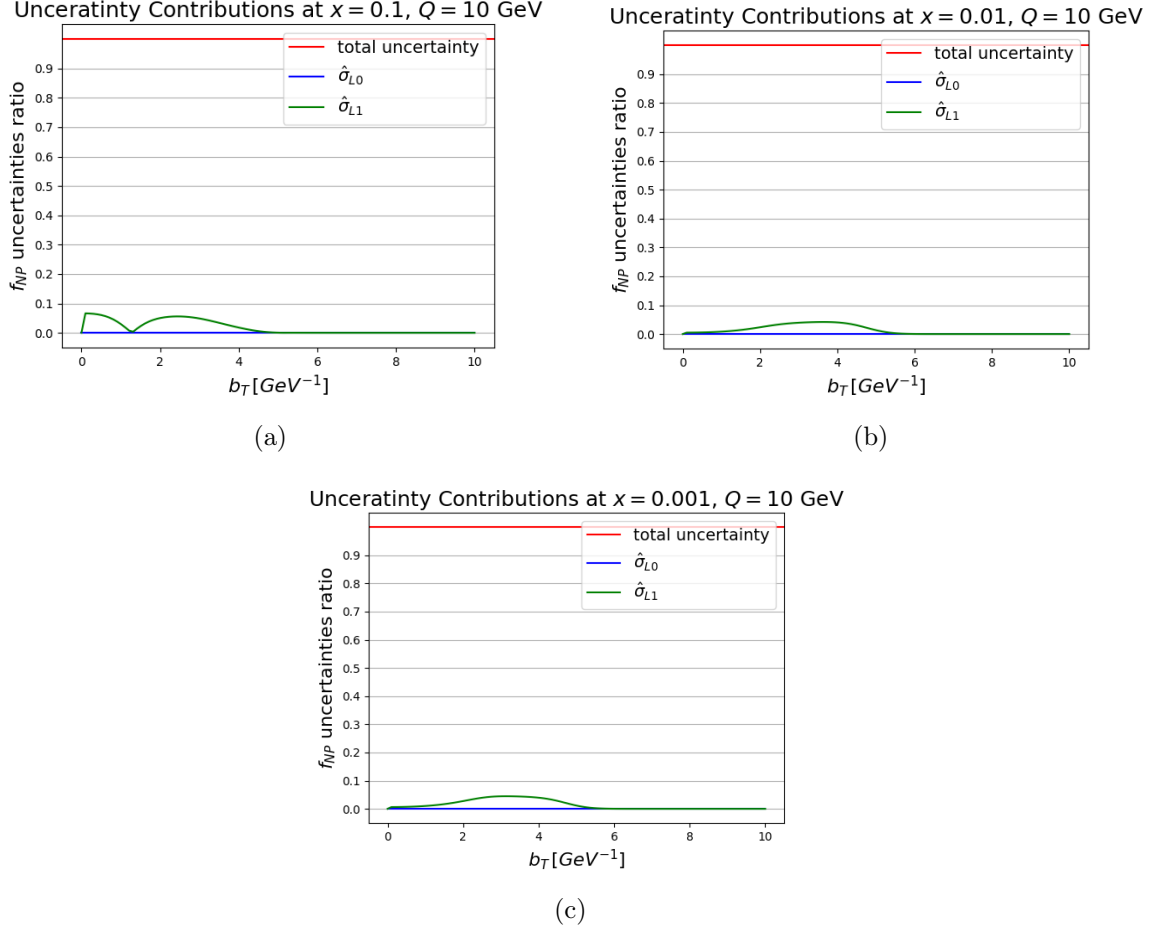


Figure 5.29: Uncertainty ratios $\hat{\sigma}_{L0}$ and $\hat{\sigma}_{L1}$ for the non-perturbative TMD PDF at $Q = 10$ GeV

Multi Closure Test of MAPTMD22

From the test on the MAPTMD22 framework, we obtained:

$$\mathcal{R}_{bv} \pm \sigma_{\mathcal{R}_{bv}} = 0.978 \pm 0.045 \quad (5.22)$$

$$\xi_{1\sigma} \pm \sigma_{\xi_{1\sigma}} = 0.688 \pm 0.010 . \quad (5.23)$$

We thus conclude that the MAPTMD22 framework is optimized and declares reliable uncertainties. This result is compatible with the observations we made for the χ^2 obtained by a Level 1 test on MAP22 (see equation (5.19)).

