UNIVERSITÀ DEGLI STUDI DI PAVIA
DOTTORATO DI RICERCA IN FISICA - XXXVI CICLO

Revealing the internal structure of the pion: a journey in phenomenology

Simone Venturini

Tesi per il conseguimento del titolo

Università degli Studi di Pavia
Dipartimento di Fisica

DOTTORATO DI RICERCA IN FISICA - XXXVI CICLO

Revealing the internal structure of the pion: a journey in phenomenology

Simone Venturini

Submitted to the Graduate School of Physics in partial fulfilment of the requirements for the degree of
DOTTORE DI RICERCA IN FISICA
DOCTOR OF PHILOSOPHY IN PHYSICS
at the
University of Pavia

Supervisor: Prof. Barbara Pasquini

Cover: *https://pixabay.com/it/illustrations/vetro-mare-cristallo-blu-sfera-7652690/*

Pion structure: a path towards phenomenology
Simone Venturini
PhD thesis - University of Pavia
Pavia, Italy, September 2023

Alla mia famiglia,
dopo 8 anni di studio,
dedico con affetto quest'ultimo tassello.

Contents

Chapter 1

Introduction

Understanding the fundamental building blocks of matter is essential for addressing crucial questions like: *What is the origin of the visible matter? How did the primordial matter evolve? How can the subatomic matter influence the behaviour of atoms? How many fundamental interactions exist, and are they comprehensively understood? How might this knowledge prove beneficial for practical applications, ultimately benefiting society?*

At present we are far from giving definitive answers to these questions; however, significant progress has been made in the past decades, especially in the field of nuclear physics [1]. The discovery of the first hadron, the proton, by Lord Rutherford [2] dates back to more than one century ago. The second hadron, the neutron, was detected by James Chadwick [3] in 1932.

Starting from the mid-nineteenth century, a multitude of other strongly interacting particles were discovered and physicists began to refer to them as a "particle zoo".

A first attempt to shed light and to explain the presence of various hadrons is represented by the *quark model*, introduced by Murray Gell-Mann. The basic assumption was the existence of just three different *quarks*, that combined together could give rise to various hadrons. The quark model led to predict the existence of new hadrons, and Gell-Mann received the Nobel Prize in Physics in 1969 *for his contributions and discoveries concerning the classification of elementary particles and their interactions.*

Physicists began to investigate how the elementary particles within hadrons could influence the properties of the hadrons themselves. The quark model soon proved insufficient in explaining some very fundamental properties of hadrons. This led to introduce other elementary particles: the *gluons*, i.e. the *mediators* of the strong interaction. Still at present, the fundamental constituents of hadrons are believed to be gluons, quarks and anti-quarks, collectively called *partons*.

Among the fundamental questions regarding the hadron properties, two problems have represented a real puzzle in the past decades: the mass and the spin of the hadrons. Unlike elementary particles, hadrons, being composite

particles, do not acquire mass through the Higgs mechanism. *Where does the mass of the hadrons come from?* One first answer came by identifying the "valence quarks" within a hadron, which are the minimal set of quarks required to account for all its quantum properties. Since quarks are elementary particles and derive their mass from the Higgs mechanism, one might - naively - expect that the problem is solved, by assuming that the mass of a hadron is the sum of the masses of its valence quarks. However, through phenomenological investigations, physicists proved that this assumption is incorrect, as the sum of the masses of the valence quarks accounts for only a minuscule fraction of the total hadron mass. A similar conundrum emerges regarding the hadron spin. *Where does the spin of the hadrons come from?* Again, the hypothesis to sum the contributions of the single spin of valence quarks revealed to be not correct: the 1987 Electron Muon Collider (EMC) measurement [4] proved that approximately only the 30% of the proton spin came from the spins of its valence quarks. This result ushered in the "spin crisis" period.

It was only with quantum chromodynamics (QCD) that physicists started to shed light on the matter. Basically, in QCD a hadron is thought to be constituted by an undefined number of partons that globally contribute to explain the quantum numbers of the parent particle. The internal structure of a hadron can be imagined as a "bubbling of energy" exchanged by virtual gluons and pairs of quarks and antiquarks arising from the QCD vacuum. The sum of this energy contribution and the masses of valence quarks constitutes the mass of the parent particle. Concerning the hadrons' spin the partons bring an intrinsic spin and they also contribute through their motion inside the parent hadron.

These issues have been further investigated by means of the parton distribution functions, representing the probability distribution of finding inside a hadron a parton endowed with specific characteristics. During the last decades, a great effort has been made for the investigation and the phenomenology-driven determination of the parton distribution functions. Their study is complicated by the non-perturbative nature of QCD, which drives to the development of theoretical models that can be fitted to experimental data, or used to obtain theoretical predictions.

The definitive and complete knowledge of the parton distribution functions would help answer the already mentioned spin and mass problems and a large number of other fundamental questions, for example: *What is the spatial distribution of partons inside the hadrons? What is the amount of momentum carried by the partons? What are the differences between a polarized or an unpolarized hadron? How does the spatial charge distribution of the partons change in a fast-moving hadron?*

All these questions have been addressed in recent years, for various hadrons, and a lot progress has been made, also thanks to the recent advances in computation power that allowed for more and more accurate lattice calculations. Some of them are near to be universally answered; other need to further work

and still remain among the main challenges of hadronic physics.

Another open question is whether there are additional particles yet to be discovered and - if this is the case - at what energy they can be accessed. Recent studies predict that due to gluons interacting with each other there may exist new undiscovered hadrons in which gluons play the same central role as the valence quarks. The frontier of these novel exotic particles is planned to be experimentally investigated, using a new cutting-edge instrumentation that could give rise to the birth of a new branch in sub-nuclear physics.

The hadrons considered in this work are the pions, which are hadrons existing in three different charges, π^+, π^- and π^0, detected in the 1950s [5, 6]. The ultimate goal of this study is to find a proper theoretical model to parametrize different parton distribution functions of these particles. The role of pions in the Standard Model (SM) is unique: primarily, they are the lightest mesons in nature, and therefore their structure is thought to be relatively simpler than that of other hadrons; secondly, they can be identified as Nambu–Goldstone pseudo-scalar bosons, associated with the spontaneous symmetry breaking of the axial-vector symmetries of the strong interaction Lagrangian; in the end, virtual pions are among the particles that are exchanged in nuclear interactions. It is therefore evident that a thorough understanding of these particles is crucial for investigating some very basic features of the SM, like the colour confinement and the dynamical chiral symmetry breaking.

The theoretical model we propose for the pion parton distribution functions serves a dual purpose. Firstly, it is used to fit existing experimental data from various semi-inclusive, inclusive and exclusive deep inelastic scattering processes. Secondly, this model enables us to make predictions about observables for future experiments.

Experimental data play a crucial role in testing the validity and efficacy of theoretical models. Conversely, theoretical models often guide the direction for future experiments. Therefore, delving into the study of the pion structure leads us down a concrete *journey in phenomenology*, bridging the gap between theory and experiment.

1.1 Content of this thesis

The theoretical model we propose to parametrize the pion parton distribution functions exploits an approach based on the Light-Front Wave Functions (LFWFs). These objects can be identified as the coefficients appearing in the N-parton Fock state expansion of the pion state and can be interpreted in an intuitive manner in the light-front quantization formalism. We include this discussion in Chapter 2, where we also introduce the mathematical definition of the parton distribution functions and the theoretical reasoning at the basis of our parametrization. The crux of the model is the presence of two distinct sets of parameters, which can be independently fitted through separate processes. The first set can be fitted to the observables related to the collinear

Parton Distribution Functions (PDFs); the second set can be - consequently - fitted to observables that are sensitive to the transverse momentum direction.

Chapter 3 is devoted to the fit of pion PDFs. In particular, we discuss all the technical details regarding the fit of the collinear set of parameters and we compare our results with other existing extractions and theoretical models.

Once the first set of parameters has been determined through the collinear fit of PDFs, the second set of parameters can be fitted in two different ways: either by fitting the data of pion electromagnetic Form Factor (FF), or the observables sensitive to the pion transverse-momentum dependent parton distributions (TMDs).

Chapter 4 concerns the parametrization and the fit of pion electromagnetic FF. As well as for the PDFs, the error bars are obtained from the replica method. The only caveat is to take into account the propagation of the errors of the collinear parameters in the fit of the FF. In light of the multiple data sets with systematic errors used for our extractions, we insert a detailed discussion on the statistical method employed in the analysis.

Chapter 5 focuses on the pion TMDs. These distributions offer an alternative approach to the FF in constraining the second set of parameters in our model. We show the fit of our model to the available experimental data and then compare the results for the fitted parameters with those coming from the fit of FF. This comparison provides a valuable consistency check, reflecting the versatility and robustness of our model in describing different aspects of the pion's internal structure. Finally, we present a fit of pion TMDs that has been performed by the MAP[1] collaboration by exploiting the Nanga Parbat [2] fitting software and we compare the two extractions.

In Chapter 6 we introduce the pion Generalized Parton Distribution Functions (GPDs) and discuss their theoretical properties. The two sets of parameters, coming from the fits of the pion PDFs and the pion FF, are used to obtain predictions for the pion GPDs. We will evolve our results from the model scale to the relevant hard scales that can be investigated experimentally, paving the way for future phenomenological applications.

In Chapter 7 we sum up the results obtained so far and we discuss the next steps to take on the *journey in phenomenology*.

[1] Multi-dimensional Analyses of Partonic distributions
[2] https://github.com/MapCollaboration/NangaParbat.

Chapter 2

Foundations

This Chapter contains the theoretical foundations at the basis of the thesis project. The starting point is the light front quantization, which provides a useful alternative to the canonical quantization of fields for the description of hadronic processes. We then present a very general introduction to parton distribution functions and hadron distribution amplitudes, which represent fundamental objects for the description of the internal structure of the hadrons in terms of their constituents. Finally, we delineate the essential procedures for building a theoretical model capable of parametrizing the different parton distribution functions of the pion.

2.1 Light Front Quantization

In any Quantum Field Theory (QFT), the first step to take is to specify the geometry of the spacetime in which the theory operates. Throughout this discussion, we consider the flat four-dimensional Minkowski spacetime. This is the geometry that is expected to hold when considering the world of subatomic particles, in the absence of heavy masses in the surrounding environment, which would otherwise introduce curvature effects.

In a four-dimensional spacetime, the controvariant coordinates of a four-vector $x^\mu = (x^0, x^1, x^2, x^3)$ and its covariant coordinates $x_\mu = (x_0, x_1, x_2, x_3)$ are linked by the metric tensor $g_{\mu\nu}$:

$$x_\mu = g_{\mu\nu} x^\nu. \tag{2.1}$$

For any pair of points x and y in the spacetime, the metric tensor induces a scalar product:

$$g(x, y) = x^\mu g_{\mu\nu} y^\nu = x^\mu y_\mu, \tag{2.2}$$

and this allows one to identify a relativistic invariant, i.e. a quantity that does not change under Lorentz transformations. Such a quantity is the spacetime interval ds, defined in terms of the spacetime distance $dx^\mu = x^\mu - y^\mu$, as

$$ds^2 = dx^\mu dx_\mu. \tag{2.3}$$

In a Minkowski space with signature $(+,-,-,-)$ the metric tensor is given by

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{2.4}$$

which is intrinsically related to the geometrical interpretation of the light-cone.

Since the speed of light c is a universal constant, it is possible to identify the 0-component of the four-vectors as the time coordinate by introducing a one-to-one correspondence $x^0 = c \cdot t$, and the spacetime interval of Eq. (2.3) becomes:

$$ds^2 = (c \cdot dt)^2 - {dx_1}^2 - {dx_2}^2 - {dx_3}^2. \tag{2.5}$$

2.1.1 The geometry of the light-cone

Let us suppose to be an observer in the origin of the four dimensional Minkowski spacetime[1]. By considering the spacetime distance between us and any other point, we can characterize the spacetime by distinguishing from three different geometrical *regions*:

$$\begin{cases} ds^2 > 0 \ (\textit{first region}); \\ ds^2 = 0 \ (\textit{second region}): \\ ds^2 < 0 \ (\textit{third region}). \end{cases} \tag{2.6}$$

Let us consider, for the moment, only those points characterized by $x^0 \geq 0$, i.e. the same as our time coordinate or living in "the future".
The spacetime points in the *first region* are in the *chronological future* of the observer, which means that it exists a future-directed chronological curve connecting the origin to them, or, more simply, it is possible to reach them with a finite velocity $v < c$. The four-vectors representing these points are referred as to *timelike*. The *second region* is the locus of the *lightlike* four-vectors. In other words, we can connect the origin of the axes to these points only with a light signal with $v = c$. The first and the second regions together constitute our *causal future*, i.e. the region of the spacetime that we can explore without violating causality. The *third region* is characteristic of Lorentzian manifolds, such as Minkowski spacetime, distinguishing them from Riemannian manifolds like Euclidean space. This region is relativistically prohibited, since it can only be explored from the origin by travelling with velocities $v > c$, thus violating causality. The four-vectors within this region are referred to as *spacelike*.

Regarding the subspace with $x^0 < 0$, the situation is similar, but this time we classify a point as belonging to the *chronological past* of the origin if $ds^2 > 0$, and to the *causal past* if $ds^2 \geq 0$.

[1] If we were not in this condition, it is always possible to apply a spacetime translation and reduce to it.

By merging the two spacetime regions $x^0 \geq 0$ and $x^0 < 0$, we can derive the geometrical structure of the Minkowski spacetime. The points separating the origin by lightlike four-vectors reside on a 4-dimensional cone: the light cone (see Fig. 2.1); the chronological past and the chronological future are contained within the walls of the cone, while the *present*, i.e. the ensemble of points with $x^0 = 0$, is a 3D hyperplane. All the other spacetime points are causally disconnected from the origin. We can observe that all the points of the present-hyperplane are spacelike, except for the origin itself[2].

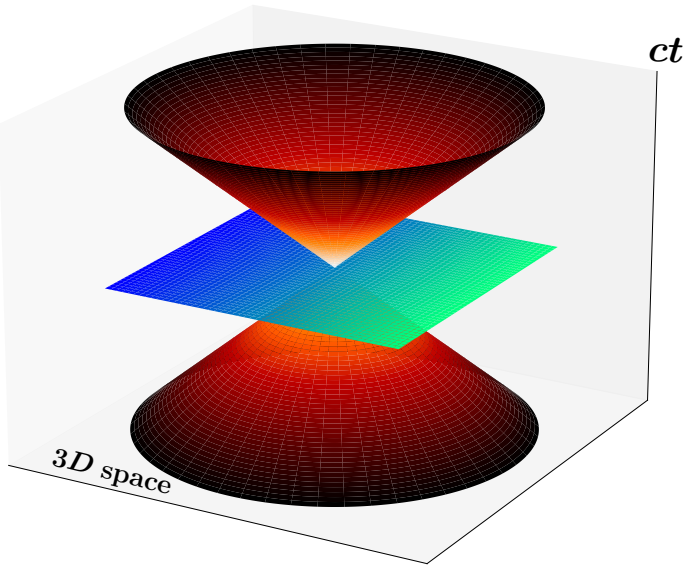


Figure 2.1: The geometry of the Minkowski spacetime in 2+1 dimensions, where the time coordinate[3] is supposed to increase upwards (one spatial dimension has been neglected to reproduce the 3D figure). With respect to the observer placed in the center of the figure, the orange light-cone contains all the lightlike four-vectors; the points inside the walls of the cone are the timelike four-vectors; all the other four-vectors are spacelike, and the blue-green plane represents the present-hyperplane.

It is a well-known fact that in Physics it is often convenient to change the reference frame, without affecting the final results. We want now to focus on the following problems: we saw that the structure of the light cone is imposed by the geometrical properties of the spacetime, and by the metric definition, but *what does it change if we vary the parametrization of the spacetime? In how many independent ways can this be done?* These questions lead us to introduce the so-called "forms of relativistic dynamics".

[2]This is the reason for which an observer cannot be in two distinguished space points at the same time.

[3]This figure, as well as the next ones with the light cones, has been reproduced by taking $c = 1$.

2.1.2 Forms of relativistic dynamics

Starting from a four-vector with the canonical coordinates $x^\mu = (x^0, x^1, x^2, x^3)$, it is always possible to apply a transformation and obtain another parametrization of the same vector, $\bar{x}^\mu$. In a general form, we can write this transformation as

$$\bar{x}^\mu \equiv \bar{x}^\mu (x^\nu) . \tag{2.7}$$

Since in Physics it is always possible to change the reference frame and return to the original one, we require that the most general transformations of Eq. (2.7) are invertible functions. In principle, there exists an infinite number of ways to build invertible transformations. However, within the context of the covariant formalism, two parametrizations that are connected by Lorentz transformations are considered equivalent. In the end it turns out that there are only five independent ways to fix such parametrizations, and they are called "forms of relativistic dynamics." These forms represent distinct frameworks to describe the behavior of physical laws and equations in relativistic systems.

Before discussing the forms of the relativistic dynamics, we are interested in studying what else varies in a Minkowski spacetime when an invertible transformation is applied. First of all, the metric tensor also transforms according to:

$$\bar{g}_{\mu\nu} = \left(\frac{\partial x^\rho}{\partial \bar{x}^\mu}\right) g_{\rho\sigma} \left(\frac{\partial x^\sigma}{\partial \bar{x}^\nu}\right) . \tag{2.8}$$

The transformation (2.8) is crucial, since it preserves the physical laws and maintains the geometric structure of the light-cone. In fact, by computing the spacetime interval in the new coordinates we find:

$$\begin{aligned} d\bar{s}^2 &= \bar{g}_{\mu\nu} d\bar{x}^\mu d\bar{x}^\nu \\ &= \frac{\partial x^\rho}{\partial \bar{x}^\mu} g_{\rho\sigma} \frac{\partial x^\sigma}{\partial \bar{x}^\nu} d\bar{x}^\mu d\bar{x}^\nu \\ &= g_{\rho\sigma} dx^\rho dx^\sigma \\ &= ds^2, \end{aligned} \tag{2.9}$$

and this property does not change the light-cone regions discussed previously. It is worth investigating what happens to the present-hyperplane. The importance of the initial conditions in Physics is well-known and spans several domains. In QFTs, the fields are quantized at "equal times" $t = t' = t_0$. In the reference frame centered at the origin of the Minkowski spacetime, the present-hyperplane $t = 0$ corresponds to the same hyperplane displayed in Fig. 2.1. An invertible transformation of coordinates can affect the geometry of the hypersurface $\bar{x}^0 = 0$. The motivation resides in the new form of the metric tensor in Eq. (2.8), which can introduce non-trivial and intricate changes to the coordinates when lowering and raising the indices.

We here briefly discuss the five representations of the hypersurfaces $\bar{x}^0 =$ const. associated to the five forms of relativistic dynamics.

1. The first invertible transformation of Eq. (2.7) is simply the identity $\mathbb{1}$, associated with the canonical coordinates (x^0, x^1, x^2, x^3) and the metric tensor of Eq. (2.4). This usual parametrization was called by Dirac the *instant form* of relativistic dynamics. The present-hyperplane coincides with the one represented in Fig. 2.1, and it evolves in time by a translation along the x^0 axis.

2. The second possibility is represented by the *front form.* The parametrization of the spacetime is obtained from the instant form by applying a linear combination of the canonical coordinates, and the resulting metric tensor is quite simple.
   Analytically, the definition of the invertible transformation is

$$\begin{gathered}\left(x^0, x^1, x^2, x^3\right) \to \left(x^+, x^-, \boldsymbol{x}_\perp\right), \\ \text{with} \begin{cases} x^+ = \frac{1}{\sqrt{2}}\left(x^0 + x^3\right) \\ x^- = \frac{1}{\sqrt{2}}\left(x^0 - x^3\right) \\ \boldsymbol{x}_\perp = \left(x^1, x^2\right) \end{cases}, \end{gathered} \tag{2.10}$$

   and the metric tensor in the new variables results to be

$$\bar{g}_{\mu\nu} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{2.11}$$

   This implies that - in the end - the resulting "present" hypersurface of the instant form differs geometrically from the original instant-form hyperplane merely by a rotation of $\pi/4$ of the latter, which makes it tangent to the light-cone structure. This is the reason why the coordinates in Eq. (2.10) are often called *light-front coordinates*. In the context of QFTs, as mentioned before, the quantization of fields is imposed at equal "times". When changing the parametrization, the notion of "time"-coordinate changes as well. The "present" hyperplane is here represented by the geometrical locus of points characterized by the condition $x^+ = 0$, which is equivalent - in form - to the condition $x^0 = 0$ in the instant form. This equivalence is what leads us to interpret x^+ as the time coordinate in the front form. Therefore, the evolution of the hyperplane in time follows the increasing direction of x^+. The geometrical representation of the front form hyperplane is depicted in Fig. 2.2.