Multi Closure Test of PV19

From the test on the PV19 framework, we obtained:

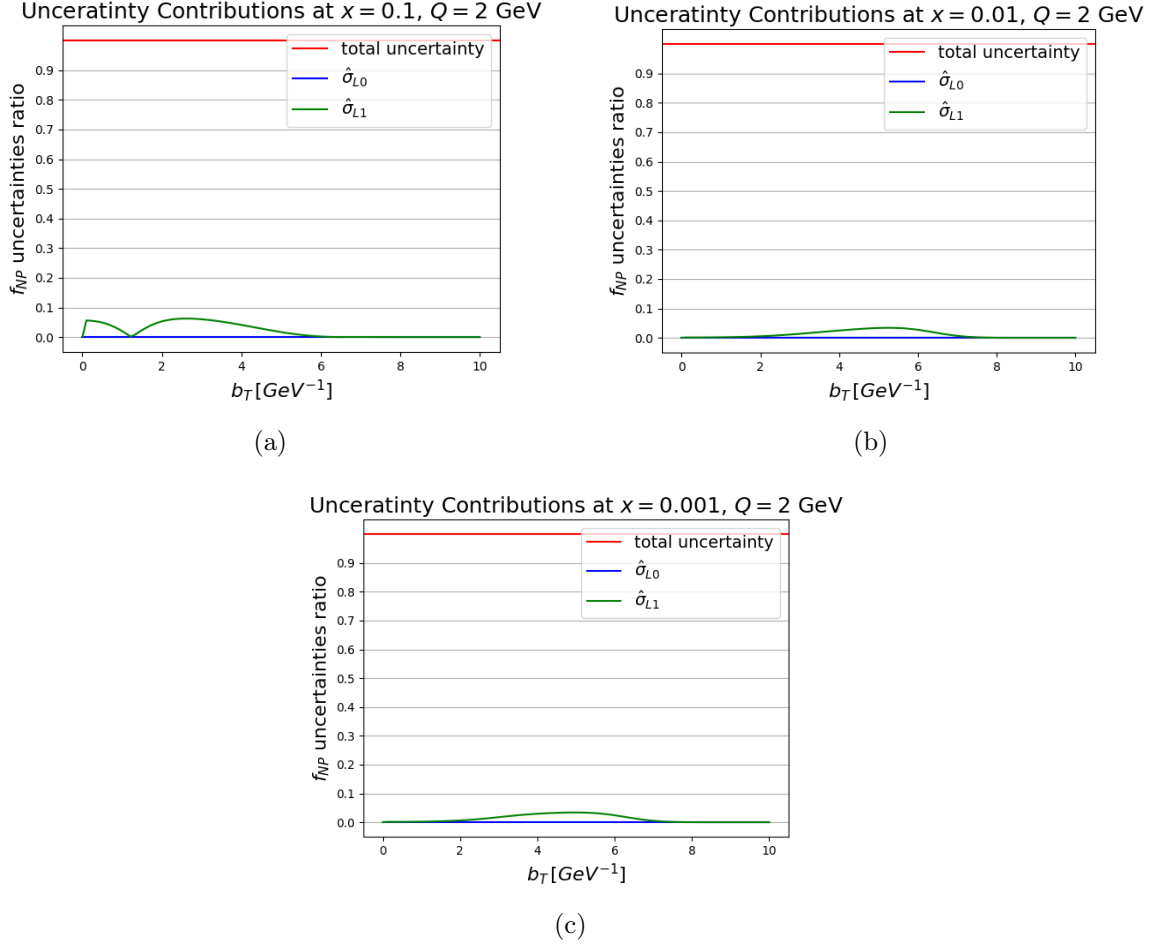


Figure 5.30: Uncertainty ratios $\hat{\sigma}_{L0}$ and $\hat{\sigma}_{L1}$ for the non-perturbative TMD PDF at $Q = 2$ GeV

$$\mathcal{R}_{bv} \pm \sigma_{\mathcal{R}_{bv}} = 1.577 \pm 0.068 \quad (5.24)$$

$$\xi_{1\sigma} \pm \sigma_{\xi_{1\sigma}} = 0.486 \pm 0.012 . \quad (5.25)$$

We thus conclude that the PV19 framework is not optimized and does not declare faithful uncertainties. This result supports our hypothesis of underfitting, which leads to uncertainty underestimation. In fact, a bias-variance ratio estimator greater than one indicates that the bias of the predictions is high compared to the variance, leading to underfitting (see figure 4.7). This result is confirmed by the quantile estimator. Since $\xi_{1\sigma}$ is lower than 0.683, we can conclude that the uncertainties of the TMDs declared by the PV19 framework are *underestimated*, meaning that what is declared to be a 68.3% confidence level is, in reality, a 48.6% confidence level, because the real TMD functional form, used to generate the pseudo-data, falls within the declared one-sigma band only in 48.6% of the cases.

CHAPTER 5. RESULTS

In the next section, we will draw conclusions based on the results obtained from the analyses discussed above.

Chapter 6

Conclusions

In this work, we analyzed, through closure testing, two fitting frameworks for extracting non-perturbative Transverse Momentum Densities (TMDs) from experimental measurements of SIDIS and Drell-Yan cross sections, MAPTMD22 [4] and PV19 [3].

The results of the tests are presented in chapter 5. Here, we summarize the findings and draw conclusions about what they imply for determining TMDs from fits to experimental data. We also explore how the tools developed for conducting closure tests on MAPTMD22 and PV19 can be applied to validate and optimize other existing fitting frameworks as well as future methodologies.

6.1 Closure Test Results

The conclusions drawn from closure testing include insights into the flexibility of the frameworks, the contributions of interpolation/extrapolation, functional, and experimental TMD uncertainties to the overall uncertainty, and the ability of the frameworks to accurately reproduce the TMD uncertainties.

Conclusions on the Frameworks' Flexibility

In the Level 0 test, we observed that both the PV19 and MAPTMD22 frameworks are sufficiently flexible when it comes to reproducing the TMD functional forms generated using their respective parameterizations, *i.e.* when the same functional form is used to *generate* and *fit* a Level 0 pseudo-dataset. However, both parameterizations show limited flexibility when tasked with reproducing a TMD functional form generated using a different parameterization than the one used for fitting. Notably, the PV19 parameterization better captures the TMD functional form constructed with MAPTMD22 than MAPTMD22 does for the TMD constructed using PV19. Based on the L0 test, it appears that the parameterization of the TMD PDF in PV19 (equation 3.16) is more flexible than that of MAPTMD22 (equation 3.12). However, the conclusions differ when we consider the results of the multi-closure test, which will be discussed below.

This result indicates that both frameworks are reliable only under the assumption

that the true TMD functional forms we aim to extract closely resemble those that can be constructed with our parameterizations. The issue with this assumption is that parameterizations with a low number of parameters, like those used in MAPTMD22 and PV19, can only reproduce a limited range of functional forms. Moreover, we have limited physical constraints on the non-perturbative TMD functional forms. That means, we do not know which functional form the true TMD should follow, which raises the risk that the true TMD functional form may lie outside the class of functional forms that can be reproduced using our parameterization. This issue is not exclusive to the MAPTMD22 and PV19 frameworks but is a general challenge faced by traditional fitting techniques for extracting TMDs from experimental data.

One possible solution to this problem is to represent the TMD functional forms using a large number of parameters via neural networks. This approach has already been successfully applied to determine the non-perturbative part of collinear PDFs by the NNPDF collaboration [12], effectively addressing the flexibility issue, since a neural network is able to explore a much larger region in the space of the TMDs functional forms than a traditional parametrization. The neural networks approach to the problem of extracting TMDs from experimental data has not been attempted yet.

Conclusions on the Uncertainty Contributions

Passing to the L1 and L2 tests, we operated under the assumption that the *real* non-perturbative TMD functional form lies within the space of forms that our parameterizations can capture. We indeed used the same parameterizations to both generate and fit the datasets, specifically testing MAP22oMAP22 and PV19oPV19. When analyzing the uncertainty contributions, we found that in MAPTMD22, the dominant sources of uncertainty are the functional and experimental uncertainties, while the interpolation/extrapolation uncertainty is suppressed. In the PV19 framework, both the functional and interpolation/extrapolation uncertainties are suppressed, with almost all the uncertainty coming from the experimental data. The conclusion which we can draw for both frameworks is that improving the predictivity of the methodologies requires more precise data. Increasing the number of data points would not significantly reduce the total uncertainty, as the interpolation/extrapolation contribution—coming from regions where no data points are available—is already subdominant in both frameworks.

Conclusions on the Multi Closure Tests

The multi-closure test allowed us to determine whether the frameworks are optimized in terms of the bias-variance tradeoff and whether the declared uncertainties faithfully represent the true TMD uncertainties. For the MAPTMD22 framework, the results from the multi closure test were satisfactory (equations (5.22), (5.23)), indicating that the model is well-optimized and reports accurate uncertainties.