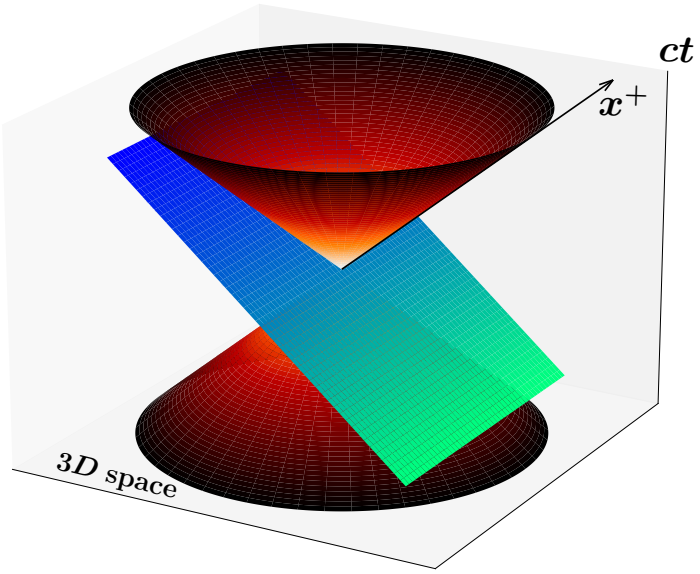


Figure 2.2: Representation of the hypersurface $x^+ = 0$, tangent to the light-cone, which evolves in the light-front time in the direction of x^+.

3. The third form is the *point form*. The spacetime coordinates are represented by the set $\bar{x}^\mu = (\tau, \omega, \theta, \phi)$, that is obtained by the canonical coordinates via the following invertible transformation:

$$\begin{cases} x^0 = \tau \cosh\omega \\ x^1 = \tau \sinh\omega \sin\theta \cos\phi \\ x^2 = \tau \sinh\omega \sin\theta \sin\phi \\ x^3 = \tau \sinh\omega \cos\theta \end{cases} . \tag{2.12}$$

The corresponding metric tensor is

$$\bar{g}_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -\tau^2 & 0 & 0 \\ 0 & 0 & -\tau^2 \sinh^2\omega & 0 \\ 0 & 0 & 0 & -\tau^2 \sinh^2\omega \sin^2\theta \end{pmatrix} . \tag{2.13}$$

The fixed-"time" hypersurface $\tau =$ const. can be rewritten in terms of the canonical coordinates by inverting Eq. (2.12):

$$\left(x^0\right)^2 - \left(x^1\right)^2 - \left(x^2\right)^2 - \left(x^3\right)^2 = \text{const.} \tag{2.14}$$

This is the analytical expression of an hyperboloid in four dimensions, that degenerates to the light cone itself for const. $= 0$. In the point form, the time coordinate has not a simple intuitive geometric representation, and it corresponds to the eigentime of the physical system. To investigate the "time" evolution of the hypersurface, it is possible to study the

problem (2.14) for different increasing values of the constant, under the conditions const. > 0 and $x^0 > 0$, as shown in Fig. 2.3.

These first three forms of relativistic dynamics were identified by Paul Dirac in 1949 [7]. The last two forms are simply variations of the point form, each associated with different hypersurfaces $\bar{x}^0 = 0$. Similar to the point form case, the *time* coordinate for these two forms has not a simple, intuitive geometric interpretation, and the evolution of the null-hypersurface can be investigated in a similar manner (see Fig. 2.3).

4. The first variation of the point form has the following expression for the fixed-"time" hypersurface, written in terms of the old canonical coordinates:

$$\begin{cases} (x^0)^2 - (x^1)^2 - (x^3)^2 = \text{const.} \\ x^0 > 0 \end{cases} . \tag{2.15}$$

5. The second variation of the point form, in the old canonical coordinates, reads:

$$\begin{cases} (x^0)^2 - (x^3)^2 = \text{const.} \\ x^0 > 0 \end{cases} . \tag{2.16}$$

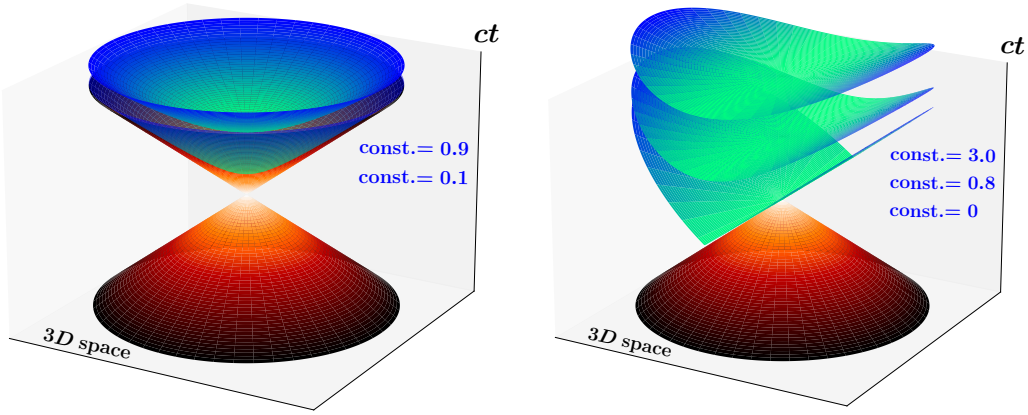


Figure 2.3: Reconstruction in 2+1 dimensions of the $\bar{x}^0 = \text{const.}$ hypersurfaces for the point form (left) and the second variation of the point form (right). In 2+1 dimensions, the first variation of the point form cannot be distinguished from the point form. Their difference could be appreciated only in 3+1 dimensions (see Eqs. (2.14) - (2.16)). Both figures show the evolution in time of the hypersurfaces, by studying their geometric behaviour for different increasing values of the constant. For the point form on the left, the const. = 0 hypersurface coincides with the light-cone.

In this work we are going to use the front form of relativistic dynamics, and from here on the symbol $g_{\mu\nu}$ will denote the metric tensor in front-form. Some

of the reasons for this choice will be clarified in the next Sections and Chapters. For now, we can say that, among all the parametrizations, the light front coordinates and the canonical ones are "privileged" since they preserve the geometry of the "present" hypersurface under "time" evolution. Moreover, the choice of the front form simplifies the computation in many-body problems in QFT, especially when dealing with deep inelastic processes.
The next question we want to answer regards the choice of the front form: *does the new parametrization affect also the canonical operators and their interpretations?*

2.1.3 Light cone operators

The operational structure of a theory emerges when one tries to quantize a classical theory by promoting the physical observables to operators. In this Section, we first introduce the observables associated to a very general classical theory, then we promote them to operators in the canonical instant form, and finally we explore the implications of re-parametrizing the spacetime in terms of the front form and how it affects the quantization process.

For a very general classical field theory, the starting point is represented by the *Lagrangian density*, that depends on the fields $\phi_r(x)$ and their derivatives $\partial_\mu \phi_r(x)$, i.e.,

$$\mathscr{L} \equiv \mathscr{L}\left(\phi_r(x), \partial_\mu \phi_r(x)\right). \tag{2.17}$$

The "time" derivatives of the fields play a special role, since they permit to introduce the momentum conjugate fields $\pi_r(x)$ associated to the fields themselves:

$$\pi_r(x) = \frac{\partial \mathscr{L}}{\partial\left(\partial_0 \phi_r(x)\right)}. \tag{2.18}$$

The Noether's theorem states, in general, that if a system has a symmetry property, then there are corresponding quantities whose values are conserved.

In case that the field theory is invariant under translations, then the four-divergence of the *energy-momentum tensor* is 0:

$$\partial_\mu T^{\mu\nu}(x) = 0, \tag{2.19}$$

where

$$T^{\mu\nu}(x) = \frac{\partial \mathscr{L}(\phi_r(x), \partial_\mu \phi_r(x))}{\partial\left(\partial_\mu \phi_r(x)\right)} \partial^\nu \phi_r(x) - g^{\mu\nu}_{\text{inst}} \mathscr{L}(\phi_r(x), \partial_\mu \phi_r(x)). \tag{2.20}$$

The metric tensor $g^{\mu\nu}_{\text{inst}}$ in the previous definition is supposed to be - for the moment - in the instant form.
We observe that the 00 component of the energy-momentum tensor is the *Hamiltonian density*, corresponding to the energy density:

$$T^{00} = \frac{\partial \mathscr{L}}{\partial(\partial_0 \phi_r)} \partial^0 \phi_r - \mathscr{L} = \mathscr{H}. \tag{2.21}$$

The other components are the energy-current density (T^{i0}), the momentum density (T^{0i}) and the momentum-current density (T^{ij}). Therefore, the four continuity equations (2.19) correspond to conservation of energy and conservation of momentum in each spatial direction. In fact, the quantities that are conserved over time are:

$$P^{\mu} = \int d^3x T^{0\mu}(x). \tag{2.22}$$

This can easily be seen by integrating over the space volume Eq. (2.19):

$$\partial_0 \int_{\mathbb{R}^3} d^3x T^{0\nu} = -\int_{\mathbb{R}^3} d^3x \partial_i T^{i\nu}. \tag{2.23}$$

The r.h.s. produces a vanishing contribution on the boundary ∂V for the Gauss-Stokes theorem. This implies that the member on the left is 0, which is equivalent to the statement that the conserved quantity is Eq. (2.22).

In case that the field theory is invariant under rotations, then the quantity whose four-divergence vanishes is the *generalized angular momentum density*:

$$\partial_\mu J^{\mu\nu\sigma}(x) = 0, \tag{2.24}$$

$$J^{\mu\nu\sigma}(x) = L^{\mu\nu\sigma}(x) + S^{\mu\nu\sigma}(x). \tag{2.25}$$

In the r.h.s. of Eq. (2.25) the orbital angular momentum and the intrinsic spin contributions appear. They are defined as

$$L^{\mu\nu\sigma}(x) = x^\nu T^{\mu\sigma}(x) - x^\sigma T^{\mu\nu}(x), \tag{2.26}$$

$$S^{\mu\nu\sigma}(x) = -i\pi^\mu_r \left(\Sigma^{\nu\sigma}\right)^r_s \phi_s(x), \tag{2.27}$$

where the spin operator is:

$$\begin{aligned}
&\left(\Sigma^{\mu\nu}\right)^r_s = 0 \text{ for spin-0 particle } \phi_r(x),\\
&\left(\Sigma^{\mu\nu}\right)^r_s = \frac{i}{2}\left([\gamma^\mu,\gamma^\nu]\right)^r_s \text{ for spin-1/2 particle } \psi_r(x),\\
&\left(\Sigma^{\mu\nu}\right)^\beta_\alpha = i\left(\delta^\mu_\alpha g^{\nu\beta} - \delta^\nu_\alpha g^{\mu\beta}\right) \text{ for spin-1 particle } A_\alpha(x).
\end{aligned} \tag{2.28}$$

Following the same reasoning as before, the conserved quantity in this context is represented by the so called *angular momentum tensor*:

$$M^{\mu\nu} = \int d^3x J^{0\mu\nu}(x) = \int d^3x \left[L^{0\mu\nu}(x) + S^{0\mu\nu}(x)\right]. \tag{2.29}$$

We are now interested in quantizing this theory, and studying the algebraic structure both in the instant form and in the front form. The ultimate objective is to interpret the meaning of the operators in the new formalism of light-front quantization.

Instant form quantization

When we promote the observables to operators, it is customary to indicate the latter with a different symbol with respect to the former. Often the symbol $\hat{O}$ represents the operator associated to the classical observable O. In our case, in order not to burden the notation, we do not distinguish the two symbols, but we just keep in mind that from here on we are referring to a quantum field theory with a set of quantum operators.

Let us introduce the *commutation relation* that links the position operator x^μ to the momentum operator P^μ:

$$\begin{aligned}[x^\mu, P^\nu] &:= x^\mu P^\nu - P^\nu x^\mu \\ &= i g^{\mu\nu}_{\text{inst}}.\end{aligned} \tag{2.30}$$

By now defining $M^{\mu\nu} = x^\mu P^\nu - x^\nu P^\mu$, we can identify the commutation relations between $M^{\mu\nu}$ and P^μ:

$$[P^\mu, P^\nu] = 0, \tag{2.31}$$

$$[P^\rho, M^{\mu\nu}] = i g^{\nu\rho}_{\text{inst}} P^\mu - i g^{\mu\rho}_{\text{inst}} P^\nu, \tag{2.32}$$

$$[M^{\rho\sigma}, M^{\mu\nu}] = i g^{\sigma\nu}_{\text{inst}} M^{\rho\mu} + i g^{\rho\mu}_{\text{inst}} M^{\sigma\nu} - i g^{\rho\nu}_{\text{inst}} M^{\sigma\mu} - i g^{\sigma\mu}_{\text{inst}} M^{\rho\nu}. \tag{2.33}$$

These equations constitute the commutation relations of the Lie algebra that generates the Poincaré group. We recall that the latter has 10 generators, corresponding to the possible transformations that we can achieve: the time translation (1), the space translations (3), the space rotations (3) and the boosts (3). In particular, the momentum operator P^μ generates the spacetime translations, while $M^{\mu\nu}$ is the generator of the rotations $J^i = \varepsilon^{ijk} M^{jk}$ and the boosts $K^i = M^{0i}$, globally constituting the Lorentz transformations:

$$M^{\mu\nu} = \begin{pmatrix} 0 & K^1 & K^2 & K^3 \\ -K^1 & 0 & J^3 & -J^2 \\ -K^2 & -J^3 & 0 & J^1 \\ -K^3 & J^2 & -J^1 & 0 \end{pmatrix}. \tag{2.34}$$

The instant form permits to interpret very intuitively the directions of application of the Lorentz transformations. In fact, the index $i = 1, 2, 3$ of P^i, J^i and K^i refers to the direction x^i of the canonical metric of the instant form, such that the boosts and the space translations are intended to act along those axes, while the rotations are applied around those axes. Finally, P^0 generates a translation in the time direction x^0. Obviously, these interpretations vary when the metric changes, and the geometric actions of these operators can become non-trivial.

We can recognize a common feature regarding some of the previous operators. Let us suppose to be in the origin of the light-cone in Fig. 2.1 and to apply a time translation by acting with P^0. After that, our space-time position changes and we end up in a point with a time-coordinate $x^0 > 0$. If we

now compare the new present-hyperplane with the initial one, we observe that they differ only by a (time) translation. Similarly, a Lorentz boost along a spatial axis modifies the temporal coordinate by introducing a "mixture" of it with spatial coordinates. This implies that, also in this case, the final present-hyperplane will be different with respect to the original one. This property is not true for the other possible transformations, i.e., the spatial translations and the rotations, that leave the time coordinate unchanged.
We define *kinematic operators* those which do not change the hypersurface $t = 0$ after their application, and *dynamical operators* the ones that vary the equal-time hypersurfaces along the time coordinate.
By summarizing, in instant form:

$$P^0, K^1, K^2, K^3 \text{ are dynamical operators;}$$
$$P^1, P^2, P^3, J^1, J^2, J^3 \text{ are kinematic operators.}$$

One may ask what kind of operator is the Hamiltonian of a physical system, and why it is not involved in the list of the previous operators. Actually, the Schrödinger equation contains the operational definition of the Hamiltonian of a system:

$$i\frac{\partial}{\partial x^0}\psi(x^0, \boldsymbol{x}) = H\psi(x^0, \boldsymbol{x}). \tag{2.35}$$

From Eq. (2.35), we can interpret the Hamiltonian as the displacement operator acting on the time coordinate, i.e. $H \equiv P^0$. This identification is also the motivation for which a quantum state evolves in time under the application of the operator e^{-iHt}: in fact, the finite generator of the translations coincides with $e^{-iP^\mu x_\mu}$, and a finite time translation is generated by $e^{-iP^0 x_0} = e^{-iP^0 x^0}$, where $t = x^0$ in the instant form.

The commutation relations introduced in Eqs. (2.32) – (2.33) can be rewritten, in a complete equivalent way, in terms of J^i and K^i as follows

$$\left[J^i, P^0\right] = 0, \qquad \left[J^i, P^l\right] = i\varepsilon^{ilm}P^m, \tag{2.36}$$
$$\left[K^i, P^0\right] = iP^i, \qquad \left[K^i, P^l\right] = -ig^{il}_{\text{inst}}P^0, \tag{2.37}$$
$$\left[J^i, J^l\right] = i\varepsilon^{ilm}J^m, \qquad \left[J^i, K^l\right] = -i\varepsilon^{ilm}K^m, \tag{2.38}$$
$$\left[K^i, K^l\right] = i\varepsilon^{ilm}J^m. \tag{2.39}$$

Front form quantization

We are interested in the differences arising from the change of the parametrization from the instant form to the front form. We know from Section 2.1.2 that the momentum operator in the light-front coordinates is expressed as

$$P^\mu = \left(P^+, P^-, \boldsymbol{P}_\perp\right), \tag{2.40}$$

where

$$P^\pm = \frac{1}{\sqrt{2}}\left(P^0 \pm P^3\right), \qquad \boldsymbol{P}_\perp = \left(P^1, P^2\right). \tag{2.41}$$

The angular momentum tensor operator $M^{\mu\nu}$ transforms like the metric tensor (2.8) with controvariant indices:

$$\bar{M}^{\mu\nu} = \left(\frac{\partial \bar{x}^{\mu}}{\partial x^{\rho}}\right) M^{\rho\sigma} \left(\frac{\partial \bar{x}^{\nu}}{\partial x^{\sigma}}\right), \tag{2.42}$$

thus producing

$$\bar{M}^{\mu\nu} = \begin{pmatrix} 0 & -K^3 & B^1 & B^2 \\ K^3 & 0 & S^1 & S^2 \\ -B^1 & -S^1 & 0 & J^3 \\ -B^2 & S^2 & -J^3 & 0 \end{pmatrix}, \tag{2.43}$$

where

$$B^1 = \frac{1}{\sqrt{2}}\left(K^1 + J^2\right), \qquad B^2 = \frac{1}{\sqrt{2}}\left(K^2 - J^1\right), \tag{2.44}$$

$$S^1 = \frac{1}{\sqrt{2}}\left(K^1 - J^2\right), \qquad S^2 = \frac{1}{\sqrt{2}}\left(K^2 + J^1\right). \tag{2.45}$$

We observe the non–trivial modification of the elements of the momentum tensor operator (2.43) compared to the instant-form given in Eq. (2.34).
The number of generators of the Poincaré group remains the same, but the new generators are: P^{μ}, K^3, J^3, $\boldsymbol{B}_{\perp} = (B^1, B^2)$ and $\boldsymbol{S}_{\perp} = (S^1, S^2)$.
The next step is to discern among these operators which are the dynamical operators and which the kinematic operators. Moreover, it is interesting to establish the new commutation relations relating the generators in the front form.

In analogy to the definition given in the instant form, the dynamical operators in the front form are those that modify the present-hyperplane $x^+ = 0$ after their application. As well as in instant form P^0 was the Hamiltonian H of the system, in the front form P^- is a dynamical operator that coincides with H. Indeed, by expanding the scalar product in the definition of the finite translation generator $e^{-iP^{\mu}x_{\mu}}$ we obtain:

$$\begin{aligned} P^{\mu}x_{\mu} &= P^+x_+ + P^-x_- - \boldsymbol{x}_{\perp}\cdot\boldsymbol{P}_{\perp} \\ &= P^+x^- + P^-x^+ - \boldsymbol{x}_{\perp}\cdot\boldsymbol{P}_{\perp}, \end{aligned} \tag{2.46}$$

and we recognize that the operator associated with the time coordinate x^+ is exactly P^-. On the contrary, P^+ is coupled to x^-, that is a spatial coordinate in the front form. Therefore, P^+ is a kinematic operator.