In contrast, the PV19 framework showed signs of underfitting when we calculated the $\xi_{1\sigma}$ and $\mathcal{R}_{bv}$ estimators (equations (5.24), (5.25)), which prove we have high bias compared to the variance of the predicted TMDs. The underfitting issue in PV19 leads to a biased declaration of the TMD uncertainties, which are underestimated. A confidence level declared to be 68% is, in reality, a 48% confidence level. Consequently, the TMD uncertainties reported by the model need to be rescaled to reflect a more accurate confidence level. The underfitting problem in PV19 may appear contradictory, given that the PV19oMAP22_{L0} test performed better than the MAP22oPV19_{L0} test (see table 5.2). This result suggested that the PV19 parameterization was more flexible than the MAPTMD22 one. However, the underfitting issue points to a lack of flexibility in the PV19 parameterization. To resolve this apparent paradox, we propose that the MAPTMD22 TMD PDF functional form used to generate the dataset for the MAP22oPV19_{L0} test is closer to the space of forms that the PV19 parameterization can capture. On the other hand, the PV19 TMD PDF used to generate the dataset for the MAP22oPV19_{L0} test may be farther from the space of forms that the MAPTMD22 parameterization can reproduce. Thus, while PV19 performed better than MAPTMD22 in a cross-test, this does not necessarily indicate that PV19 is generally more flexible. However, this assumption has yet to be tested. It would be interesting to explore the size of the functional space that both the PV19 and MAPTMD22 parameterizations can capture, as this would provide a more definitive comparison of their flexibility.

6.2 Future Applications of This Work

The implementation and execution of closure testing on the PV19 and MAPTMD22 methodologies do not conclude with these two specific tests. A significant result of this work is the potential to apply closure testing to future fitting methodologies and enhance them based on these statistical tests. Closure testing is a crucial tool for validating and assessing the robustness of a framework and should be considered an integral part of the methodology. In chapter 3, we discussed how the frameworks developed by the MAP Collaboration (including PV19 and MAPTMD22) address the extraction of TMDs from experimental data, the process is summarized in figure 3.2. The implementation of closure testing in a fitting framework involves adding a step to the process: the "validation-optimization" step. This step means testing the methodology on a pseudo-dataset generated from a known input and fine-tuning the framework until it can accurately reproduce both the underlying *true* law (the input used to generate the L0 dataset), and the associated TMD uncertainties. Only after optimizing the methodology based on closure test results should it be used to fit real datasets.

Thanks to this work, the validation-optimization step has been partially integrated into the Nanga Parbat¹ suite. Specifically, the processes for generating Level 0, Level 1, and Level 2 pseudo-data can be reused to test other frameworks. When a new

¹Nanga Parbat is the software developed by the MAP Collaboration, designed to run fits for TMD determination using various fitting frameworks, including different parameterizations and optimization algorithms.

framework with its own parameterization is used within Nanga Parbat to fit real data, it produces data points predicted by the TMDs constructed from that parameterization. These correspond to the Level 0 (non-fluctuated) data points. The process of constructing the Level 0 dataset from these predictions is straightforward and can be executed using the same code, regardless of the specific parameterization. Similarly, the code for data fluctuation, which generates the Level 1 and Level 2 datasets, is also framework-independent, *i.e.* does not depend on the parametrization used to generate the dataset. This means it can be applied to produce new pseudo-datasets for testing future frameworks. Additionally, the estimators calculated through a multi-closure test, $\xi_{1\sigma}$ and $\mathcal{R}_{bv}$, are defined in the data-space rather than in the space of TMDs (see section 4.5). This allows us to compute $\xi_{1\sigma}$ and $\mathcal{R}_{bv}$ independently of the framework being tested and reuse the same code developed for testing the MAPTMD22 and PV19 frameworks on future methodologies. The parts of this work which are not directly transferable to other frameworks are the parametrization-dependent components. These include the uncertainty characterization (section 5.3) and the calculation of the Mean Squared Distance estimator (equation 5.4), both of which require specific TMD functional forms as input.

Integrating closure testing into future frameworks will ensure that the methodologies used to extract non-perturbative TMDs from experimental data are both validated and optimized, leading to more reliable determinations of TMDs. The reliability of the predicted TMDs is crucial to further develop our knowledge of non-perturbative QCD, test our current understanding of the standard model and search for new physics beyond it.

List of Figures

2.1	Energy dependence of the strong coupling $\alpha_s(\mu)$, image from Ref. [9]. . .	7
2.2	In DIS, a high-energy electron interacts with partons inside a proton, causing the proton to break up. Only the final-state electron is measured. . .	8
2.3	In light-front coordinates, the axes Y^+ and Y^- are parallel to the light cone. In these plots, the particles are approximated to be massless. . .	10
2.4	PDF set for the proton extracted from a global dataset using NNPDF4.0 . . .	13
2.5	The handbag diagram represent the hadronic part of the interaction, that is the scattering between a virtual photon and a single constituent of the hadron.	14
2.6	Initial and final particles involved in SIDIS in 3D space	17
2.7	Transverse momenta in the hadronic part of SIDIS, image from [4] . . .	18
2.8	Initial and final particles involved in Drell-Yan in 3D space	20
2.9	Diagrammatic description of the Drell-Yan short-range interaction . . .	21
2.10	Relevant Drell-Yan momenta	21
2.11	Functional form of b_* at $Q = 3.2$ GeV	25
3.1	Kinematic coverage in (x, Q^2) space of the dataset used in MAPTMD22, image from [4]	31
3.2	Flowchart of the optimization process in the analyzed fitting frameworks. In this chart, the model $\mathcal{G}$ is represented by the label "QCD theory", the input $u^{(n)}$ is represented by the "TMD parametrization", and the predicted dataset $\mathcal{G}(u^{(n)})$ is represented by "theoretical predictions". . .	37
4.1	L0 data for a physical observable $\frac{d\sigma}{d q_T }$. The points are perfectly distributed on the theoretical functional form of the observable.	41
4.2	L1 data for a physical observable $\frac{d\sigma}{d q_T }$ (in yellow). The white points represent the L0 data, the yellow points are the fluctuated data.	42
4.3	L2 data for $\frac{d\sigma}{d q_T }$. The L2 dataset is composed of N_{rep} Monte Carlo replicas of each L1 data-point.	43
4.4	Interpolation/extrapolation uncertainty. This uncertainty is associated with the variability of the functional form in the intervals between data points, as well as in kinematic regions not covered by the dataset. . . .	46
4.5	Functional uncertainty. This uncertainty is associated to the fluctuations of the dataset and is given by the spread of a large number of predictions. All the predicted functional forms approximate equivalently well the L1 data distribution.	47

LIST OF FIGURES

4.6	The spread of the set of predictions $\{u_*\}$ in a L2 test. This represent the total uncertainty.	48
4.7	The behavior of bias and variance depending on the model complexity and the distribution of the predictions of a data point depending on the values of bias and variance. Image from Ref. [46]	52
5.1	Real and predicted $f_{NP}(x = 0.1, b_T, \zeta = 100 \text{ GeV}^2)$ in MAP22oPV19 _{L0}	66
5.2	Real and predicted $f_{NP}(x = 0.1, b_T, \zeta = 4 \text{ GeV}^2)$ in MAP22oPV19 _{L0}	67
5.3	χ^2 distribution in MAP22oPV19 _{L0}	67
5.4	χ^2 distribution in the PV19oMAP22 _{L0} test.	68
5.5	Real and predicted $f_{NP}(x = 0.1, b_T, \zeta = 100 \text{ GeV}^2)$ in the PV19oMAP22 _{L0} test.	69
5.6	χ^2 distribution in the MAP22oMIX24 _{L0} test.	71
5.7	Real and predicted $f_{NP}(x = 0.1, b_T, \zeta = 100 \text{ GeV}^2)$ in the MAP22oMIX24 _{L0} test.	71
5.8	Plots at $Q = 10 \text{ GeV}$, $x = 0.1$	75
5.9	Plots at $Q = 10 \text{ GeV}$, $x = 0.01$	76
5.10	Plots at $Q = 10 \text{ GeV}$, $x = 0.001$	77
5.11	Plots at $Q = 10 \text{ GeV}$, $z = 0.3$	78
5.12	Plots at $Q = 10 \text{ GeV}$, $z = 0.6$	79
5.13	Plots at $Q = 2 \text{ GeV}$, $x = 0.1$	80
5.14	Plots at $Q = 2 \text{ GeV}$, $x = 0.01$	81
5.15	Plots at $Q = 2 \text{ GeV}$, $x = 0.001$	82
5.16	Plots at $Q = 2 \text{ GeV}$, $z = 0.3$	83
5.17	Plots at $Q = 10 \text{ GeV}$, $z = 0.6$	84
5.18	Real and predicted $f_{NP}(x = 0.1, b_T, \zeta = 10^4 \text{ GeV}^2)$ in MAPTMD22. This is the Q-region of the LHC data, the fact that the total uncertainty is small compared to the other plots confirms that the uncertainty decreases in the kinematic regions where the data are more precise.	85
5.19	Uncertainty ratios $\hat{\sigma}_{L0}$ and $\hat{\sigma}_{L1}$ for the non-perturbative TMD PDF at $Q = 10 \text{ GeV}$	86
5.20	Uncertainty ratios $\hat{\sigma}_{L0}$ and $\hat{\sigma}_{L1}$ for the non-perturbative TMD FF at $Q = 10 \text{ GeV}$	86
5.21	Uncertainty ratios $\hat{\sigma}_{L0}$ and $\hat{\sigma}_{L1}$ for the non-perturbative TMD PDF at $Q = 2 \text{ GeV}$	87
5.22	Uncertainty ratios $\hat{\sigma}_{L0}$ and $\hat{\sigma}_{L1}$ for the non-perturbative TMD FF at $Q = 2 \text{ GeV}$	88
5.23	Plots at $Q = 10 \text{ GeV}$, $x = 0.1$	89
5.24	Plots at $Q = 10 \text{ GeV}$, $x = 0.01$	90
5.25	Plots at $Q = 10 \text{ GeV}$, $x = 0.001$	91
5.26	Plots at $Q = 2 \text{ GeV}$, $x = 0.1$	92
5.27	Plots at $Q = 2 \text{ GeV}$, $x = 0.01$	93
5.28	Plots at $Q = 2 \text{ GeV}$, $x = 0.001$	94
5.29	Uncertainty ratios $\hat{\sigma}_{L0}$ and $\hat{\sigma}_{L1}$ for the non-perturbative TMD PDF at $Q = 10 \text{ GeV}$	96