Concerning the operators (2.44) and (2.45), it is possible to show that B^1 and B^2 are kinematic, while S^1 and S^2 are dynamical. This is due to the separate action of a boost and a rotation: when observed at the level of the old instant form coordinates, the boost mixes the spatial coordinate x^i with the time coordinate x^0, while the rotation results in a mixing of the other two space coordinates (with a $\pm$ sign depending on the considered axis). The net effect is to leave $x^+ = 0$ untouched after the application of both B^1 and B^2 [8].

Differently, the action of S^1 and S^2 changes the present-hyperplane in the front form. The boost K^3, that was a dynamical operator in the instant form, becomes a kinematic operator in the front form, as it can be proven by acting on a four-vector $(x^+, x^-, \boldsymbol{x}_\perp)$ and then showing that $K^3 x^+ = 0 \iff x^+ = 0$.

By summarizing, in front form:

$$P^-, S^1, S^2 \text{ are dynamical operators;}$$
$$P^+, P^1, P^2, K^3, J^3, B^1, B^2 \text{ are kinematic operators.}$$

We notice that a simple change of parametrization can vary the number of kinematic and dynamical operators, as well as the classification of some of the operators involved.

Among the previous operators, we still do not know the significance of P^+. If we write the on-shell relation $M^2 = P_\mu P^\mu = 2P^+P^- - \boldsymbol{P}_\perp^2$ and solve it for P^-, we find:

$$P^- = \frac{\boldsymbol{P}_\perp^2}{2P^+} + \frac{M^2}{2P^+}. \tag{2.47}$$

This equation suggests the interpretation of P^+ as a mass parameter, after having identified P^- as the Hamiltonian.
The new algebraic commutation rules are the following ones:

$$\begin{aligned}
&\left[P^\mu, P^\nu\right] = \left[J^3, P^+\right] = \left[J^3, P^-\right] = \left[B^i, P^+\right] = 0,\\
&\left[P^-, B^i\right] = iP_\perp^i, \quad \left[J^3, P_\perp^i\right] = i\varepsilon^{il}P_\perp^l,\\
&\left[J^3, B^i\right] = -i\varepsilon^{il}B^l, \quad \left[P_\perp^i, B^l\right] = -i\delta^{il}P^+,
\end{aligned} \tag{2.48a}$$

$$\begin{aligned}
&\left[B^i, B^l\right] = \left[S^i, P^-\right] = \left[S^i, S^l\right] = \left[J^3, K^3\right] = \left[P_\perp^i, K^3\right] = 0,\\
&\left[B^i, K^3\right] = -iB^i, \quad \left[P^\pm, K^3\right] = -iP^\pm,\\
&\left[S^i, K^3\right] = iS^i, \quad \left[J^3, S^i\right] = -i\varepsilon^{il}S^l,\\
&\left[S^i, B^l\right] = -i\left(\delta^{il}K^3 + \varepsilon^{il}J^3\right), \quad \left[P^+, S^i\right] = -iP_\perp^i,\\
&\left[P_\perp^i, S^l\right] = -i\delta^{il}P^+.
\end{aligned} \tag{2.48b}$$

The first three relations (2.48a) contain just the operators P^μ, $\boldsymbol{B}_\perp$ and J^3. The set of these three operators constitute a subgroup of the Poincaré group that is *isomorphic* to the 2-dimensional subgroup of the non-relativistic Galilean transformations. This allows us, more formally than before, to identify P^- with the Hamiltonian, P^+ with the mass, $\boldsymbol{P}_\perp$ with translations, $\boldsymbol{B}_\perp$ with Galilean boosts and J^3 with rotations in the transverse plane of coordinates x^1 and x^2.

Before exploiting the advantages of this isomorphism, we want to stress that the structure of Eqs. (2.22) and (2.29) remains unchanged - in form - in the light-front coordinates:

$$P^\mu = \int d^3x T^{+\mu}(x). \tag{2.49}$$

$$M^{\mu\nu} = \int d^3x J^{+\mu\nu}(x) = \int d^3x \left[L^{+\mu\nu}(x) + S^{+\mu\nu}(x)\right]. \tag{2.50}$$

The isomorphism we identified leads to a parallelism between the dynamics in the transverse plane in the light-front formalism and the non-relativistic dynamics in 2 dimensions.

If we consider a non-relativistic composite physical system in 2 dimensions, whose Hamiltonian is invariant under translations, then it is convenient to separate the dynamics into two parts: the motion of the center of mass and the motion of the constituents relative to the center of mass. By identifying the position vectors $\mathbf{r}_i$ of all the N massive fragments, the position vector of the center of mass results to be:

$$\mathbf{R} = \sum_{i=1}^{N} \rho_i \mathbf{r}_i, \tag{2.51}$$

where ρ_i is the ratio between the mass of the i-th constituent and mass of the total system: $\rho_i = \dfrac{m_i}{M}$. The center of mass position is related to the generator of the Galilean boosts via

$$\mathbf{B} = -M\mathbf{R}. \tag{2.52}$$

By now exploiting the isomorphism, we can establish analogous relations, but with the introduction of the proper corresponding operators defined on the transverse plane in the front form. Thus, we identify the *center of transverse momentum* of the partons as

$$\boldsymbol{R}_\perp = \sum_{i=1}^{N} x_i \boldsymbol{r}_{\perp i}, \tag{2.53}$$

with the following relation to the generators (2.44)

$$\boldsymbol{R}_\perp = -\frac{1}{P^+}\boldsymbol{B}_\perp, \tag{2.54}$$

where $\boldsymbol{r}_{\perp i}$ is the position of the i-th parton in the transverse plane and $x_i = \dfrac{p_i^+}{P^+}$ is the fraction of collinear momentum carried by the i-th parton.

This discussion is very general, and it holds for each classical field theory with a Lagrangian density like (2.17). From the very last part of this Section, we learned that in Physics, when we deal with composite systems, it is often convenient to break down the problem into its constituent parts. We want now to consider the specific case of QCD, where we can describe hadron states in terms of the partons they contain. However, QCD adds complexity to the situation, since the number of hadron constituents is neither known nor fixed. In principle, to reconstruct the complete hadron state, one would have to consider an infinite number of partons, leading us to the concept of the Fock state decomposition, which will be discussed in Section 2.1.5.

2.1.4 Foundations of QCD

QCD is the QFT that describes the dynamics and the interactions of *coloured* particles. The colour is a quantum number that can take 3 different values, i.e., the colour number is $N_C = 3$. This means that the associated symmetry group is SU(3), i.e., the group of the unitary 3×3 matrices with determinant equal to one. From the general theory, the number of gauge bosons is the same as the number of the group generators, which is $N_C^2 - 1 = 8$.
The group generators are $T^a = \dfrac{\lambda^a}{2}$, with $a = 1, \ldots, 8$, and λ^a are the hermitian and traceless Gell-Mann matrices:

$$\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$\lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \lambda^5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \qquad \lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$
$$\lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \qquad \lambda^8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{2.55}$$

The previous matrices satisfy the following commutation relations

$$\left[\lambda^a, \lambda^b\right] = 2i f^{ab}{}_c \lambda^c, \tag{2.56}$$

where the structure constants of the $\mathfrak{su}(3)$ Lie algebra are

$$\begin{aligned} &f^{12}{}_3 = 2, \\ &f^{14}{}_7 = f^{24}{}_6 = f^{25}{}_7 = f^{34}{}_5 = 1, \\ &f^{15}{}_6 = f^{36}{}_7 = -1, \\ &f^{67}{}_8 = \sqrt{3}. \end{aligned} \tag{2.57}$$

The QCD is a pure Yang-Mills theory, without any spontaneous symmetry breaking, and this implies that all the 8 gauge bosons, that are the *gluons*, are massless.
The fermions of the QCD are the *quarks* and the *antiquarks* which, as elementary particles, acquire mass from the Higgs mechanism. Quarks and antiquarks exist in 6 different *flavours*: up (u), down (d), charm (c), strange (s), top (t) and bottom (b). Each quark of fixed flavour is characterized by a certain colour c that can assume exactly N_c distinct values.

The *quark field* of flavour q and colour c depending on the position x is denoted as $\Psi_c^q(x)$. The Lagrangian density including only the free fermion fields reads

$$\mathscr{L}_{\text{QCD}}^{\text{free}} = \sum_q \sum_{c,c'} \bar{\Psi}_{c'}^q(x) \left(i\partial\!\!\!/ - m\right) \Psi_c^q(x), \tag{2.58}$$

where the symbol $\not{\partial}$ stays for $\gamma^\mu \partial_\mu$ and γ^μ are the Dirac matrices in the so-called Standard Representation:

$$\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},$$
$$\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, \qquad \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{2.59}$$

In the light-cone formalism γ^0 and γ^3 are replaced by:

$$\gamma^\pm = \frac{1}{\sqrt{2}}(\gamma^0 \pm \gamma^3). \tag{2.60}$$

It is also worth introducing the definition of another 4×4 matrix that will be used in the discussion, and that combines all the Dirac matrices: this is $\gamma_5 := i\gamma^0\gamma^1\gamma^2\gamma^3$.

The free Lagrangian density (2.58) is invariant under a global transformation $\Psi^q_a(x) \to U^a{}_b \Psi^q_a(x)$, where $U^a{}_b$ is an element of SU(3) generated by T^a. The *gauge protocol* assures that if we promote the global transformation to a local transformation $U \to U(x)$, and we replace the canonical derivative ∂_μ with the covariant derivative D_μ, then we can obtain the expression for the complete Lagrangian density $\mathcal{L}_{\text{QCD}}$, which results invariant under the application of $U(x)$. The definition of the covariant derivative includes the eight *gauge fields* $A_{\mu\, a}(x)$, with the colour index $a = 1, \ldots, 8$:

$$D_\mu = \partial_\mu + i g_s T^a A_{\mu\, a}(x). \tag{2.61}$$

The dynamics of the gluon fields enters in the QCD Lagrangian density through a self-contraction of the gluon *field strength tensor*. The complete controvariant form of the latter is:

$$G_{\mu\nu\, a}(x) = \partial_\mu A_{\nu\, a}(x) - \partial_\nu A_{\mu\, a}(x) - g_s f^{bc}{}_a A_{\mu\, b}(x) A_{\nu\, c}(x). \tag{2.62}$$

Finally, the QCD Lagrangian density reads:

$$\mathcal{L}_{\text{QCD}} = \sum_q \sum_{c,c'} \bar{\Psi}^q_{c'}(x) \left(i\not{D} - m \right) \Psi^q_c(x) - \frac{1}{4} G^{\mu\nu}{}_a(x) G_{\mu\nu\, a}(x). \tag{2.63}$$

Both Quantum Electrodynamics (QED) and QCD are *gauge theories*, i.e., they are invariant under certain gauge-fixing conditions. There are many different ways to fix the gauge, and the choice is often guided by the need to simplify the practical calculations. In the light-front approach the most natural gauge choice is the *axial gauge*:

$$n_\mu A^\mu{}_a(x) = 0, \tag{2.64}$$

$$n^\mu = (0, 1, \mathbf{0}_\perp), \tag{2.65}$$

which is equivalent to $A^+_a(x) = 0$. This latter condition is called *light-cone gauge* [9].

As well as the theory of QED, QCD (2.63) contains the interactions between the fermion fields and the gauge bosons. In addition to that and differently from QED, it predicts the existence of interactions among gluon themselves, with the possibility of having vertices involving 3 and 4 fields. All these theoretical interaction processes have been tested and observed at particle colliders with excellent accuracy, thus providing a proof that QCD is a very fundamental and solid theory at the basis of the standard model.
Nevertheless, there are still several open problems related to QCD. First of all, we introduced 8 massless gluons, that are the mediators of the strong interaction. Usually, if the gauge bosons of a QFT are massless, then the interaction that they mediate has a long (infinite) range of action, but this is not the case of QCD, where the strong interaction is characterized by a *short range*, typically of the order of 1 fm. This is due to the most mysterious and distinctive feature of this theory: the *colour confinement*. The colour confinement obliges all the particles endowed with colours to be confined inside a colour-less bound state. In other words, the free fields of quarks, antiquarks and gluons cannot exist in nature apart from inside the hadrons. At the moment nobody knows why does the colour confinement exist and a mathematical proof is lacking.
To this regard, we want to stress that the only way to get access to the properties of the free coloured fields is to use very energetic probes that can penetrate inside the hadrons and interact with quarks, antiquarks and gluons. The higher the energy of the probe, the smaller the distance it can resolve.

Another feature of QCD is the behaviour of the coupling constant α_S as function of the energy scale. As the energy scale increases, the value of the coupling constant decreases, thus causing the interactions between particles to become asymptotically weaker. This property is called *asymptotic freedom*, and it is exactly what it guarantees the applicability of the techniques of perturbative calculations (pQCD). On the opposite limit, when the energy scale decreases, the value of α_s becomes bigger and all the pQCD predictions fail. The mass energy of the hadrons (around 1 GeV) lies in a region in which the value of α_S is far from being small, and other alternatives to pQCD are needed for implementing the calculations. The light-front quantization provides an efficient tool for this purpose, since it is the starting point to reach a description in terms of the Light-Front Wave Function (LFWF) representation [10] that provides an intuitive physical interpretation of a large class of parton distribution functions. This is exactly the path we aim to pursue throughout the next Sections.

Apart from the quantum numbers specific to the QCD sector, the partons involved have also an intrinsic spin. In the framework of light-front quantization, the appropriate operator to introduce is the *light-front helicity* [11],

i.e.,

$$j_z = J^3 - \boldsymbol{R}_\perp \times \boldsymbol{P}_\perp, \tag{2.66}$$

which can be interpreted as an intrinsic angular momentum, being the difference between the total and the orbital angular momentum in the x^3 direction. Let us consider the momentum operator P^μ and the light-front helicity operator (2.66). We denote as $|p, \lambda\rangle$ a state which is an eigenstate of both these operators:

$$P^\mu |p, \lambda\rangle = p^\mu |p, \lambda\rangle, \tag{2.67}$$

$$j_z |p, \lambda\rangle = \lambda |p, \lambda\rangle. \tag{2.68}$$

It is possible to prove [8, 11] that the definition in Eq. (2.66) reduces to the ordinary helicity, which is interpreted as the projection of the total spin along the momentum direction, in the infinite-momentum limit $p^+ \to \infty$.

2.1.5 Fock state decomposition

To properly describe the internal structure of hadrons in the framework of relativity, we must identify a suitable basis of states that represents an irreducible representation of the Poincaré group. In the light-front dynamics the operators $\{P^+, \boldsymbol{P}_\perp, j_z, M, \boldsymbol{J}^2\}$ form a complete set of commuting self-adjoint operators. Additional quantum numbers such as flavour and colour, play no special role in the description of the transformation properties under the action of the Poincaré group and can be added to characterize the internal structure of the particle state [12]. Accordingly, the i-th parton is represented in terms of a colour c_i, a helicity λ_i, a longitudinal momentum p_i^+, a transverse momentum $\boldsymbol{p}_{\perp i}$ and, in case of a fermion, a flavour q_i.

For a system of N-interacting partons, we introduce the Fock-state decomposition of the Hilbert space as

$$\mathcal{F} := \mathrm{H}_0 \oplus \bigoplus_{N=1}^{\infty} \mathrm{H}^{\otimes N}, \tag{2.69}$$

where H_0 is the state with zero partons, while H contains the single-particle states, that can be quarks, antiquarks and gluons.
The N-particle *Fock state* is then represented as

$$|N, w_1^{c_1\lambda_1}, ..., w_N^{c_N\lambda_N}\rangle\,, \tag{2.70}$$

where $w_i = (p_i^+, \boldsymbol{p}_{\perp i})$.

The Fock states are *complete*. This implies that if we sum over all the discrete variables and integrate over all the continuous ones, we recover the identity:

$$\sum_N \not\int dw\, |N, w_1^{c_1\lambda_1}, ..., w_N^{c_N\lambda_N}\rangle \langle N, w_1^{c_1\lambda_1}, ..., w_N^{c_N\lambda_N}| = \mathbb{1}. \tag{2.71}$$

The symbol $\not\sum\!\!\!\!\!\!\int$ represents the combination of summing over helicities and colours and integrating over momenta, according to:

$$\not\sum\!\!\!\!\!\!\int dw = \sum_{\lambda_i}\sum_{c_i}\int\prod_{i=1}^{n} dp_i^+ \frac{d^2\boldsymbol{p}_{\perp i}}{(16\pi^3)^{n-1}}\delta\Big(p^+ - \sum_{i=1}^{n} p_i^+\Big)\delta^{(2)}\Big(\boldsymbol{p}_\perp - \sum_{i=1}^{n}\boldsymbol{p}_{\perp i}\Big). \tag{2.72}$$

The N-particle state in Eq. (2.70) is built from the QCD light-front vacuum state $|0\rangle$ by acting with the creation operators of the free-fields:

$$|N, w_1^{c_1\lambda_1}, ..., w_N^{c_N\lambda_N}\rangle = \prod_{j=1}^{N_q} q^\dagger_{c_j\lambda_j}(w_j)\prod_{l=1}^{N_{\bar q}}\bar q^\dagger_{c_l\lambda_l}(w_l)\prod_{m=1}^{N_{gl}} g^\dagger_{c_m\lambda_m}(w_m)\,|0\rangle\,, \tag{2.73}$$

where $N_q + N_{\bar q} + N_{gl} = N$ and

$q^\dagger_{c_i\lambda_i}(w_i)$ creates a quark with colour c_i, light-cone helicity λ_i flavour q and momentum $(p_i^+, \boldsymbol{p}_{\perp i})$;

$\bar q^\dagger_{c_i\lambda_i}(w_i)$ creates an antiquark with colour c_i, light-cone helicity λ_i flavour q and momentum $(p_i^+, \boldsymbol{p}_{\perp i})$;

$g^\dagger_{c_i\lambda_i}(w_i)$ creates a gluon with colour c_i, light-cone helicity λ_i and momentum $(p_i^+, \boldsymbol{p}_{\perp i})$. (2.74)

We recall that the colour indices for the gluons can take 8 different values, while the colours of the (anti)quarks exist only in 3 different types. In this Section we do not distinguish the two symbols for (anti)quarks and gluons in order not to burden the notation. In the next Sections we will return to distinguish the colour a_i of the i-th gluon from the colour c_i of the i-th fermion.

The creation ladder operators and the corresponding annihilation operators obey the following commutations and anti-commutations relations in light-front quantization

$$\Big[g_{c'\lambda'}(w_i'), g^\dagger_{c\lambda}(w_i)\Big] = 16\pi^3 p_i^+\delta\left(p_i'^+ - p_i^+\right)\delta^{(2)}\left(\boldsymbol{p}'_{\perp i} - \boldsymbol{p}_{\perp i}\right)\delta_{c'c}\delta_{\lambda'\lambda}, \tag{2.75}$$

$$\Big\{q'_{c'\lambda'}(w_i'), q^\dagger_{c\lambda}(w_i)\Big\} = 16\pi^3 p_i^+\delta\left(p_i'^+ - p_i^+\right)\delta^{(2)}\left(\boldsymbol{p}'_{\perp i} - \boldsymbol{p}_{\perp i}\right)\delta_{q'q}\delta_{c'c}\delta_{\lambda'\lambda}, \tag{2.76}$$

$$\Big\{\bar q'_{c'\lambda'}(w_i'), \bar q^\dagger_{c\lambda}(w_i)\Big\} = 16\pi^3 p_i^+\delta\left(p_i'^+ - p_i^+\right)\delta^{(2)}\left(\boldsymbol{p}'_{\perp i} - \boldsymbol{p}_{\perp i}\right)\delta_{q'q}\delta_{c'c}\delta_{\lambda'\lambda}. \tag{2.77}$$

The Fock states $|N, w_1^{c_1\lambda_1}, ..., w_N^{c_N\lambda_N}\rangle$ are eigenstates of the momentum operator P^μ, with eigenvalues p^+ and $\boldsymbol{p}_\perp$, which are obtained from the N-parton momenta as

$$\sum_{i=1}^{N} p_i^+ = p^+, \quad \sum_{i=1}^{N}\boldsymbol{p}_{\perp i} = \boldsymbol{p}_\perp, \tag{2.78}$$

where $p_i^+ > 0$. The eigenstate equation, applied to the vacuum state of H_0, produces:

$$P^+ |0\rangle = 0, \tag{2.79}$$
$$\boldsymbol{P}_\perp |0\rangle = \boldsymbol{0}. \tag{2.80}$$

We observe that the property $p_i^+ > 0$ is crucial to interpret the vacuum QCD state $|0\rangle$ as the state with zero particles and zero momentum. On the contrary, if the positive condition of the collinear momentum of the partons were not satisfied, a certain state with $N \neq 0$ particles but 0 momentum in the collinear direction would contribute to the total vacuum state, and the previous identification would fail [13].

Thanks to the completeness relation of the set of Fock states and to the properties (2.78), we can rewrite the general hadron state in terms of the parton constituents:

$$|p^+, \boldsymbol{p}_\perp, \Lambda\rangle = \sum_N \sum\!\!\!\!\!\!\int \psi^\Lambda_{N,\beta}(r) \, |N, w_1^{c_1\lambda_1}, ..., w_N^{c_N\lambda_N}\rangle \, . \tag{2.81}$$

The coefficients appearing in Eq. (2.81) are the LFWFs of the hadron which represent the probability amplitudes of finding N partons inside the hadron with helicity Λ, each possessing specific momenta w_i, collectively denoted with r, and discrete quantum numbers c_i, λ_i and, for quarks and antiquarks, flavour q_i, organized in the index β:

$$\psi^\Lambda_{N,\beta}(r) \equiv \langle p^+, \boldsymbol{p}_\perp, \Lambda | N, w_1^{c_1\lambda_1}, ..., w_N^{c_N\lambda_N}\rangle \, . \tag{2.82}$$

Given the structure of the front-form, in which the transverse boost operators (2.44) are kinematic, the LFWFs do not depend on the momentum of the parent hadron, but only on the parton momenta *relative* to the composite particle, i.e.

$$\psi^\Lambda_{N,\beta}(r) \equiv \psi^\Lambda_{N,\beta}\left(x_1, \boldsymbol{k}_{\perp 1}, \ldots, x_N, \boldsymbol{k}_{\perp N}\right) \tag{2.83}$$

where

$$x_i = \frac{p_i^+}{p^+}, \qquad \boldsymbol{k}_{\perp i} = \boldsymbol{p}_{\perp i} - x_i \boldsymbol{p}_\perp. \tag{2.84}$$

It is possible to exploit Eq. (2.84) to change the transverse integration variables in Eq. (2.72) from $\boldsymbol{p}_{\perp i}$ to $\boldsymbol{k}_{\perp i}$ and, accordingly, to rewrite the Fock states as $|N, \tilde{w}_1^{c_1\lambda_1}, ..., \tilde{w}_N^{c_N\lambda_N}\rangle$, where now $\tilde{w}_i = \left(p_i^+, \boldsymbol{k}_{\perp i} + x_i \boldsymbol{p}_\perp\right)$. This also implies that Eq. (2.81) is replaced by

$$|p^+, \boldsymbol{p}_\perp, \Lambda\rangle = \sum_N \sum_{c_i} \sum_{\lambda_i} \int \prod_{i=1}^N \mathrm{d}[i] \psi^\Lambda_{N,\beta}(r) \, |N, \tilde{w}_1^{c_1\lambda 1}, ..., \tilde{w}_N^{c_N\lambda N}\rangle \, , \tag{2.85}$$

where the integration measure reads

$$\prod_{i=1}^N \mathrm{d}[i] = [dx]_N \left[d^2\boldsymbol{k}_\perp\right]_N , \tag{2.86}$$

$$[dx]_N = \prod_{i=1}^{N} \frac{dx_i}{\sqrt{x_i}} \delta\left(1 - \sum_{i=1}^{N} x_i\right), \tag{2.87}$$

$$\left[d^2\boldsymbol{k}_\perp\right]_N = \frac{1}{[2(2\pi)^3]^{N-1}} \prod_{i=1}^{N} d^2\boldsymbol{k}_{\perp i} \delta^{(2)}\left(\sum_{i=1}^{N} \boldsymbol{k}_{\perp i}\right). \tag{2.88}$$

2.1.6 Free fields and their quantization

The quantization of free fields consists in promoting the fields of a classical theory (2.17) to operators and then imposing commutation relations (for bosons) or anti-commutation relations (for fermions) between the fields and their conjugates (2.18). These relations are intended to be true only on the same "time" hypersurface, thus they depend on the specific choice of the form of dynamics. In our case, we are interested in the fields of QCD in the front-form. This implies that the (anti)commuting relations for the fields are specified for $x^+ = x'^+$, where x and x' are the four-positions of the two fields involved. This is exactly what defines the *light-front quantization*.

The purpose of this Section is to write explicit expressions for the free fields of QCD, i.e., the quark field, the antiquark field and the gluon fields that appear in Eq. (2.63). These definitions will exploit the quantum ladder operators in (2.74), that provide a way to generate from the QCD light-front vacuum the partons with certain quantum numbers and momenta.
We firstly derive the expressions for the conjugate of the fermion fields (2.18) from the QCD Lagrangian density (2.63). We obtain that the conjugate field for $\bar{\Psi}^q_c(x)$ vanishes, while the one for $\Psi^q_c(x)$ is:

$$\pi^q(x) = i\bar{\Psi}^q_{c'}(x)\gamma^+ := i{\Psi^q_{c'}}^\dagger(x)\gamma^0\gamma^+. \tag{2.89}$$

This equation suggests that $\Psi^q_c(x)$ and $\bar{\Psi}^q_c(x)$ are not dynamically independent fields and each of them can be derived from the other.
In the case of the boson field, the situation is more complicated. This can be seen also in instant-form, where the conjugate fields $\pi^g_{\mu\, a}(x)$ associated to $A_{\mu\, a}(x)$ computed from the QCD Lagrangian density vanish. To overcome this problem, the free-field term $-\frac{1}{4}G^{\mu\nu}{}_a(x)G_{\mu\nu\, a}(x)$ is modified with the addition of a four-divergence, that leaves the equations of motion (EOM) unchanged. The problem, then, is to introduce certain gauge-fixing conditions to suppress the non-physical components of the gluon fields. The light-cone gauge (2.64) - (2.65) does not fix the gauge completely, as the number of the independent physical components for each $a = 1, \ldots, 8$ must be 2 (and not 3). One possible solution in the front form is to require, in addition to $A^+{}_a(x) = 0$, some boundary conditions for the transverse components of the gluon fields $\boldsymbol{A}_{\perp\, a}(\pm\infty^-)$, thanks to which $A^-{}_a(x)$ is no longer an independent degree of freedom [14].

By recalling the discussion about the light-front helicity at the end of Section 2.1.4, we introduce the quantum number of light-front helicity into the

QCD fields and arrive at the following light-front quantization conditions:

$$\left\{\Psi^{q'}_{c'\lambda'}(x'),\Psi^{q\dagger}_{c\lambda}(x)\right\}\Big|_{x'^+=x^+}=\delta(x'^- - x^-)\delta^{(2)}(\boldsymbol{x'}_\perp-\boldsymbol{x}_\perp)\delta_{q'q}\delta_{c'c}\delta_{\lambda'\lambda}\,, \tag{2.90}$$

$$\left[A^{i'}{}_{a'\lambda'}(x'),-\frac{\partial}{\partial x^-}A^{i}{}_{a\lambda}(x)\right]\Big|_{x'^+=x^+}=\delta(x'^- - x^-)\delta^{(2)}(\boldsymbol{x'}_\perp-\boldsymbol{x}_\perp)\delta_{i'i}\delta_{a'a}\delta_{\lambda'\lambda}\,, \tag{2.91}$$

where the upper latin index of the gluon field identifies one of the two components of $\boldsymbol{A}_{\perp\,a}(x)$.

We give now the general explicit expressions for the QCD fields depending on the 4-dimensional position x:

$$\Psi^{q}_{c\lambda}(x)=\int\frac{dk^+d^2\boldsymbol{k}_\perp}{2k^+(2\pi)^3}\Theta\left(k^+\right)\left[e^{-ik\cdot x}\varphi^{q}_{c\lambda}(w)+e^{ik\cdot x}\varphi^{\bar q}_{c\lambda}(w)\right], \tag{2.92}$$

$$\bar\Psi^{q}_{c\lambda}(x)=\int\frac{dk^+d^2\boldsymbol{k}_\perp}{2k^+(2\pi)^3}\Theta\left(k^+\right)\left[e^{-ik\cdot x}\bar\varphi^{\bar q}_{c\lambda}(w)+e^{ik\cdot x}\bar\varphi^{q}_{c\lambda}(w)\right], \tag{2.93}$$

$$A^{i}{}_{a\lambda}(x)=\int\frac{dk^+d^2\boldsymbol{k}_\perp}{2k^+(2\pi)^3}\Theta\left(k^+\right)T^a\left\{g_{a\lambda}(w)\epsilon^i_\lambda(w)e^{-ik\cdot x}+g^\dagger_{a\lambda}(w)\epsilon^{i*}_\lambda(w)e^{ik\cdot x}\right\}. \tag{2.94}$$

In these equations, w represents the three-momentum $w=(k^+,\boldsymbol{k}_\perp)$, and Θ is the Heaviside Theta Function:

$$\Theta\left(k^+\right)=\begin{cases}1 & \text{if } k^+>0;\\ 0 & \text{otherwise.}\end{cases} \tag{2.95}$$

The operators appearing in the fermionic fields are combinations of the ladder operators (2.74) and the Dirac spinors:

$$\begin{aligned}\varphi^{q}_{c\lambda}(w)&=u_\lambda(w)q_{c\lambda}(w),\\ \varphi^{\bar q}_{c\lambda}(w)&=v_\lambda(w)\bar q^\dagger_{c\lambda}(w),\\ \bar\varphi^{\bar q}_{c\lambda}(w)&=\bar u_\lambda(w)q^\dagger_{c\lambda}(w),\\ \bar\varphi^{q}_{c\lambda}(w)&=\bar v_\lambda(w)\bar q_{c\lambda}(w).\end{aligned} \tag{2.96}$$

The Dirac spinors in the light-cone formalism are

$$u_\uparrow(w)=\frac{1}{\sqrt{2^{3/2}k^+}}\begin{pmatrix}\sqrt{2}k^++m\\ k_R\\ \sqrt{2}k^+-m\\ k_R\end{pmatrix},\qquad u_\downarrow(w)=\frac{1}{\sqrt{2^{3/2}k^+}}\begin{pmatrix}-k_L\\ \sqrt{2}k^++m\\ k_L\\ -\sqrt{2}k^++m\end{pmatrix},$$

$$v_\uparrow(w)=\frac{-1}{\sqrt{2^{3/2}k^+}}\begin{pmatrix}-k_L\\ \sqrt{2}k^+-m\\ k_L\\ -\sqrt{2}k^+-m\end{pmatrix},\qquad v_\downarrow(w)=\frac{-1}{\sqrt{2^{3/2}k^+}}\begin{pmatrix}\sqrt{2}k^+-m\\ k_R\\ \sqrt{2}k^++m\\ k_R\end{pmatrix}, \tag{2.97}$$

where $k_{R/L} = \frac{1}{\sqrt{2}}(k_x \pm ik_y)$.
As observed previously, the two fields in Eqs. (2.92) and (2.93) are not independent, since the second one can be derived from the first one by taking the adjoint and then multiplying it by γ^0. Vice-versa, the quark field can be derived from the antiquark field by multiplying the latter by γ^0 and then by computing the adjoint.

Actually, the independent dynamical fermionic fields are represented by the so-called *good* light-front components of the fields (see, for instance, Ref. [15]). They are obtained from Eqs. (2.92) - (2.93) via a projection through the operators $\mathscr{P}_\pm$:

$$\Psi^{q+}_{c\lambda}(x) \equiv \mathscr{P}_+ \Psi^q_{c\lambda}(x), \tag{2.98}$$
$$\bar{\Psi}^{q+}_{c\lambda}(x) \equiv \bar{\Psi}^q_{c\lambda}(x)\mathscr{P}_-, \tag{2.99}$$
$$\mathscr{P}_\pm = \frac{1}{2}\gamma^\mp\gamma^\pm, \tag{2.100}$$
$$\mathscr{P}_+ + \mathscr{P}_- = \mathbb{1}, \tag{2.101}$$
$$\mathscr{P}_\pm^\dagger = \mathscr{P}_\pm. \tag{2.102}$$

Only by projecting the expressions (2.92) and (2.93) on the good components, we obtain the physical independent fermionic fields:

$$\Psi^{q+}_{c\lambda}(x) = \int \frac{dk^+ d^2\mathbf{k}_\perp}{2k^+(2\pi)^3}\Theta\left(k^+\right)\left[e^{-ik\cdot x}\varphi^{q+}_{c\lambda}(w) + e^{ik\cdot x}\varphi^{\bar{q}+}_{c\lambda}(w)\right], \tag{2.103}$$
$$\bar{\Psi}^{q+}_{c\lambda}(x) = \int \frac{dk^+ d^2\mathbf{k}_\perp}{2k^+(2\pi)^3}\Theta\left(k^+\right)\left[e^{-ik\cdot x}\bar{\varphi}^{\bar{q}+}_{c\lambda}(w) + e^{ik\cdot x}\bar{\varphi}^{q+}_{c\lambda}(w)\right], \tag{2.104}$$

where the operators φ^{q+} and $\bar{\varphi}^{\bar{q}+}$ involve now the good components of the spinors, i.e.

$$u^+_\uparrow(w) = \sqrt{\frac{k^+}{2^{1/2}}}\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \quad u^+_\downarrow(w) = \sqrt{\frac{k^+}{2^{1/2}}}\begin{pmatrix}0\\1\\0\\-1\end{pmatrix},$$
$$v^+_\uparrow(w) = \sqrt{\frac{k^+}{2^{1/2}}}\begin{pmatrix}0\\-1\\0\\1\end{pmatrix}, \quad v^+_\downarrow(w) = \sqrt{\frac{k^+}{2^{1/2}}}\begin{pmatrix}-1\\0\\-1\\0\end{pmatrix}. \tag{2.105}$$

Generally, the fermion fields appearing in the QCD calculations are represented by Eqs. (2.92) - (2.93) and (2.103) - (2.104) with a sum over the colour and helicity indices:

$$\Psi^q(x) = \sum_{c,\lambda}\Psi^q_{c\lambda}(x), \tag{2.106}$$

$$\bar{\Psi}^q(x) = \sum_{c,\lambda} \bar{\Psi}^q_{c\lambda}(x), \tag{2.107}$$

$$\Psi^{q+}(x) = \sum_{c,\lambda} \Psi^{q+}_{c\lambda}(x), \tag{2.108}$$

$$\bar{\Psi}^{q+}(x) = \sum_{c,\lambda} \bar{\Psi}^{q+}_{c\lambda}(x). \tag{2.109}$$

The gluon fields in Eq. (2.94) have only transverse components, i.e., they correspond to the two genuine independent degrees of freedom. The polarization vectors appearing in their definition are expressed in light-front formalism as

$$\epsilon^\mu_\lambda(w) = \left(0, \frac{\boldsymbol{\epsilon}_\perp(\mu)\cdot\boldsymbol{k}_\perp}{k^+}, \boldsymbol{\epsilon}_\perp(\mu)\right), \tag{2.110}$$

with

$$\boldsymbol{\epsilon}_\perp(\uparrow / \downarrow) = \mp\frac{1}{\sqrt{2}}(1, i). \tag{2.111}$$

2.2 Parton Distribution Functions

All the elements presented so far form the basis for describing hadrons in terms of their constituents. A prime example of how we can use all the theoretical machinery previously developed is in the study of parton distribution functions. These objects embody the *journey in phenomenology*, as their theoretical definitions and predictions need to be compared to phenomenological extractions of the cross sections in various scattering experiments.
In this Section we present the landscape of the different types of parton distribution functions and specify the most general ones, from which all the others can be derived under certain specific limits [16–20].

2.2.1 Generalized parton correlator

Let us consider an interaction process between a hadron and a probe particle. In the initial state the hadron is described by the four-momentum $p = (p^+, p^-, \boldsymbol{p}_\perp)$ and the light-front helicity Λ. In the most general case, after the interaction, the momentum and the light-front helicity of the hadron change and become p' and Λ'.

We introduce the *hadron average momentum* and the *momentum transfer* between the initial and final states:

$$P := \frac{p + p'}{2}, \tag{2.112}$$

$$\Delta := p' - p. \tag{2.113}$$

The on-shell conditions $p^2 = p'^2 = M^2$ impose the following relations

$$P^2 = M^2 - \frac{\Delta^2}{4}, \qquad P \cdot \Delta = 0. \tag{2.114}$$

The very basic assumption of the parton model consists in the hypothesis that the interaction takes place between the probe particle and only one single parton inside the hadron, which is the *active parton*. All the other partons act as *spectators* in the scattering process. This assumption is reasonable when the resolution power of the probe is sufficient to distinguish the internal structure of the hadron. This is associated to the energy scale at which the process takes place. If this energy is too small, the hadron is just seen as a structure-less point-like object and the parton model predictions fail. We assume from now that the energy scale is sufficient to support the parton model hypothesis.
We consider the case in which the active parton is a quark carrying an initial momentum $k - \frac{\Delta}{2}$ and ending up with a final momentum $k + \frac{\Delta}{2}$.

The process is completely described only by taking into account the physics of the probe, the physics of the hadron and their interaction. Here we focus merely on the hadronic part, where we can introduce the most general, *fully unintegrated quark-quark correlator function* between hadron states in light-front helicity basis [16, 20]:

$$\Phi_{\Lambda\Lambda'}(k, \Delta; P) = \int \frac{d^4\zeta}{(2\pi)^4} e^{ik\zeta} \langle p', \Lambda' | \, \bar{\Psi}^q(0) \mathcal{U}_{(0,\zeta)} \Psi^q(\zeta) \, | p, \Lambda \rangle \, . \tag{2.115}$$

The schematic illustration of the generalized parton correlator is represented in Fig. 2.4.

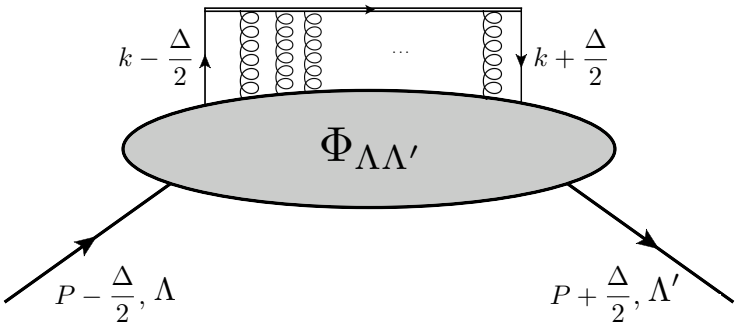


Figure 2.4: Schematic illustration of the generalized parton correlator within the parton model assumption. The effect of the gauge-link is represented by the emission of the soft gluons between the active quark and the remnant.

The definition (2.115) is simply a Fourier Transform (FT) of a matrix element containing the final and initial states of the hadron, two fermion fields

localized in two different positions, and a *gauge link operator* $\mathcal{U}_{(0,\zeta)}$ connecting the two points 0 and ζ. The role of the gauge link will be explained at the end of this Section. The variables Δ, P, Λ and Λ' can be - in principle - experimentally determined. Vice-versa, it is not possible to measure the momentum of the active quark before and after the interaction due to the colour confinement.