LIST OF FIGURES

5.30	Uncertainty ratios $\hat{\sigma}_{L0}$ and $\hat{\sigma}_{L1}$ for the non-perturbative TMD PDF at $Q = 2$ GeV	97
------	--	----

List of Tables

3.1	Drell-Yan measurements included in PV19, table from [3]	29
3.2	Drell-Yan measurements included in MAPTMD22, table from [4]	30
3.3	SIDIS measurements included in MAPTMD22, table from [4]	31
3.4	Levels of accuracy in the models, depending on the truncation orders in the perturbative content of the factorized cross sections. The accuracy reached by $\mathcal{G}_{\text{MAP22, DY}}$ and $\mathcal{G}_{\text{PV19}}$ is N^3LL ; the accuracy of $\mathcal{G}_{\text{MAP22, SIDIS}}$ is N^3LL^-	33
3.5	Inputs parametrization of the MAPTMD22 and PV19 frameworks.	35
4.1	Levels of data fluctuations in closure testing	43
4.2	Uncertainty layers introduced by each level of fluctuation.	49
4.3	Input parametrizations used to generate three different set of L0 pseudo-data.	54
4.4	Values of the parameters used as input to generate the pseudodata.	55
4.5	Models and inputs used to generate the three L0 pseudo-datasets $\mathcal{D}_{L0}^{\text{MAP22}}$, $\mathcal{D}_{L0}^{\text{PV19}}$, and $\mathcal{D}_{L0}^{\text{MIX24}}$	56
4.6	Uncertainties associated to the i -th data point of the experimental dataset $\mathcal{D}$, these uncertainties are transferred to the i -th point of the L0, L1 and L2 pseudo-datasets.	57
4.7	Table including all the performed L0 closure tests.	59
4.8	Table including all the performed L0, L1, and L2 closure tests for the uncertainty characterization of MAPTMD22 and PV19.	60
4.9	Table including the performed multi closure tests. We performed each test on a total of 5000 L2 dataset (100 L2 fluctuations starting from 50 different L1 datasets).	60
5.1	Ceres-Minuit comparison.	63
5.2	Results of the tests performed on the L0 datasets.	72
5.3	Plots of the uncertainties of the TMD PDFs and FFs predicted by MAPTMD22 in different kinematic regions.	73
5.4	Plots of the uncertainty ratios $\hat{\sigma}_{L0}$ and $\hat{\sigma}_{L1}$ of the TMD PDFs and FFs predicted by MAPTMD22 in different kinematic regions.	85
5.5	Plots of the uncertainties of the TMD PDFs predicted by PV19 in different kinematic regions.	95
5.6	Plots of the uncertainty ratios $\hat{\sigma}_{L0}$ and $\hat{\sigma}_{L1}$ of the TMD PDFs predicted by PV19 in different kinematic regions.	95