Let us trace the correlator (2.115) over the Dirac space by inserting an operator Γ generated from the *Dirac basis* $\mathcal{D} = \{\mathbb{1}, \gamma^\mu, \gamma_5, \gamma^\mu\gamma_5, i\sigma^{\mu\nu}\}$, i.e.

$$\begin{aligned}\Phi^{[\Gamma]}_{\Lambda\Lambda'}(k,\Delta;P) &= \frac{1}{2}\mathrm{Tr}[\Phi_{\Lambda\Lambda'}(k,\Delta;P)\Gamma] \\ &= \frac{1}{2}\int \frac{d^4\zeta}{(2\pi)^4} e^{ik\cdot\zeta} \langle p',\Lambda'|\, \bar{\Psi}^q_j(0)\mathcal{U}_{(0,\zeta)}\Gamma_{ji}\Psi^q_i(\zeta)\,|p,\Lambda\rangle\,, \end{aligned} \tag{2.116}$$

where i and j are Dirac indexes. By varying Γ, the corresponding hadron correlator provides complementary information on the partonic spin structure inside the hadron.

Due to the non-perturbative nature of hadronic physics, it is not possible to completely compute the quark-quark correlator analytically. To overcome this problem, one possibility is to parametrize it in terms of measurable quantities and then reconstruct the analytical structure using phenomenological methods [18].

The choice of the Dirac operator between $\gamma^+, \gamma^+\gamma_5, i\sigma^{j+}\gamma_5$, selects the helicity configuration of the active parton in the initial and in the final states, thus allowing one to account for contributions to the correlator from various helicity-transforming processes. These contributions enter in the description of the hadronic process at *twist*-2. What we mean with "twist" is actually the *dynamical twist*, as introduced in Ref. [21], i.e., the value $t \in \mathbb{N}$ at which an operator enters at the level of the cross-section according to

$$\left(\frac{M}{\sqrt{Q^2}}\right)^{t-2}, \tag{2.117}$$

where Q^2 is the energy scale of the process. From this definition, we define $t = 2$ as the *leading-twist*.

We now return on the question regarding the presence of the gauge-link $\mathcal{U}_{(0,\zeta)}$ in the definition of the correlator. Let us consider a local gauge transformation acting on the fermionic fields $\Psi(x)$ of a generic QFT:

$$\Psi(x) \to e^{-ig\alpha(x)}\Psi(x), \tag{2.118}$$

where g is the the coupling constant between the fermion and the gauge field of the theory $A^\mu(x)$. The effect of the same transformation on the field $\bar{\Psi}(x)$ changes the sign in the exponential. It is clear that if the two fields are defined in two different points, then the effect of the gauge transformation on the product of the two fields is:

$$\bar{\Psi}(x)\Psi(y) \to \bar{\Psi}(x)\Psi(y)e^{ig(\alpha(x)-\alpha(y))}. \tag{2.119}$$

This is a problem for the gauge invariance of the theory. What breaks the gauge-invariance is the bi-locality of the field operators. If $x = y$, the gauge invariance would be preserved.
For this reason, it is necessary to interpose an object between the two fermionic fields that transforms under (2.118) in a way that the gauge invariance is restored. In other words, we need an operator $\mathcal{U}_{(x,y)}$ s.t.

$$\mathcal{U}_{(x,y)} \to e^{-ig\alpha(x)}\mathcal{U}_{(x,y)}e^{ig\alpha(y)}. \tag{2.120}$$

The gauge-link, also known as *Wilson line*, transforms exactly in the proper way, and it is defined as

$$\mathcal{U}_{(x,y)} = \mathcal{P}\exp\Big[ig\int_x^y d\eta^\mu A_\mu(\eta)\Big], \tag{2.121}$$

where $\mathcal{P}$ denotes the path ordering of the integral from x to y. In the case of an Abelian theory like QED the path ordering is redundant. On the contrary it is crucial for those theories whose generators do not commute, like QCD. This term explicitly depends on the path chosen to connect the two space-time points, thus introducing the same path-dependence to the correlator (2.115). The expression of the term (2.121) arises from the final state interactions [22]. Basically, after the active quark has been extracted from the hadron, all the remaining part interacts with it by exchanging soft gluons. The path ordering takes into account the exact order in which the gluons are exchanged. This can be seen by expanding the definition of the gauge-link in a Taylor series of the coupling constant g:

$$\begin{aligned}\mathcal{U}_{(x,y)} = \mathbb{1} &\\ &+ (ig)\int_x^y d\eta_1^\mu A_\mu(\eta_1)\\ &+ (ig)^2\int_x^y d\eta_1^\mu \int_x^{\eta_1} d\eta_2^\mu A_\mu(\eta_1)A_\mu(\eta_2)\\ &+ \dots,\end{aligned} \tag{2.122}$$

where we wrote explicitly the terms associated to the exchange of one and two gluons. In principle, the number of gluons exchanged is infinite, as Eq. (2.122) suggests. The schematic representation of the gauge link, seen as exchange of gluons, is reported in Fig. 2.4. At present, there is no evidence of a process from which the unintegrated correlator can be extracted phenomenologically. Instead, it is possible to get access to the k^--integrated correlator

$$\begin{aligned}&\Phi^{[\Gamma]}_{\Lambda\Lambda'}(x,\mathbf{k}_\perp,\Delta;P)\\ &= \int dk^-\,\Phi^{[\Gamma]}_{\Lambda\Lambda'}(k,\Delta;P)\\ &= \int \frac{d\zeta^- d^2\boldsymbol{\zeta}_\perp}{2(2\pi)^3}e^{i\left(k^+\zeta^- - \boldsymbol{\zeta}_\perp\cdot\mathbf{k}_\perp\right)}\left\langle p',\Lambda'\right|\bar{\Psi}^q(0)\mathcal{U}_{(0,\zeta)}\Gamma\Psi^q(\zeta)\left|p,\Lambda\right\rangle\Big|_{\zeta^+=0},\end{aligned} \tag{2.123}$$

where $x = k^+/P^+$ is the fraction of average longitudinal momentum carried by the active quark. This is the mathematical object at the basis of the most general parton distribution functions, which are the Generalized Transverse-Momentum Dependent Distribution functions (GTMDs) [18, 23, 24].

In principle, the choice of the path connecting 0 and ζ for the Wilson line depends on the physical process one wants to study. The condition $\zeta^+ = 0$ imposed in Eq. (2.123) brings us to consider the gauge-link just in the longitudinal direction and in the transverse direction:

$$\mathcal{U}^L(0^-, \zeta^-; \boldsymbol{\eta}_\perp) = \mathscr{P} \exp\Big[ig \int_{0^-}^{\zeta^-} d\eta^- A^+(0, \eta^-, \boldsymbol{\eta}_\perp)\Big], \tag{2.124}$$

$$\mathcal{U}^T(\mathbf{0}_\perp, \boldsymbol{\zeta}_\perp; \eta^-) = \mathscr{P} \exp\Big[ig \int_{\mathbf{0}_\perp}^{\boldsymbol{\zeta}_\perp} d^2\boldsymbol{\eta}_\perp \cdot \boldsymbol{A}_\perp(0, \eta^-, \boldsymbol{\eta}_\perp)\Big]. \tag{2.125}$$

The appropriate composition of these two gauge-links form the operator to be included in Eq. (2.123) for the specific process under consideration [25]. In case of Semi-Inclusive Deep Inelastic Scattering (SIDIS) and Drell-Yan (DY), the gauge-links are obtained as

$$\begin{cases} \mathcal{U}^L(0^-, +\infty^-; \mathbf{0}_\perp)\mathcal{U}^T(\mathbf{0}_\perp, \boldsymbol{\infty}_\perp; +\infty^-) & \\ \quad \times \mathcal{U}^T(\boldsymbol{\infty}_\perp, \boldsymbol{\zeta}_\perp; +\infty^-)\mathcal{U}^L(+\infty^-, \zeta^-; \boldsymbol{\zeta}_\perp) & \text{for SIDIS,} \\ \mathcal{U}^L(0^-, -\infty^-; \mathbf{0}_\perp)\mathcal{U}^T(\mathbf{0}_\perp, \boldsymbol{\infty}_\perp; -\infty^-) & \\ \quad \times \mathcal{U}^T(\boldsymbol{\infty}_\perp, \boldsymbol{\zeta}_\perp; -\infty^-)\mathcal{U}^L(-\infty^-, \zeta^-; \boldsymbol{\zeta}_\perp) & \text{for DY.} \end{cases} \tag{2.126}$$

By recalling the light-front gauge condition (2.64), we can practically ignore the longitudinal gauge-link, since $\mathcal{U}^L(0^-, \zeta^-; \boldsymbol{\eta}_\perp) \equiv \mathbb{1}$. On the contrary, the transverse gauge-link cannot be naively neglected. Indeed, if we use the advanced boundary conditions for the transverse components of the gauge fields, then $\mathcal{U}^T(\mathbf{0}_\perp, \boldsymbol{\zeta}_\perp; \eta^-) \equiv \mathbb{1}$, but the LFWFs acquire a complex phase that encodes the effects of the final state interactions [22, 26, 27]. This is the reason for which the LFWFs are treated as complex quantities in all the general theoretical expressions we will derive. Nevertheless, in the end, we will neglect the effects of the complex phase and we will present a model that parametrizes the LFWFs as pure real functions.

2.2.2 Landscape of parton distributions

All the parton distribution functions can be derived from the "mother" distributions, which are the GTMDs, by taking specific limits of the k^--integrated correlator (2.123). The GTMDs depend on the 3 variables, that span a six-dimensional space and can be grouped in the set $\mathcal{V} = \{x, \boldsymbol{k}_\perp, \Delta\}$. Consequently, each particular limit defines a function that depends on a specific subset of $\mathcal{V}$. Therefore, the number of possible distributions that can be obtained is the same as the number of possible subsets of $\mathcal{V}$, i.e., 8. The various subsets, defining the parton distribution functions are: $\mathcal{V}$, $\{x, \boldsymbol{k}_\perp\}$, $\{x, \Delta\}$,

$\{\boldsymbol{k}_\perp, \Delta\}$, $\{x\}$, $\{\Delta\}$, $\{\boldsymbol{k}_\perp\}$ and $\emptyset$. The axes in Fig. 2.5 represent the three physical variables of $\mathcal{V}$. The different parton distribution functions, that we introduce in the following, are represented as black points whose coordinates coincide with their variable dependence.

TMDs

The Transverse-Momentum Dependent Parton Distributions (TMDs) are obtained from the GTMDs by imposing $\Delta = 0$. The latter condition means that the TMDs can be accessed in those processes in which there is no momentum transfer between the initial and final states of the hadron, as in the case of semi-inclusive processes. Thanks to their dependence, these parton distribution functions can provide a three-dimensional picture of the hadron in the momentum space $(x, \boldsymbol{k}_\perp)$.

GPDs

The integral of the correlator (2.123) over the transverse momentum of the active parton $\boldsymbol{k}_\perp$ produces the Generalized Parton Distributions (GPDs), which are functions depending on the subset $(x, \Delta^+, \boldsymbol{\Delta}_\perp)$. We observe that the 2-dimensional variable $\boldsymbol{\Delta}_\perp$, representing the transfer of transverse momentum, can always be Fourier-transformed in the position parameter $\boldsymbol{b}_\perp$. Then, the GPDs can give a simultaneous description in both momentum and position space for the partons inside the hadron [28, 29].

TMFFs

The Transverse-Momentum Dependent Form Factors (TMFFs) can be obtained from the GTMDs by integrating over x. These functions depend on $(\boldsymbol{k}_\perp, \Delta)$, as they parametrize non-diagonal matrix elements in momentum space of local operators.

PDFs

The GPDs and the TMDs are completely independent distribution functions, since they are obtained from the GTMDs by taking different kinematic limits. However, they have a common limit to the collinear Parton Distribution Functions (PDFs). In fact, the PDFs can be obtained in two ways: either from the GPDs for $\Delta \to 0$, or from the integration over $\boldsymbol{k}_\perp$ of the TMDs [4]. PDFs are functions only on x, thus providing a probability density of extracting a parton with a fraction x of longitudinal momentum from the hadron.

[4] The integral over the transverse momentum cannot be literally taken since it requires a suitable regularization of the ultraviolet region, which leads to modifications of this simple picture. More details can be found in the literature, see, e.g., Ref. [30].

FFs

The Form Factors (FFs) can be obtained by integrating the GPDs over the longitudinal x and thus are functions only on the transfer momentum Δ. Another way to recover the FFs is to integrate the TMFFs over the transverse momentum $\boldsymbol{k}_\perp$. The form factors can be extracted only by considering exclusive processes, and reconstructing the momentum transfer from the final state.

TMSDs

Like the FFs and the PDFs, the other parton distribution functions depending only on one element of the set $\mathcal{V}$ are the Transverse-Momentum Dependent Spin Densities (TMSDs), which are functions of $\boldsymbol{k}_\perp$. These distributions can be obtained either from the TMFFs by imposing $\Delta = 0$, or from the integral over x of the TMDs.

Charges

The very last quantities we can obtain are the "Charges". These represent global properties of the hadrons, i.e., quantities that can be measured by looking at the hadron as a whole. They are obtained from the GTMDs by taking subsequently, in any order, the three following operations: $\int dx$, $\int d^2\boldsymbol{k}_\perp$, $\Delta = 0$.

All the relations linking the different parton distribution functions are reported in Fig. 2.5.

Wigner distributions

The Fourier transform of the GTMDs from momentum space to position space gives rise to a class of functions that are the Wigner distributions. They provide the most general one-body information of partons in both momentum and position space. The most general six-dimensional Wigner distributions were first introduced in Refs. [31, 32]. However, in this case the physical interpretation is plagued by relativistic corrections. This issue has been solved in the light-front formalism by integrating over the longitudinal spatial dimension, leading to five-dimensional phase-space distributions [24, 33, 34], depending on the set $(x, \boldsymbol{k}_\perp, \boldsymbol{b}_\perp)$.

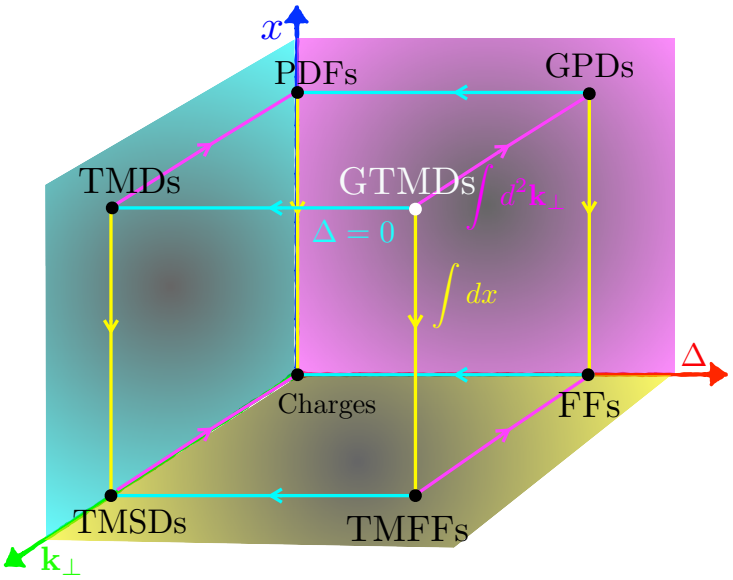


Figure 2.5: Representation of the projections of the GTMDs into parton distributions and form factors. The GTMDs, depending on x, $\boldsymbol{k}_\perp$ and Δ are represented as a white point. All the other distributions that are obtained from the GTMDs within proper limits are represented as black points. The cyan arrows give the forward limit in the hadron momentum, the pink arrows correspond to integrating over the quark transverse-momentum and the yellow arrows project out the longitudinal momentum of quarks. The variables on which the distributions depend coincide with their coordinates w.r.t. to the three axis x, Δ and $\boldsymbol{k}_\perp$.

The different parton distribution functions contain, in general, complementary information regarding the internal structure of the hadrons. The only effective way to hope to build a realistic picture of the hadrons is to combine these pieces of information by studying both the exclusive and inclusive processes through a phenomenological approach.

The final goal of this thesis is to develop a theoretical model to parametrize all the different types of parton distribution functions in the case of the pion. This will involve performing fits to experimental data for the FFs, conducting fits for observables sensitive to the PDFs and the TMDs, and then providing predictions for the GPDs.

2.3 Pion Distribution Amplitudes

The parton distribution functions are associated to certain probabilities of extracting a parton from the hadron with a specific momentum in the initial state and then re-inserting it back with a final momentum. In this sense, the parton distribution functions encode information about individual active

partons.

However, it is possible to introduce another class of objects, the so called Distribution Amplitudes (DAs), that enter the description of exclusive processes at large momentum transfer [35] and provide insights into the momentum fraction distributions of multiple partons in collinear directions simultaneously. In particular, they refer to a specific Fock state, with a fixed number of constituents, and they can be denoted as $d(x_1, x_2, \ldots, x_N)$, where x_i is the fraction of the collinear momentum of the i-th parton.

The DAs can be computed as hadron-to-vacuum matrix elements of non-local gauge-invariant operators built of quark and gluon fields at light-like separations. In the case of a pion with valence quark of flavour q and valence antiquark of flavour q' the matrix elements related to the leading-twist DA reads

$$\langle 0|\, \bar{\Psi}^{q'}(-z)\Gamma \mathcal{U}_{(-z,z)} \Psi^{q}(z)\, |\pi(p)\rangle\,, \tag{2.127}$$

where $\Gamma \in \mathcal{D}$ and $\mathcal{U}_{(-z,z)} \equiv \mathbb{1}$. The dependence of the DAs on the renormalization scale is understood.

As specified in Section 2.2.1, the twist order changes accordingly to the choice of Γ which involves the contribution of different components of the fields (good/bad). We refer here to the two-particle DAs at twist-2, representing the probability amplitude of finding in the pion a valence quark with light cone momentum fraction x and a valence antiquark with fraction $(1-x)$.

Such a DA, $d(x, 1-x) \equiv \phi_\pi(x)$, is produced by selecting the Dirac operator $\gamma^+\gamma_5$ [36, 37], according to

$$\phi_\pi(x) = \frac{1}{f_\pi} \int \frac{dz^-}{\pi} e^{iz^- p^+ (2x-1)} \langle 0|\, \bar{\Psi}^{q'}(-z)\gamma^+\gamma_5 \Psi^{q}(z)\, |\pi(p)\rangle\,, \tag{2.128}$$

where f_π is the pion decay constant, which guarantees the normalization

$$\int_0^1 \phi_\pi(x) dx = 1. \tag{2.129}$$

The leading-twist DAs for pions, kaons and ϕ mesons have been recently determined from lattice calculations [38, 39]. Moreover, there are different works on pion DAs up to twist-4 [35, 40–43].

2.4 Theoretical Model

This Section represents the core of the whole project as it outlines the steps needed to developing our theoretical model for parametrizing the various types of pion parton distribution functions.

As mentioned in Section 2.1.5, to recover the complete state of the pion, we should consider an infinite expansion of Fock states, but this is - in practice - impossible.

Therefore, we limit our expansion to the first leading orders, encompassing the components with 2 to 4 partons. Accordingly, the pion state is represented as follows

$$|\pi(P)\rangle = |\pi(P)\rangle_{q\bar{q}'} + |\pi(P)\rangle_{q\bar{q}'g} + |\pi(P)\rangle_{q\bar{q}'gg} + \sum_{\{\mathfrak{s}\bar{\mathfrak{s}}\}} |\pi(P)\rangle_{q\bar{q}'\{\mathfrak{s}\bar{\mathfrak{s}}\}}\,, \tag{2.130}$$

where $q\bar{q}' = u\bar{d}, d\bar{u}$, and the sum in $\{\mathfrak{s}\bar{\mathfrak{s}}\}$ runs over the N_f -flavour pairs of the sea quarks ($u\bar{u}$, $d\bar{d}$, and $s\bar{s}$ at the model scale). This choice is driven by the intention to parametrize all the pion PDFs effectively, including the valence PDF, the gluon PDF and the sea PDF, as outlined in Chapter 3.

To our knowledge, the expansion in the Fock space up to the two-gluon component represents the largest basis that has been used so far in light-front model calculations of the pion PDFs and e.m. form factor. The states in Eq. (2.130) represent the lowest-order contributions to fulfil this task. Although the gluon PDF could be computed by including only the $q\bar{q}'g$ state, the presence of $q\bar{q}'gg$ completes the states with 4 partons in the Fock expansion. The inclusion of additional Fock states with a higher number of partons is not expected to significantly increase the quality of the fits and the accessible information of parton distribution functions. Furthermore, such inclusion would significantly increase the level of complexity in the calculations. Therefore, limiting the expansion to these selected states proves to be a practical and effective approach for our purposes.

2.4.1 Pion Light Front Wave Amplitudes

For a generic hadron with helicity Λ, the angular momentum conservation is governed by the following constraint

$$\Lambda = l_z + \lambda, \tag{2.131}$$

where the total parton orbital angular momentum (OAM) l_z and the total parton helicity λ are summed over all the partons:

$$\begin{cases} l_z = \sum_{i=1}^{N} l_{zi} \\ \lambda = \sum_{i=1}^{N} \lambda_i \end{cases} . \tag{2.132}$$

Accordingly, each Fock component of the pion state can be classified in terms of l_z or, alternatively, in terms of λ, as outlined in Ref. [44].

This classification is performed in terms of Light Front Wave Amplitudes (LFWAs), which are derived from the LFWFs by factorizing the dependence on the discrete quantum numbers, such as colour, flavour, and helicity, and the $\boldsymbol{k}_\perp$-dependence, which explicitly corresponds to the OAM content of the Fock-state (for more details on the derivation, see Ref. [45]).

In the case of a pion with $\Lambda = 0$, the Fock expansion of Eq. (2.130) comprises 94 independent LFWAs, corresponding to all possible combinations of parton helicities. As it will be clarified in this Section, the model we are going to develop introduces approximately two parameters for each LFWA. If we were to consider all 94 LFWAs, we would introduce a set of around 200 parameters. Considering that the number of experimental data points to be fitted is approximately the same, and that the level of computational complexity would be significant, the fit would result practically unfeasible. For simplicity, we limit our study to consider only the projection on the $l_z = 0$ component, i.e.,

$$|\pi(p)\rangle^{l_z=0} = |\pi(p)\rangle^{l_z=0}_{q\bar{q}'} + |\pi(p)\rangle^{l_z=0}_{q\bar{q}'g} + |\pi(p)\rangle^{l_z=0}_{q\bar{q}'gg} + \sum_{\{\jmath\bar{\jmath}\}} |\pi(p)\rangle^{l_z=0}_{q\bar{q}'\{\jmath\bar{\jmath}\}} \,. \tag{2.133}$$

Among all the independent LFWAs resulting from the projection in Eq. (2.133), we choose to neglect those multiplied by factors dependent on the parton transverse momenta. This is because we consistently neglect the contributions from higher-twist DAs which enter the parametrization of these LFWAs as well as the LFWAs with $|l_z| = 1$. In this way, we arrive at a set of 7 independent LFWAs, facilitating a more straightforward and direct connection between the LFWAs and the pion DAs, as it will be discussed in Section 2.4.2.

We report here the four Fock-state contributions including the final 7 independent LFWAs:

$$|\pi(p)\rangle^{l_z=0}_{q\bar{q}'} = \int \mathrm{d}[1]\mathrm{d}[2] \frac{\delta_{c_1c_2}}{\sqrt{3}} \psi^{(1)}_{q\bar{q}'}(1,2) \left[q^\dagger_{c_1\uparrow}(w_1) \bar{q}'^\dagger_{c_2\downarrow}(w_2) - q^\dagger_{c_1\downarrow}(w_1) \bar{q}'^\dagger_{c_2\uparrow}(w_2) \right] |0\rangle \,, \tag{2.134}$$

$$|\pi(p)\rangle^{l_z=0}_{q\bar{q}'g} = \int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3] \frac{T^a_{c_1c_2}}{2} \psi^{(1)}_{q\bar{q}'g}(1,2,3) \times \left[(q\bar{q}')^\dagger_{A,1}\, g^\dagger_{a\downarrow}(w_3) - (q\bar{q}')^\dagger_{A,-1}\, g^\dagger_{a\uparrow}(w_3) \right] |0\rangle \,, \tag{2.135}$$

$$|\pi(p)\rangle^{l_z=0}_{q\bar{q}'gg} = \int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3]\mathrm{d}[4] \frac{\delta_{c_1c_2}\delta^{ab}}{\sqrt{24}} \Big\{ \psi^{(1)}_{q\bar{q}'gg}(1,2,3,4)\, (q\bar{q}')^\dagger_{A,0}\, (gg)^\dagger_{S,0} + \psi^{(2)}_{q\bar{q}'gg}(1,2,3,4)\, (q\bar{q}')^\dagger_{S,0}\, (gg)^\dagger_{A,0} \Big\} |0\rangle \,, \tag{2.136}$$

$$|\pi(p)\rangle^{l_z=0}_{q\bar{q}'\{\jmath\bar{\jmath}\}} = \int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3]\mathrm{d}[4] \frac{\delta_{c_1c_2}\delta_{c_3c_4}}{3} \Big\{ \psi^{(1)}_{q\bar{q}'\jmath\bar{\jmath}}(1,2,3,4)\, (q\bar{q}')^\dagger_{A,0}\, (\jmath\bar{\jmath})^\dagger_{S,0} + \psi^{(2)}_{q\bar{q}'\jmath\bar{\jmath}}(1,2,3,4)\, (q\bar{q}')^\dagger_{S,0}\, (\jmath\bar{\jmath})^\dagger_{A,0}$$

$$+\,\psi^{(3)}_{q\bar{q}'s\bar{s}}\,(1,2,3,4)\left[(q\bar{q}')^{\dagger}_{A,1}\,(s\bar{s})^{\dagger}_{A,-1}-(q\bar{q}')^{\dagger}_{A,-1}\,(s\bar{s})^{\dagger}_{A,1}\right]\Bigg\}\,|0\rangle\,. \tag{2.137}$$

These states contain the following operators:

$$(q\bar{q}')^{\dagger}_{S,0}=q^{\dagger}_{c_1\uparrow}(w_1)\bar{q}'^{\dagger}_{c_2\downarrow}(w_2)+q^{\dagger}_{c_1\downarrow}(w_1)\bar{q}'^{\dagger}_{c_2\uparrow}(w_2), \tag{2.138}$$

$$(s\bar{s})^{\dagger}_{S,0}=s^{\dagger}_{c_3\uparrow}(w_3)\bar{s}^{\dagger}_{c_4\downarrow}(w_4)+s^{\dagger}_{c_3\downarrow}(w_3)\bar{s}^{\dagger}_{c_4\uparrow}(w_4), \tag{2.139}$$

$$(q\bar{q}')^{\dagger}_{A,0}=q^{\dagger}_{c_1\uparrow}(w_1)\bar{q}'^{\dagger}_{c_2\downarrow}(w_2)-q^{\dagger}_{c_1\downarrow}(w_1)\bar{q}'^{\dagger}_{c_2\uparrow}(w_2), \tag{2.140}$$

$$(s\bar{s})^{\dagger}_{A,0}=s^{\dagger}_{c_3\uparrow}(w_3)\bar{s}^{\dagger}_{c_4\downarrow}(w_4)-s^{\dagger}_{c_3\downarrow}(w_3)\bar{s}^{\dagger}_{c_4\uparrow}(w_4), \tag{2.141}$$

$$(q\bar{q}')^{\dagger}_{A,1}=q^{\dagger}_{\uparrow c_1}(w_1)\bar{q}'^{\dagger}_{c_2\uparrow}(w_2), \tag{2.142}$$

$$(s\bar{s})^{\dagger}_{A,1}=s^{\dagger}_{\uparrow c_3}(w_3)\bar{s}^{\dagger}_{c_4\uparrow}(w_4), \tag{2.143}$$

$$(q\bar{q}')^{\dagger}_{A,-1}=q^{\dagger}_{\downarrow c_1}(w_1)\bar{q}'^{\dagger}_{c_2\downarrow}(w_2), \tag{2.144}$$

$$(s\bar{s})^{\dagger}_{A,-1}=s^{\dagger}_{\downarrow c_3}(w_3)\bar{s}^{\dagger}_{c_4\downarrow}(w_4), \tag{2.145}$$

$$(gg)^{\dagger}_{S,0}=g^{\dagger}_{a\uparrow}(w_3)g^{\dagger}_{b\downarrow}(w_4)+g^{\dagger}_{a\downarrow}(w_3)g^{\dagger}_{b\uparrow}(w_4), \tag{2.146}$$

$$(gg)^{\dagger}_{A,0}=g^{\dagger}_{a\uparrow}(w_3)g^{\dagger}_{b\downarrow}(w_4)-g^{\dagger}_{a\downarrow}(w_3)g^{\dagger}_{b\uparrow}(w_4). \tag{2.147}$$

The model we are developing is based on the parametrization of the 7 LFWAs: $\psi^{(1)}_{q\bar{q}'}$, $\psi^{(1)}_{q\bar{q}'g}$, $\psi^{(1)}_{q\bar{q}'gg}$, $\psi^{(2)}_{q\bar{q}'gg}$, $\psi^{(1)}_{q\bar{q}'s\bar{s}}$, $\psi^{(2)}_{q\bar{q}'s\bar{s}}$, $\psi^{(3)}_{q\bar{q}'s\bar{s}}$. From now on, we will denote the general LFWA with $\psi^{(i)}_{\beta}\equiv\psi^{(i)}_{\beta}\,(1,2,\ldots,N)$, where β is a label referring to the parton composition $\{q\bar{q}',q\bar{q}'g,q\bar{q}'gg,q\bar{q}'s\bar{s}\}$, while the integers $i=1,\ldots,N$ are collective indices referring to $(x_i,\boldsymbol{k}_{\perp i})$. These collective indices distinguish the relative kinematics of the LFWAs from the arguments of the ladder operators $w_i=(p^+_i,\boldsymbol{p}_{\perp i})$, with the transverse variables linked via Eq. (2.84), and $p^+_i=x_ip^+$. The integration measure in Eqs. (2.134) - (2.137) is defined in terms of the measure in Eq. (2.86).

From a mathematical point of view, the LFWAs are functions depending only on the fractions of momentum in the collinear direction x_i and on the relative transverse momenta $\boldsymbol{k}_{\perp i}$ of the partons. Without loss of generality [5], it is always possible to express $\psi^{(i)}_{\beta}$ in the following general functional form

$$\psi^{(i)}_{\beta}(x_i,\boldsymbol{k}_{\perp i})=\phi^{(i)}_{\beta}(x_i)\Omega^{(i)}_{\beta}(x_i,\boldsymbol{k}_{\perp i}). \tag{2.148}$$

The model dependence enters in the next step, by specifying the analytic form of the functions $\phi^{(i)}_{\beta}$ and $\Omega^{(i)}_{\beta}$.

2.4.2 Modelling the LFWAs

The very general expression of Eq. (2.148) is here specified, on the basis of the following rationale. The pion DAs introduced in Section (2.3), as well as the

[5] In the most trivial case $\phi^{(i)}_{\beta}\equiv 1$ and $\Omega^{(i)}_{\beta}\equiv\psi^{(i)}_{\beta}$.

functions $\phi_\beta^{(i)}$, depend solely on the fraction of momentum x_i carried by the partons in the collinear direction. Our aim is to establish a clear and direct correspondence between the pion DAs and $\phi_\beta^{(i)}$, once fixed the twist order and the number of constituent partons. By doing so, the functions $\phi_\beta^{(i)}$ acquire a non-trivial physical meaning. This connection between the LFWAs and the pion DAs facilitates a deeper understanding and meaningful interpretation of the LFWAs in the context of pion structure and dynamics.

To achieve this goal, we need to compute the matrix elements of Eq. (2.127), for the Fock-states in Eqs. (2.134) - (2.137). The most general result of these calculations, which are intended to be valid at the level of bare operators, can be summarized in relations of the following form

$$\int \left[d^2\boldsymbol{k}_\perp\right]_N \psi_\beta^{(i)}\left(1,2,\ldots,N\right) = \sum_j a_{ij} d_j(x_1,...,x_N), \tag{2.149}$$

where $d_j(x_1,...,x_N)$ are the pion DAs. As evident from Eq. (2.149), the correspondences between the $\phi_\beta^{(i)}$ and the DAs are affected by the presence of the integrals in the transverse measures of the functions $\Omega_\beta^{(i)}$ contained in $\psi_\beta^{(i)}$. Therefore, the model we are going to introduce must satisfy the following condition:

1. the integrals of the functions $\Omega_\beta^{(i)}$ in the transverse measure, for all β, are finite multiplicative factors.

Let us suppose, for the moment, that these multiplicative factors are all equal to 1, i.e. that the $\Omega_\beta^{(i)}$ are normalized to unity:

$$\int \left[d^2\boldsymbol{k}_\perp\right]_N \Omega_\beta^{(i)}\left(x_1,\boldsymbol{k}_{\perp 1},x_2,\boldsymbol{k}_{\perp 2},\ldots,x_N,\boldsymbol{k}_{\perp N}\right) = 1. \tag{2.150}$$

Then, the left-hand side of Eq. (2.149) reduces to:

$$\int \left[d^2\boldsymbol{k}_\perp\right]_N \psi_\beta^{(i)}\left(1,2,\ldots,N\right) = \phi_\beta^{(i)}(x_1,...,x_N) \tag{2.151}$$

and we obtain the direct identification $\phi_\beta^{(i)}(x_1,...,x_N) = \sum_j a_{ij} d_j(x_1,...,x_N)$. Such a normalization condition is always possible, if we do not assume specific boundary conditions for $\phi_\beta^{(i)}$. In fact, given Eq. (2.149), we can always introduce $\phi_\beta^{(i)}(x_1,...,x_N) = \sum_j a_{ij} d_j(x_1,...,x_N)$ and write

$$\psi_\beta^{(i)}(x_1,\boldsymbol{k}_{\perp 1},...,x_N,\boldsymbol{k}_{\perp N}) = \left(\frac{\psi_\beta^{(i)}(x_1,\boldsymbol{k}_{\perp 1},...,x_N,\boldsymbol{k}_{\perp N})}{\phi_\beta^{(i)}(x_1,...,x_N)}\right)\phi_\beta^{(i)}(x_1,...,x_N). \tag{2.152}$$

The function in brackets has exactly the right dependence on the arguments to be identified with $\Omega_\beta^{(i)}(x_1,\boldsymbol{k}_{\perp 1},...,x_N,\boldsymbol{k}_{\perp N})$, and it satisfies the normalization (2.150).

In the case where $\Omega_\beta^{(i)}$ is not normalized to 1, the relationship between $\phi_\beta^{(i)}$ and the DAs would change only by factors corresponding to the norms of $\Omega_\beta^{(i)}$. However, the general structure of Eq. (2.149) would remain unaffected.

To better clarify this abstract reasoning, we present the explicit result at twist-2 in Section 2.4.3, where we connect our $\phi^{(1)}_{q\bar{q}'}(x_1, 1-x_1)$ to $d_2(x_1, 1-x_1)$. The latter coincides to the already mentioned $\phi_\pi(x_1)$ reported in Eq. (2.128).

Model for the transverse-momentum dependence

There are many ways to introducing functions depending on x_i and $\boldsymbol{k}_{\perp i}$, such that the requirement of finite integrals in transverse direction is satisfied. The Brodsky-Lepage-Huang prescription [46] achieves this with the following dependency

$$\Omega_\beta^{(i)}\left(x_1, \boldsymbol{k}_{\perp 1}, x_2, \boldsymbol{k}_{\perp 2}, \ldots, x_N, \boldsymbol{k}_{\perp N}\right) = \frac{\left(16\pi^2 a_\beta^{(i)2}\right)^{N-1}}{\prod_{i=1}^{N} x_i} \exp\left(-a_\beta^{(i)2} \sum_{i=1}^{N} \frac{\boldsymbol{k}_{\perp i}^2}{x_i}\right), \tag{2.153}$$

where $a_\beta^{(i)}$ are free parameters that need to be fitted. Since these $a_\beta^{(i)}$ are associated with the part of the LFWAs that depends on the transverse momentum, they collectively form the set of "transverse parameters", hereafter denoted as $\mathcal{A}^T$. More explicitly:

$$\mathcal{A}^T = \{a^{(1)}_{q\bar{q}'}, a^{(1)}_{q\bar{q}'g}, a^{(1)}_{q\bar{q}'gg}, a^{(2)}_{q\bar{q}'gg}, a^{(1)}_{q\bar{q}'s\bar{s}}, a^{(2)}_{q\bar{q}'s\bar{s}}, a^{(3)}_{q\bar{q}'s\bar{s}}\}, \tag{2.154}$$

where

$a^{(1)}_{q\bar{q}'}$ is the transverse parameter for the $q\bar{q}'$ state,
$a^{(1)}_{q\bar{q}'g}$ is the transverse parameter for the $q\bar{q}'g$ state,
$a^{(i=1,2)}_{q\bar{q}'gg}$ are the transverse parameters for the $q\bar{q}'gg$ independent states,
$a^{(i=1,2,3)}_{q\bar{q}'s\bar{s}}$ are the transverse parameters for the $q\bar{q}'s\bar{s}$ independent states.

The functions in Eq. (2.153) satisfy the normalization condition in Eq. (2.150). Moreover, the integrals of their squares in the transverse direction give the following result

$$\int \left[d^2\boldsymbol{k}_\perp\right]_N \Omega_\beta^{(i)2} = \frac{\left(8\pi^2 a_\beta^{(i)2}\right)^{N-1}}{\prod_{i=1}^{N} x_i}. \tag{2.155}$$

This latter property poses an issue for our intended approach. As we will discuss in Section 3, the collinear parton distribution functions can be expressed in terms of integrals of the squares of the LFWAs. However, the presence of

the transverse parameters in the r.h.s. of Eq. (2.155) implies that the collinear PDFs are parametrized in terms of both the parameters in the longitudinal-momentum part of the LFWAs and the transverse set $\mathscr{A}^T$. This contradicts our goal, as we aim to develop a model where the collinear PDFs are solely parametrized by a set of "longitudinal parameters", while the transverse parameters enter only at the level of those parton distribution functions dependent on the transverse-momentum direction, such as the TMDs, the FFs and the GPDs.

Therefore, for our purposes, it is necessary to modify the model in Eq. (2.153) in a manner that eliminates any dependence on $\mathscr{A}^T$ in the integrals like (2.155).

To achieve this, we adopt the following expression for the functions $\Omega_\beta^{(i)}$

$$\Omega_\beta^{(i)}\left(x_1, \boldsymbol{k}_{\perp 1}, x_2, \boldsymbol{k}_{\perp 2}, \ldots, x_N, \boldsymbol{k}_{\perp N}\right) = \frac{\left(4\sqrt{2}\pi a_\beta^{(i)}\right)^{N-1}}{\prod_{i=1}^{N} x_i} \exp\left(-a_\beta^{(i)2} \sum_{i=1}^{N} \frac{\boldsymbol{k}_{\perp i}^2}{x_i}\right), \tag{2.156}$$

which implies the following normalizations

$$\int [d^2\boldsymbol{k}_\perp]_N \Omega_\beta^{(i)} = \frac{1}{\left(2\sqrt{2}\pi a_\beta^{(i)}\right)^{N-1}}, \tag{2.157}$$

$$\int [d^2\boldsymbol{k}_\perp]_N \Omega_\beta^{(i)2} = \frac{1}{\prod_{i=1}^{N} x_i}. \tag{2.158}$$

The latter equation guarantees that the collinear PDFs will not depend on the transverse set $\mathscr{A}^T$, while Eq. (2.157) ensures finite normalizations for the functions $\Omega_\beta^{(i)}$. We notice that the normalization constants on the r.h.s. of Eq. (2.157) now explicitly depend on the transverse parameters. This implies that the relationships we obtain between the DAs and our collinear dependency will exhibit a dependence on the parameters $a_\beta^{(i)}$ through global normalization factors.

There is no way around this problem: either the DAs or the PDFs will inevitably have some degree of dependence on the transverse parameters a_β. This means that there will always be traces of the transverse structure in the collinear part. We emphasize that our choice is to have the collinear PDFs completely independent of $\mathscr{A}^T$, which is exactly the meaning of Eq. (2.158).

Summarizing, the second condition our model fulfils is:

2. the integrals of the functions $\Omega_\beta^{(i)2}$ over the transverse measure must not depend on the set of parameters used to model the functions $\Omega_\beta^{(i)}$.