Bibliography

- [1] J. C. Collins, “What exactly is a parton density?,” *Acta Phys. Polon. B*, vol. 34, p. 3103, 2003.
- [2] J. Collins, *Foundations of Perturbative QCD*, vol. 32 of *Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology*. Cambridge University Press, 7 2023.
- [3] A. Bacchetta, V. Bertone, C. Bissolotti, G. Bozzi, F. Delcarro, F. Piacenza, and M. Radici, “Transverse-momentum-dependent parton distributions up to N³LL from Drell-Yan data,” *JHEP*, vol. 07, p. 117, 2020.
- [4] A. Bacchetta *et al.*, “Unpolarized transverse momentum distributions from a global fit of Drell-Yan and semi-inclusive deep-inelastic scattering data,” *JHEP*, vol. 10, p. 127, 2022.
- [5] L. Del Debbio *et al.*, “Bayesian approach to inverse problems: an application to NNPDF closure testing,” *Eur. Phys. J. C*, vol. 82, no. 4, p. 330, 2022.
- [6] R. D. Ball *et al.*, “Parton distributions for the LHC Run II,” *JHEP*, vol. 04, p. 040, 2015.
- [7] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley, 1995.
- [8] B. Grinstein, “Introductory lectures on QCD,” in *2nd CERN-CLAF School of High Energy Physics*, pp. 27–56, 6 2003.
- [9] S. Navas *et al.*, “Review of particle physics,” *Phys. Rev. D*, vol. 110, no. 3, p. 030001, 2024.
- [10] R. E. Taylor, “The Discovery of the point - like structure of matter,” *Phil. Trans. Roy. Soc. Lond. A*, vol. 359, pp. 225–240, 2001.
- [11] R. Brock *et al.*, “Handbook of perturbative QCD; Version 1.1: September 1994,” 9 1994.
- [12] R. D. Ball *et al.*, “The path to proton structure at 1% accuracy,” *Eur. Phys. J. C*, vol. 82, no. 5, p. 428, 2022.

BIBLIOGRAPHY

- [13] T.-J. Hou, J. Gao, T. J. Hobbs, K. Xie, S. Dulat, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, I. Sitiwaldi, D. Stump, and C.-P. Yuan, “New cteq global analysis of quantum chromodynamics with high-precision data from the lhc,” *Phys. Rev. D*, vol. 103, p. 014013, Jan 2021.
- [14] Bailey, S., Cridge, T., Harland-Lang, L. A., Martin, A. D., and Thorne, R. S., “Parton distributions from lhc, hera, tevatron and fixed target data: Msht20 pdfs,” *Eur. Phys. J. C*, vol. 81, no. 4, p. 341, 2021.
- [15] Z. Zhang, “Hera inclusive neutral and charged current cross sections and a new pdf fit, herapdf 2.0,” 2015.
- [16] A. Accardi, X. Jing, J. F. Owens, and S. Park, “Light quark and antiquark constraints from new electroweak data,” *Phys. Rev. D*, vol. 107, p. 113005, Jun 2023.
- [17] T. C. Rogers, “An overview of transverse-momentum-dependent factorization and evolution,” *Eur. Phys. J. A*, vol. 52, no. 6, p. 153, 2016.
- [18] I. Scimemi, “A short review on recent developments in TMD factorization and implementation,” *Adv. High Energy Phys.*, vol. 2019, p. 3142510, 2019.
- [19] J. Collins, “Overview of TMD Factorization and Evolution,” *Int. J. Mod. Phys. Conf. Ser.*, vol. 37, p. 1560021, 2015.
- [20] A. Bacchetta, “Where do we stand with a 3-D picture of the proton?,” *Eur. Phys. J. A*, vol. 52, no. 6, p. 163, 2016.
- [21] A. Bacchetta, M. Diehl, K. Goeke, A. Metz, P. J. Mulders, and M. Schlegel, “Semi-inclusive deep inelastic scattering at small transverse momentum,” *JHEP*, vol. 02, p. 093, 2007.
- [22] A. Signori, *Flavor and Evolution Effects in TMD Phenomenology: Manifestation of Hadron Structure in High-Energy Scattering Processes*. PhD thesis, Vrije U., Amsterdam, 2016.
- [23] G. Moreno *et al.*, “Dimuon Production in Proton - Copper Collisions at $\sqrt{s} = 38.8\text{-GeV}$,” *Phys. Rev. D*, vol. 43, pp. 2815–2836, 1991.
- [24] A. S. Ito *et al.*, “Measurement of the Continuum of Dimuons Produced in High-Energy Proton - Nucleus Collisions,” *Phys. Rev. D*, vol. 23, pp. 604–633, 1981.
- [25] 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 $\sqrt{s} = 1.8\text{ TeV}$,” *Phys. Rev. Lett.*, vol. 84, pp. 845–850, 2000.
- [26] T. Aaltonen *et al.*, “Transverse momentum cross section of e^+e^- pairs in the Z -boson region from $p\bar{p}$ collisions at $\sqrt{s} = 1.96\text{ TeV}$,” *Phys. Rev. D*, vol. 86, p. 052010, 2012.