Having established the parametrization of the functions $\Omega_\beta^{(i)}$, we can now explain how we model the pure collinear part.

Model for the longitudinal momentum-fraction dependence

Given that the integrals of the $\Omega_\beta^{(i)}$ are just finite factors (see Eq. (2.157)[6]), Eq. (2.149) provides the connection between the functions $\phi_\beta^{(i)}$ and the DAs. For this reason, the third and last condition we require is:

3. the x_i-dependence of the functions $\phi_\beta^{(i)}$ is inferred from the pion DAs.

Following the analysis of Ref. [43], one can construct the expansion of the leading-twist pion DAs in the representations of the collinear subgroup of the conformal transformations on the light-cone. We simplify this expansion by limiting it to the leading contribution, which corresponds with the lowest conformal spin and is referred to as the asymptotic DAs. For a generic N-parton state, these asymptotic DAs are expressed as follows:

$$\prod_{i=1}^{N} x_i^{2j_i-1}, \tag{2.159}$$

where j_i is the conformal spin of the i-th parton, which is $j_i = 1$ for quarks and anti-quarks and $j_i = 3/2$ for gluons. In this way, the functions $\phi_\beta^{(i)}$, for all the Fock states, take the following form

$$\begin{aligned}
\phi^{(1)}_{q\bar{q}'}(x_1,x_2) &= N^{(1)}_{q\bar{q}'}x_1x_2\\
\phi^{(1)}_{q\bar{q}'g}(x_1,x_2,x_3) &= N^{(1)}_{q\bar{q}'g}x_1x_2x_3^2\\
\phi^{(1)}_{q\bar{q}'gg}(x_1,x_2,x_3,x_4) &= N^{(1)}_{q\bar{q}'gg}x_1x_2(x_3x_4)^2,\\
\phi^{(2)}_{q\bar{q}'gg}(x_1,x_2,x_3,x_4) &= N^{(2)}_{q\bar{q}'gg}x_1x_2x_3x_4(x_3^2-x_4^2),\\
\phi^{(1)}_{q\bar{q}'s\bar{s}}(x_1,x_2,x_3,x_4) &= N^{(1)}_{q\bar{q}'s\bar{s}}x_1x_2x_3x_4,\\
\phi^{(2)}_{q\bar{q}'s\bar{s}}(x_1,x_2,x_3,x_4) &= N^{(2)}_{q\bar{q}'s\bar{s}}x_1x_2x_3x_4,\\
\phi^{(3)}_{q\bar{q}'s\bar{s}}(x_1,x_2,x_3,x_4) &= N^{(3)}_{q\bar{q}'s\bar{s}}x_1x_2x_3x_4.
\end{aligned} \tag{2.160}$$

We observe the different functional form of $\phi^{(1)}_{q\bar{q}'gg}$ and that of $\phi^{(2)}_{q\bar{q}'gg}$ which results from the different symmetry property of the corresponding LFWAs under the exchange of the last two arguments pertaining to the gluons. Furthermore, the three functions $\phi^{(1)}_{q\bar{q}'s\bar{s}}$, $\phi^{(2)}_{q\bar{q}'s\bar{s}}$, and $\phi^{(3)}_{q\bar{q}'s\bar{s}}$ associated with the Fock-state in Eq. (2.137) correspond to different helicity configurations of four partons and are all symmetric under the exchange of the last two entries. As long as we consider unpolarized PDFs with the extant database, we lack sensitivity to these distinct helicity configurations. Consequently, we make the assumption that these three functions share the same dependence on the longitudinal momentum fractions, albeit with different normalization factors. This assumption simplifies the parametrization and reduces the complexity of the model while still maintaining adequate accuracy for our specific purposes.

[6]Under the hypothesis $a_\beta^{(i)} \neq 0$.

Unfortunately the parametrization (2.160) leads to poor results for the fit of the PDFs. For this reason we decided to change the asymptotic expansion for the states $q\bar{q}'$ and $q\bar{q}'g$ by including the first beyond-asymptotic term in the expansion. Moreover, for the $q\bar{q}'$ state, we modified also the exponent of the asymptotic expansion, by introducing a variable exponent γ_q to be fitted to data.

The choice to include only the first term beyond the asymptotic expansion and solely for the first two Fock states was guided by the desire to avoid introducing too many free parameters into the model, considering the limited number of available data points for the fits. We also experimented with varying the exponents of the asymptotic states, and it was observed that the quality of the fit remained mostly unchanged, even when introducing other variable exponents for different Fock states.

Actually, we performed various tests by including additional terms (up to 2 beyond the asymptotic ones for each single Fock state). We also made other trials by introducing additional variable exponents for different Fock states, but all these attempts did not significantly impact the quality of the fit.

The next question about the modelling of the parametrization for the DAs is: *How did you introduce the next-to-asymptotic terms?*

For each Fock-state component, we seek a suitable complete set of orthonormal polynomials $p_m(x_1, x_2, \cdots, x_N)$ of degree m to perform the expansion of the DAs. This set should contain the asymptotic term and therefore satisfy the following orthonormality conditions

$$\begin{cases} \int [dx]_N \prod_{i=1}^{N} \sqrt{x_i}\, x_i^{2j_i-1} p_m(x_1, x_2, \ldots, x_N) = 0, \\ \int [dx]_N \prod_{i=1}^{N} \sqrt{x_i}\, p_m(x_1, x_2, \ldots, x_N) p_l(x_1, x_2, \ldots, x_N) = 0 \text{ for } m \neq l. \end{cases} \tag{2.161}$$

In the case of $q\bar{q}'g$, the relevant polynomials are a combination of the Jacobi polynomials. For the $q\bar{q}'$ state with varied asymptotic exponent, the first equation in (2.161) involves $x_1^{\gamma_q} x_2^{\gamma_q}$ instead of $x_1 x_2$, and the required polynomials are a combination of Gegenbauer polynomials with variable dimensionality. The choice to employ a variable exponent was motivated by the observation that an expansion in Gegenbauer polynomials with variable dimensionality typically exhibits faster convergence than an expansion with fixed dimensionality [47].

More specifically, the desired polynomials for $q\bar{q}'$ and $q\bar{q}'g$ are as follows

$$p_m(x_1, x_2) = \frac{C_m^{\gamma_q+\frac{1}{2}}(x_1) + C_m^{\gamma_q+\frac{1}{2}}(x_2)}{2}\,, \tag{2.162}$$

$$p_m(x_1, x_2, x_3) = (1-x_3)^k \mathscr{P}_{m-2k}^{(2,2k-3)}(1-2x_3)\, \frac{\mathscr{P}_k^{(1,1)}\left(\frac{x_2-x_1}{1-x_3}\right) + \mathscr{P}_k^{(1,1)}\left(\frac{x_1-x_2}{1-x_3}\right)}{2}. \tag{2.163}$$

In the previous equations, $C_j^{\alpha}(x)$ are Gegenbauer polynomials and $\mathscr{P}_j^{(\alpha,\beta)}(x)$ are Jacobi polynomials. The index $k \in \mathbb{N}$ in Eq. (2.163) satisfies $k \leq \frac{m}{2}$, and this is due to the conservation of total angular momentum. In fact, since we are considering the pion state with the projection of OAM $l_z = 0$ and since the pion is a spinless hadron, only the conformal spins of the partons are relevant. The total contribution of angular momentum can be decomposed as $J = j_1 + j_2 + j_3 + n = \frac{7}{2} + n$, with $n \in \mathbb{N}$. For the degeneracy of the quantum number, we have $j = j_1 + j_2 + j_3 + k$ with $0 \leq k \leq n$. By setting $n = m - k$, we precisely arrive at $k \leq \frac{m}{2}$, where $m \in \mathbb{N}$. The coefficients that appear in the Jacobi polynomials in Eq. (2.163) are, in fact, combinations of the conformal spins of the partons.

It is noteworthy that in both (2.162) and (2.163) the only integers m that render those expressions non-vanishing are the even numbers.

On the basis of all the previous discussion, the final model for the collinear part of the LFWAs is:

$$\phi^{(1)}_{q\bar{q}'}(x_1, x_2) = N^{(1)}_{q\bar{q}'}(x_1 x_2)^{\gamma_q}\left(1 + d_{q1} C_2^{(\gamma_q+1/2)}(x_1 - x_2)\right) \tag{2.164}$$

$$\phi^{(1)}_{q\bar{q}'g}(x_1, x_2, x_3) = N^{(1)}_{q\bar{q}'g} x_1 x_2 x_3^2 \left[1 + d_{g1}(3 - 7x_3)\right], \tag{2.165}$$

$$\phi^{(1)}_{q\bar{q}'gg}(x_1, x_2, x_3, x_4) = N^{(1)}_{q\bar{q}'gg} x_1 x_2 (x_3 x_4)^2, \tag{2.166}$$

$$\phi^{(2)}_{q\bar{q}'gg}(x_1, x_2, x_3, x_4) = N^{(2)}_{q\bar{q}'gg} x_1 x_2 x_3 x_4 (x_3^2 - x_4^2), \tag{2.167}$$

$$\phi^{(1)}_{q\bar{q}'s\bar{s}}(x_1, x_2, x_3, x_4) = N^{(1)}_{q\bar{q}'s\bar{s}} x_1 x_2 x_3 x_4, \tag{2.168}$$

$$\phi^{(2)}_{q\bar{q}'s\bar{s}}(x_1, x_2, x_3, x_4) = N^{(2)}_{q\bar{q}'s\bar{s}} x_1 x_2 x_3 x_4, \tag{2.169}$$

$$\phi^{(3)}_{q\bar{q}'s\bar{s}}(x_1, x_2, x_3, x_4) = N^{(3)}_{q\bar{q}'s\bar{s}} x_1 x_2 x_3 x_4, \tag{2.170}$$

The norms $N^{(i)}_{\beta}$ in Eqs. (2.164) - (2.170) are defined in such a way as to guarantee the normalization of the pion state which reads

$$\langle \pi(p') | \pi(p) \rangle = 2(2\pi)^3 p^+ \delta\left(p'^+ - p^+\right) \delta\left(\boldsymbol{p}'_\perp - \boldsymbol{p}_\perp\right). \tag{2.171}$$

Accordingly, they are given by

$$\begin{aligned} N^{(1)}_{q\bar{q}'} = \cos(\alpha_1) 4^{\gamma_q} \sqrt{\frac{\Gamma\left(\frac{5}{2} + 2\gamma_q\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(2\gamma_q\right)}} \\ \Big\{ (4\gamma_q + 3)[1 + 4\gamma_q + 2d_{q1}(1 - \gamma_q)(1 + 2\gamma_q)] \\ + d_{q1}^2 (1 + 2\gamma_q)^2 (3 + 4\gamma_q + 3\gamma_q^2) \Big\}^{-1/2}, \end{aligned} \tag{2.172}$$

$$N^{(1)}_{q\bar{q}'g} = 6\sqrt{210}\sin(\alpha_1)\cos(\alpha_2)\left(18 - 18 d_{g1} + 29 d_{g1}^2\right)^{-1/2}, \tag{2.173}$$

$$N^{(1)}_{q\bar{q}'gg} = 2\sqrt{\frac{7}{6}} N^{(2)}_{q\bar{q}'gg} = 30\sqrt{77}\sin(\alpha_1)\sin(\alpha_2)\cos(\alpha_3), \tag{2.174}$$

$$N^{(1)}_{q\bar{q}'s\bar{s}} = N^{(2)}_{q\bar{q}'s\bar{s}} = \frac{1}{\sqrt{2}} N^{(3)}_{q\bar{q}'s\bar{s}} = 2\sqrt{35}\sin(\alpha_1)\sin(\alpha_2)\sin(\alpha_3), \tag{2.175}$$

where $\Gamma(x)$ is the Euler Gamma function.

The free parameters we introduced for the collinear part can be grouped in a set which we refer to as the "longitudinal set" of parameters and denote with $\mathcal{A}^L$ to distinguish it from $\mathcal{A}^T$. The longitudinal set comprises the following parameters:

$$\mathcal{A}^L = \{\gamma_q, d_{q1}, d_{g1}, \alpha_1, \alpha_2, \alpha_3\}, \tag{2.176}$$

where

γ_q is the varied asymptotic exponent for the state $q\bar{q}'$;
d_{q1} is the weight of the first term beyond the asymptotic one for $q\bar{q}'$;
d_{g1} is the weight of the first term beyond the asymptotic one for $q\bar{q}'g$;
$\alpha_1, \alpha_2, \alpha_3$ regulate the relative weights of the different Fock components.

Summarizing, the model we have implemented so far consists of two sets of parameters, $\mathcal{A}^L$ and $\mathcal{A}^T$, with a total number of 13 parameters. The underlying logic is to first fix the set of longitudinal parameters by fitting the pion PDFs. Once these parameters have been determined, we can use them as inputs in the definitions of other parton distribution functions or form factors that also depend on the transverse set. Subsequently, the transverse parameters can be determined by fitting the experimental data related to these functions. This sequential approach allows us to efficiently and effectively determine the optimal values for both sets of parameters and achieve a comprehensive description of the parton dynamics in the pion.

2.4.3 Explicit link between $\phi^{(1)}_{q\bar{q}'}$ and ϕ_π

In this Section, we work out the relation in Eq. (2.149) for the leading-twist pion DA.

Throughout the calculation, we will obtain a model-independent equation that connects the two-partons LFWA to the pion DA at twist-2 (see Eq.(2.193)). Subsequently, we will proceed to detail the model introduced in Section 2.4.2, thus obtaining the desired relation between $\phi^{(1)}_{q\bar{q}'}$ and ϕ_π.

We observe that Eq. (2.128) is completely equivalent[7] to the relation:

$$f_\pi p^+ \int_0^1 dx' e^{ip^+ z^-(2x'-1)} \phi_\pi(x') = \langle 0| \bar{\Psi}^{q'}(z)\gamma^+\gamma_5 \Psi^q(-z) |\pi(p)\rangle . \tag{2.177}$$

By considering this form, we can insert the definition of the pion state with 2 partons of Eq. (2.134) into the right-hand side. We now show that the operator

[7] This statement can be proven by multiplying both sides of Eq. (2.177) by $e^{ip^+ z^-(2x-1)}$ and by taking the integrals $\int dz^-$.

$\gamma^+\gamma_5$ automatically selects the good components of the fields. Starting from:

$$\langle 0|\,\bar{\Psi}^{q'}(z)\gamma^+\gamma_5\Psi^q(-z)\,|\pi(p)\rangle \tag{2.178}$$

and using the following properties of the Dirac matrices

$$\gamma^\mu\gamma_5 = -\gamma_5\gamma^\mu, \tag{2.179}$$
$$\gamma^\pm\gamma^\pm = 0, \tag{2.180}$$
$$\gamma^0\gamma^+ = \gamma^-\gamma^0, \tag{2.181}$$

we can write:

$$\begin{aligned}
&\langle 0|\bar{\Psi}^{q'}(z)\gamma^+\gamma_5\Psi^q(-z)\,|\pi(p)\rangle \\
&\quad= \langle 0|\,\Psi^{q'\,\dagger}(z)\gamma^0\mathbb{1}\gamma^+\gamma_5\mathbb{1}\Psi^q(-z)\,|\pi(p)\rangle \\
&\quad= \langle 0|\,\Psi^{q'\,\dagger}(z)\gamma^0\frac{1}{2}(\gamma^+\gamma^- + \gamma^-\gamma^+)\gamma^+\gamma_5\frac{1}{2}(\gamma^+\gamma^- + \gamma^-\gamma^+)\Psi^q(-z)\,|\pi(p)\rangle \\
&\quad= \langle 0|\,\Psi^{q'\,\dagger}(z)\gamma^0\frac{1}{2}\gamma^+\gamma^-\gamma^+\gamma_5\frac{1}{2}\gamma^-\gamma^+\Psi^q(-z)\,|\pi(p)\rangle \\
&\quad= \langle 0|\,\Psi^{q'\,\dagger}(z)\frac{1}{2}\gamma^-\gamma^+\gamma^0\gamma^+\gamma_5\frac{1}{2}\gamma^-\gamma^+\Psi^q(-z)\,|\pi(p)\rangle \\
&\quad= \langle 0|\,\Psi^{q'\,\dagger}(z)\mathscr{P}_+\gamma^0\gamma^+\gamma_5\mathscr{P}_+\Psi^q(-z)\,|\pi(p)\rangle \\
&\quad= \langle 0|\,\bar{\Psi}^{q'+}(z)\gamma^+\gamma_5\Psi^{q+}(-z)\,|\pi(p)\rangle\,.
\end{aligned} \tag{2.182}$$

We now consider the matrix element (2.178) and insert the Fourier expansion for the fields with the good components of Eqs. (2.108) - (2.109). This leads to the following four contributions

$$\sum_{\lambda,\lambda'}\sum_{c,c'} u^{+\,\dagger}_{\lambda'}(w')\gamma^0\gamma^+\gamma_5 u^+_\lambda(w) q'^\dagger_{c'\lambda'}(w') q_{c\lambda}(w) e^{i(k'^+ + k^+)z^-}\delta_{cc'}, \tag{2.183}$$

$$\sum_{\lambda,\lambda'}\sum_{c,c'} u^{+\,\dagger}_{\lambda'}(w')\gamma^0\gamma^+\gamma_5 v^+_\lambda(w) q'^\dagger_{c'\lambda'}(w') \bar{q}^\dagger_{c\lambda}(w) e^{i(k'^+ - k^+)z^-}\delta_{cc'}, \tag{2.184}$$

$$\sum_{\lambda,\lambda'}\sum_{c,c'} v^{+\,\dagger}_{\lambda'}(w')\gamma^0\gamma^+\gamma_5 u^+_\lambda(w) \bar{q}'_{c'\lambda'}(w') q_{c\lambda}(w) e^{-i(k'^+ - k^+)z^-}\delta_{cc'}, \tag{2.185}$$

$$\sum_{\lambda,\lambda'}\sum_{c,c'} v^{+\,\dagger}_{\lambda'}(w')\gamma^0\gamma^+\gamma_5 v^+_\lambda(w) \bar{q}'_{c'\lambda'}(w') \bar{q}^\dagger_{c\lambda}(w) e^{-i(k'^+ + k^+)z^-}\delta_{cc'}. \tag{2.186}$$

The next step is to show that the only non-vanishing quantity is Eq. (2.185). Calculating this quantity is instructive to understand why the other contributions vanish.