- [27] B. Abbott *et al.*, “Measurement of the inclusive differential cross section for Z bosons as a function of transverse momentum in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV,” *Phys. Rev. D*, vol. 61, p. 032004, 2000.
- [28] V. M. Abazov *et al.*, “Measurement of the shape of the boson transverse momentum distribution in $p\bar{p} \rightarrow Z/\gamma^* \rightarrow e^+e^- + X$ events produced at $\sqrt{s}=1.96$ -TeV,” *Phys. Rev. Lett.*, vol. 100, p. 102002, 2008.
- [29] V. M. Abazov *et al.*, “Measurement of the Normalized $Z/\gamma^* \rightarrow \mu^+\mu^-$ Transverse Momentum Distribution in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV,” *Phys. Lett. B*, vol. 693, pp. 522–530, 2010.
- [30] R. Aaij *et al.*, “Measurement of the forward Z boson production cross-section in pp collisions at $\sqrt{s} = 7$ TeV,” *JHEP*, vol. 08, p. 039, 2015.
- [31] R. Aaij *et al.*, “Measurement of forward W and Z boson production in pp collisions at $\sqrt{s} = 8$ TeV,” *JHEP*, vol. 01, p. 155, 2016.
- [32] R. Aaij *et al.*, “Measurement of the forward Z boson production cross-section in pp collisions at $\sqrt{s} = 13$ TeV,” *JHEP*, vol. 09, p. 136, 2016.
- [33] S. Chatrchyan *et al.*, “Measurement of the Rapidity and Transverse Momentum Distributions of Z Bosons in pp Collisions at $\sqrt{s} = 7$ TeV,” *Phys. Rev. D*, vol. 85, p. 032002, 2012.
- [34] V. Khachatryan *et al.*, “Measurement of the transverse momentum spectra of weak vector bosons produced in proton-proton collisions at $\sqrt{s} = 8$ TeV,” *JHEP*, vol. 02, p. 096, 2017.
- [35] G. Aad *et al.*, “Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector,” *JHEP*, vol. 09, p. 145, 2014.
- [36] G. 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,” *Eur. Phys. J. C*, vol. 76, no. 5, p. 291, 2016.
- [37] P. L. McGaughey *et al.*, “Cross-sections for the production of high mass muon pairs from 800-GeV proton bombardment of H-2,” *Phys. Rev. D*, vol. 50, pp. 3038–3045, 1994. [Erratum: *Phys.Rev.D* 60, 119903 (1999)].
- [38] 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,” *Phys. Rev. D*, vol. 99, no. 7, p. 072003, 2019.
- [39] A. M. Sirunyan *et al.*, “Measurements of differential Z boson production cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *JHEP*, vol. 12, p. 061, 2019.
- [40] G. 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,” *Eur. Phys. J. C*, vol. 80, no. 7, p. 616, 2020.

BIBLIOGRAPHY

- [41] A. Airapetian *et al.*, “Multiplicities of charged pions and kaons from semi-inclusive deep-inelastic scattering by the proton and the deuteron,” *Phys. Rev. D*, vol. 87, p. 074029, 2013.
- [42] M. Aghasyan *et al.*, “Transverse-momentum-dependent Multiplicities of Charged Hadrons in Muon-Deuteron Deep Inelastic Scattering,” *Phys. Rev. D*, vol. 97, no. 3, p. 032006, 2018.
- [43] J. Nocedal and S. Wright, *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering, Springer New York, 2006.
- [44] S. Agarwal, K. Mierle, and T. C. S. Team, “Ceres Solver,” 10 2023.
- [45] F. a. James, “MINUIT: Function Minimization and Error Analysis Reference Manual,” 1998. CERN Program Library Long Writeups.
- [46] I. Reznikov, “Stop using the same image in bias-variance trade-off explanation.” Medium, november 2021. <https://medium.com/@ivanreznikov/stop-using-the-same-image-in-bias-variance-trade-off-explanation-691997a94a54>.