The net effect of the leading-twist operator $\gamma^0\gamma^+\gamma_5$ on the the good spinors defined in Eq. (2.105) is

$$\gamma^0\gamma^+\gamma_5 \begin{cases} u^+_\uparrow(w) \\ u^+_\downarrow(w) \\ v^+_\uparrow(w) \\ v^+_\downarrow(w) \end{cases} = \sqrt[4]{2}\sqrt{k^+} \begin{cases} (1 \quad 0 \quad 1 \quad 0)^T \\ (0 \quad -1 \quad 0 \quad 1)^T \\ (0 \quad 1 \quad 0 \quad -1)^T \\ (-1 \quad 0 \quad -1 \quad 0)^T \end{cases}. \tag{2.187}$$

The scalar products inside Eqs. (2.183) - (2.186) between the adjoint good spinors and the results of Eq. (2.187) simply introduce a dependence on k^+ and k'^+ based on the relative sign of the helicities, according to

$$u_{\lambda'}^{+\dagger}(w')\gamma^0\gamma^+\gamma_5 u_{\lambda}^{+}(w) = \begin{cases} 2\sqrt{k^+k'^+} \text{ if } \lambda = \lambda' = \uparrow \\ -2\sqrt{k^+k'^+} \text{ if } \lambda = \lambda' = \downarrow \\ 0 \text{ otherwise} \end{cases},$$

$$u_{\lambda'}^{+\dagger}(w')\gamma^0\gamma^+\gamma_5 v_{\lambda}^{+}(w) = \begin{cases} 2\sqrt{k^+k'^+} \text{ if } \lambda = \uparrow, \lambda' = \downarrow \\ -2\sqrt{k^+k'^+} \text{ if } \lambda = \downarrow, \lambda' = \uparrow \\ 0 \text{ otherwise} \end{cases},$$

$$v_{\lambda'}^{+\dagger}(w')\gamma^0\gamma^+\gamma_5 u_{\lambda}^{+}(w) = \begin{cases} 2\sqrt{k^+k'^+} \text{ if } \lambda = \downarrow, \lambda' = \uparrow \\ -2\sqrt{k^+k'^+} \text{ if } \lambda = \uparrow, \lambda' = \downarrow \\ 0 \text{ otherwise} \end{cases},$$

$$v_{\lambda'}^{+\dagger}(w')\gamma^0\gamma^+\gamma_5 v_{\lambda}^{+}(w) = \begin{cases} 2\sqrt{k^+k'^+} \text{ if } \lambda = \lambda' = \downarrow \\ -2\sqrt{k^+k'^+} \text{ if } \lambda = \lambda' = \uparrow \\ 0 \text{ otherwise} \end{cases}. \tag{2.188}$$

The presence of the 2-Fock state $|\pi(p)\rangle_{q\bar{q}'}^{l_z=0}$ introduces additional quark and anti-quark ladder operators that can be combined with the ladder operators in (2.185) using the anti-commutation relations in Eqs. (2.76) - (2.77):

$$\begin{aligned}
&\langle 0|\bar{q}'_{c'\lambda'}(w')q_{c\lambda}(w)\left[q^\dagger_{c_1\uparrow}(w_1)\bar{q}'^\dagger_{c_2\downarrow}(w_2) - q^\dagger_{c_1\downarrow}(w_1)\bar{q}'^\dagger_{c_2\uparrow}(w_2)\right]|0\rangle \\
&= \langle 0|\,\bar{q}'_{c'\lambda'}(w')\Big\{\left[q_{c\lambda}(w)q^\dagger_{c_1\uparrow}(w_1) + q^\dagger_{c_1\uparrow}(w_1)q_{c\lambda}(w)\right]\bar{q}'^\dagger_{c_2\downarrow}(w_2) \\
&\qquad - \left[q_{c\lambda}(w)q^\dagger_{c_1\downarrow}(w_1) + q^\dagger_{c_1\downarrow}(w_1)q_{c\lambda}(w)\right]\bar{q}'^\dagger_{c_2\uparrow}(w_2)\Big\}|0\rangle \\
&= \langle 0|\,\bar{q}'_{c'\lambda'}(w')\Big\{(16\pi^3)p_1^+\delta(k^+ - p_1^+)\delta^{(2)}\left(\boldsymbol{k}_\perp - \boldsymbol{p}_{\perp 1}\right) \\
&\qquad \times \left(\delta_{\lambda\uparrow}\bar{q}'^\dagger_{c_2\downarrow}(w_2) - \delta_{\lambda\downarrow}\bar{q}'^\dagger_{c_2\uparrow}(w_2)\right)\delta_{cc_1}\Big\}|0\rangle \\
&= \langle 0|\,(16\pi^3)p_1^+\delta(k^+ - p_1^+)\delta^{(2)}\left(\boldsymbol{k}_\perp - \boldsymbol{p}_{\perp 1}\right)\delta_{cc_1} \\
&\qquad \times \Big\{\delta_{\lambda\uparrow}\left[\bar{q}'_{c'\lambda'}(w')\bar{q}'^\dagger_{c_2\downarrow}(w_2) - \bar{q}'^\dagger_{c_2\downarrow}(w_2)\bar{q}'_{c'\lambda'}(w')\right] \\
&\qquad\quad - \delta_{\lambda\downarrow}\left[\bar{q}'_{c'\lambda'}(w')\bar{q}'^\dagger_{c_2\uparrow}(w_2) - \bar{q}'^\dagger_{c_2\uparrow}(w_2)\bar{q}'_{c'\lambda'}(w')\right]\Big\}|0\rangle \\
&= (16\pi^3)^2 p_1^+p_2^+\delta(k^+ - p_1^+)\delta(k'^+ - p_2^+)\delta^{(2)}\left(\boldsymbol{k}_\perp - \boldsymbol{p}_{\perp 1}\right)\delta^{(2)}\left(\boldsymbol{k}'_\perp - \boldsymbol{p}_{\perp 2}\right) \\
&\qquad \times \left[\delta_{\lambda\uparrow}\delta_{\lambda'\downarrow} - \delta_{\lambda\downarrow}\delta_{\lambda'\uparrow}\right]\delta_{cc_1}\delta_{c'c_2}.
\end{aligned} \tag{2.189}$$

In the very last part of the previous computation, we used the normalization of the 0-particle Fock state, i.e., $\langle 0|0\rangle = 1$. The matching between two

0-particle Fock states is due to the fact that the number of quark and antiquark creator ladder operators in the construction of the pion state with two partons is the same as the number of annihilation quark and antiquark ladder operators in (2.185). On the contrary, the combination of ladder operators in the contributions (2.183), (2.184) and (2.186) does not match that of the pion state in the 2-particles Fock state. This implies, in the end, some scalar products between $|0\rangle$ and $\langle N \neq 0|$, which are 0 due to the orthogonality of the Fock states. By exploiting what we found in (2.188) and the Krönecker deltas over the helicity, we obtain:

$$\begin{aligned}
&\langle 0| \sum_{\lambda,\lambda'} \sum_{c,c'} v^{+\dagger}_{\lambda'}(w')\gamma^0\gamma^+\gamma_5 u^+_{\lambda}(w)\bar{q}'_{c'\lambda'}(w')q_{c\lambda}(w)e^{-i(k'^+-k^+)z^-}\delta_{cc'}\,|\pi(p)\rangle^{l_z=0}_{q\bar{q}'} \\
&\quad = -\int \frac{dk^+ d^2\boldsymbol{k}_\perp}{2k^+(2\pi)^3}\frac{dk'^+ d^2\boldsymbol{k}'_\perp}{2k'^+(2\pi)^3}\sum_{c,c'} 4\sqrt{k^+k'^+}e^{-i(k'^+-k^+)z^-} \\
&\qquad \times \int d[1]d[2]\sum_{c_1c_2}\psi^{(1)}_{q\bar{q}'}(1,2)\frac{\delta_{c_1c_2}}{\sqrt{3}}(16\pi^3)^2 p_1^+ p_2^+ \\
&\qquad\quad \times \delta(k^+ - p_1^+)\delta(k'^+ - p_2^+)\delta^{(2)}\left(\boldsymbol{k}_\perp - \boldsymbol{p}_{\perp 1}\right)\delta^{(2)}\left(\boldsymbol{k}'_\perp - \boldsymbol{p}_{\perp 2}\right)\delta_{cc_1}\delta_{c'c_2}\delta_{cc'}.
\end{aligned} \tag{2.190}$$

After performing the sums over colour indices and the integrals over the momentum variables, we obtain the final result:

$$\begin{aligned}
&\langle 0|\bar{\Psi}^{q'}(z)\gamma^+\gamma_5\Psi^q(-z)\,|\pi(p)\rangle^{l_z=0}_{q\bar{q}'} \\
&\quad = 4\sqrt{3}p^+\int dx_1[d^2\boldsymbol{k}_\perp]_2 e^{ip^+z^-(2x_1-1)}\psi^{(1)}_{q\bar{q}'}(x_1, 1-x_1, \boldsymbol{k}_{\perp 1}, \boldsymbol{k}_{\perp 2}).
\end{aligned} \tag{2.191}$$

The significance of this result lies in the ability to establish a connection between the pion 2-parton LFWA and the twist-2 pion DA via Eq. (2.177):

$$\int \left[d^2\boldsymbol{k}_\perp\right]_2 \psi^{(1)}_{q\bar{q}'}\left(x, 1-x, \boldsymbol{k}_{\perp 1}, \boldsymbol{k}_{\perp 2}\right) = \frac{1}{4\sqrt{3}} f_\pi \phi_\pi(x). \tag{2.192}$$

We stress that this equation is model-independent, as it has been derived from very general principles of QCD.

From this point onwards, one can proceed to explicitly express a parametrization for the $\psi^{(1)}_{q\bar{q}'}$ thus obtaining model-dependent results.

In our case, we apply the model introduced in Section 2.4.2. As previously shown, the underlying logic of our model was to find a direct link between the part of the LFWAs that depends only on the fraction of collinear momentum x_i and the pion DAs. With the introduction of this parametrization, thanks to the property in Eq. (2.157), we find the following simple and direct relation

$$\phi^{(1)}_{q\bar{q}'}\left(x, 1-x\right) = \frac{\pi}{\sqrt{6}} a_2 f_\pi \phi_\pi\left(x\right). \tag{2.193}$$

This relation provides a way to predict the behaviour of ϕ_π once all the free parameters of the model have been determined by a fit on independent

observables.

The discussion presented so far prompts a spontaneous question: *"How can we parametrize the different parton distribution functions, with a model provided for the LFWAs?"*. The answer lies in the LFWA overlap representation of parton distribution functions, which has been extensively explored in previous studies [48–51]. In the next Chapters, we will present the specific LFWA overlap representation formulas for the different parton distribution functions analyzed in this work.

Chapter 3

Pion collinear parton distribution functions

This Chapter focuses on the parametrization and the fit of the pion PDFs. In the initial part we introduce the definition of the PDFs in terms of the collinear quark-quark correlator, their LFWA overlap representation and explicit expression within the model introduced in Section 2.4.2. In the second part, we report the fit procedure with the data selection, and we present the results in comparison with existing extractions in literature.

3.1 Collinear correlators and parton distribution functions

The collinear quark-quark correlator is derived from the general correlator in Eq. (2.123) by integrating over the quark transverse momentum and taking the forward matrix element with $p' = p$. In the case of spin-0 target such as the pion, the only non vanishing contribution is obtained for $\Gamma \equiv \gamma^+$, which corresponds to the following definition of the unpolarized PDF

$$f_1^q(x) = \int \frac{d\zeta^-}{2(2\pi)} e^{ik^+\zeta^-} \langle \pi(p)|\bar{\Psi}^q(0)\gamma^+\Psi^q(\zeta)|\pi(p)\rangle\Big|_{\substack{\zeta^+=0\\ \boldsymbol{\zeta}_\perp=\mathbf{0}}}. \tag{3.1}$$

In our analysis, thanks to the presence of the Fock states with one and two gluons in (2.135) - (2.136), we are able to define also the collinear parton distribution function in which the active parton is a gluon.

The gluon PDF is defined in a similar manner to Eq. (3.1), but involving the matrix element of two gluon field strength tensors (2.62)

$$f_1^g(x) = \frac{1}{xp^+} \int \frac{d\zeta^-}{2\pi} e^{ik^+\zeta^-} \langle \pi(p)|G^{+i}{}_a(0)G_{+i\,a}(\zeta)|\pi(p)\rangle\Big|_{\substack{\zeta^+=0\\ \boldsymbol{\zeta}_\perp=\mathbf{0}}}. \tag{3.2}$$

In Eq. (3.2) and hereafter we use the Einstein convention for repeated colour as well spatial indices.

We stress again that the parametrization that we built for the pion state contains only 4 Fock states, but those are the essential elements in order to obtain a description for all the PDFs. To the best of our knowledge, it represents the most inclusive model concerning the number of Fock components included in the analysis of the pion internal structure.

The equation (3.1) can be used to find the PDF of a quark and antiquark of a certain fixed flavour. In this Chapter, as well as in Chapter 4, we refer to positive pions π^+, using the notation f_1^q.

By neglecting electroweak corrections and the difference in the quark masses, charge symmetry imposes the following relations

$$\begin{aligned} f^u_{1,\pi^+}(x) &= f^{\bar{d}}_{1,\pi^+}(x) \\ &= f^d_{1,\pi^-}(x) = f^{\bar{u}}_{1,\pi^-}(x) \\ &= 2f^u_{1,\pi^0}(x) = 2f^{\bar{u}}_{1,\pi^0}(x) = 2f^d_{1,\pi^0}(x) = 2f^{\bar{d}}_{1,\pi^0}(x). \end{aligned} \tag{3.3}$$

The model scale is set at $\mu_0 = 0.85$ GeV. Such a scale has been fixed phenomenologically, to yield the best-fit value for the χ^2-fit of the PDFs when the scale is varied within a range of values below the charm-mass threshold (see Section 3.4.2), thus ignoring the presence of the sea pairs $c\bar{c}, t\bar{t}$ and $b\bar{b}$ at the initial scale.
We also assume a SU(3)-symmetric sea, i.e,

$$f^d_1(x) = f^{\bar{u}}_1(x) = f^s_1(x) = f^{\bar{s}}_1(x). \tag{3.4}$$

We can now identify other four significant PDFs that are often used in the literature: the *u-valence* $f_1^{u_v}(x)$, the *d-valence* $f_1^{d_v}(x)$, the *total valence* $f_1^v(x)$ and the *total Sea* $f_1^S(x)$. Their definitions are:

$$f^{u_v}_1(x) = f^u_1(x) - f^{\bar{u}}_1(x), \tag{3.5}$$
$$f^{d_v}_1(x) = f^d_1(x) - f^{\bar{d}}_1(x), \tag{3.6}$$
$$f^v_1(x) = f^{u_v}_1(x) - f^{d_v}_1(x), \tag{3.7}$$
$$f^S_1(x) = 2f^d_1(x) + 2f^{\bar{u}}_1(x) + f^s_1(x) + f^{\bar{s}}_1(x). \tag{3.8}$$

Thanks to the symmetries (3.3) and (3.4), the total valence and the total Sea can be rewritten as:

$$f^v_1(x) = 2f^{u_v}_1(x), \tag{3.9}$$
$$f^S_1(x) = 6f^d_1(x). \tag{3.10}$$

As it is obvious from the previous equations, the four distributions (3.5) - (3.8) are not all independent. In literature different sets of independent PDFs are studied. In the upcoming Sections, we will discuss the parametrization for $f_1^v(x)$, $f_1^g(x)$, $f_1^S(x)$. We will also present the model's outcomes for these specific PDFs or, alternatively, for $f_1^{u_v}(x) = f_1^v(x)/2$.

3.2 LFWA overlap representation of the PDFs

The LFWA overlap representation of the PDFs is obtained by substituting the pion states with their expansions in terms of Fock states (2.134) - (2.137) into the correlators defined in Eqs. (3.1) - (3.2). The resulting functions depend on a superposition of LFWAs, that we parametrized in Section 2.4.2. Since there is no momentum transfer between the initial and the final pion state ($\Delta = 0$), the LFWA overlap representation of the PDFs will exclusively consist of the squares of the various LFWAs.

For the total valence PDF, all the Fock states contribute in the matrix elements (3.1), and the expression of $f_1^v(x)$ can be decomposed in terms of the individual contributions as follows

$$f_1^v(x) = f^v_{1,u\bar{d}}(x) + f^v_{1,u\bar{d}g}(x) + f^v_{1,u\bar{d}gg}(x) + \sum_{\{s\bar{s}\}} f^v_{1,u\bar{d}\{s\bar{s}\}}(x). \tag{3.11}$$

In the case of the total Sea, only the Fock states of Eq. (2.137) contribute, while the gluon PDF is obtained from the Fock-states in Eqs. (2.135) and (2.136). We therefore can write

$$f_1^S(x) = 2\sum_{\{s\bar{s}\}} f^S_{1,u\bar{d}\{s\bar{s}\}}(x), \tag{3.12}$$

$$f_1^g(x) = f^g_{1,u\bar{d}g}(x) + f^g_{1,u\bar{d}gg}(x). \tag{3.13}$$

By using the expressions of each Fock state in terms of LFWAs, we obtain the following representation for each contribution:

$$f^v_{1,u\bar{d}}(x) = 4\int \mathrm{d}[1]\mathrm{d}[2]\sqrt{x_1x_2}\delta(x-x_1)|\psi^{(1)}_{u\bar{d}}(1,2)|^2, \tag{3.14}$$

$$f^v_{1,u\bar{d}g}(x) = 4\int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3]\sqrt{x_1x_2x_3}\delta(x-x_1)|\psi^{(1)}_{u\bar{d}g}(1,2,3)|^2, \tag{3.15}$$

$$\begin{aligned} f^v_{1,u\bar{d}gg}(x) = 16\int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3]\mathrm{d}[4]\sqrt{x_1x_2x_3x_4}\delta(x-x_1) \\ \times\left[|\psi^{(1)}_{u\bar{d}gg}(1,2,3,4)|^2 + |\psi^{(2)}_{u\bar{d}gg}(1,2,3,4)|^2\right], \end{aligned} \tag{3.16}$$

$$\begin{aligned} f^v_{1,u\bar{d}\{s\bar{s}\}}(x) = 8\int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3]\mathrm{d}[4]\sqrt{x_1x_2x_3x_4}\delta(x-x_1) \\ \times\left[|\psi^{(1)}_{u\bar{d}s\bar{s}}(1,2,3,4)|^2 + |\psi^{(2)}_{u\bar{d}s\bar{s}}(1,2,3,4)|^2 + \frac{1}{2}|\psi^{(3)}_{u\bar{d}s\bar{s}}(1,2,3,4)|^2\right], \end{aligned} \tag{3.17}$$

$$\begin{aligned} f^S_{1,u\bar{d}\{s\bar{s}\}}(x) = 4\int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3]\mathrm{d}[4]\sqrt{x_1x_2x_3x_4}\delta(x-x_3) \\ \times\left[|\psi^{(1)}_{u\bar{d}s\bar{s}}(1,2,3,4)|^2 + |\psi^{(2)}_{u\bar{d}s\bar{s}}(1,2,3,4)|^2 + \frac{1}{2}|\psi^{(3)}_{u\bar{d}s\bar{s}}(1,2,3,4)|^2\right], \end{aligned} \tag{3.18}$$

$$f^g_{1,u\bar{d}g}(x) = 2\int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3]\sqrt{x_1x_2x_3}\delta(x-x_3)|\psi^{(1)}_{u\bar{d}g}(1,2,3)|^2, \tag{3.19}$$

$$f^g_{1,u\bar{d}gg}(x) = 16\int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3]\mathrm{d}[4]\sqrt{x_1x_2x_3x_4}\delta(x-x_3) \times\left[|\psi^{(1)}_{u\bar{d}gg}(1,2,3,4)|^2 + |\psi^{(2)}_{u\bar{d}gg}(1,2,3,4)|^2\right]. \tag{3.20}$$

3.3 Explicit model for pion PDFs

We are now in the position to implement the model developed in Section 2.4.2 for the LFWAs to find the parametrization of the PDFs in Eqs. (3.14) - (3.20). The results for the three independent PDFs are as follows

$$\begin{aligned} f^v_1(x,\mu_0^2) &= \mathcal{C}^v_{u\bar{d}}(x\bar{x})^{2\gamma_q-1}\left[1+d_{q1}(1+2\gamma_q)\left(1+\gamma_q(x-\bar{x})^2-6x\bar{x}\right)\right]^2 \\ &+ \mathcal{C}^v_{u\bar{d}g}x\bar{x}^5\left[3+18xd_{g1}-10\bar{x}d_{g1}+13d^2_{g1}+14xd^2_{g1}(x-4\bar{x})\right] \\ &+ \mathcal{C}^v_{u\bar{d}gg}x\bar{x}^9 + \sum_{\{s\bar{s}\}}\mathcal{C}^v_{u\bar{d}\{s\bar{s}\}}x\bar{x}^5, \end{aligned} \tag{3.21}$$

$$\begin{aligned} f^g_1(x,\mu_0^2) &= \mathcal{C}^g_{u\bar{d}g}(x\bar{x})^3\left[1+d_{g1}(3-7x)\right]^2 + \mathcal{C}^g_{u\bar{d}gg_1}x^3\bar{x}^7 \\ &+ \mathcal{C}^g_{u\bar{d}gg_2}x\bar{x}^5(5-20x-6x^2+52x^3+95x^4), \end{aligned} \tag{3.22}$$

$$f^S_1(x,\mu_0^2) = 2\sum_{\{s\bar{s}\}}\mathcal{C}^S_{u\bar{d}\{s\bar{s}\}}x\bar{x}^5, \tag{3.23}$$

where $\bar{x}=1-x$ and the $\mathcal{C}$-coefficients are related to the norms of Eqs. (2.172) - (2.175) through the following expressions

$$\begin{aligned} &\mathcal{C}^v_{u\bar{d}} = 4N^{(1)\,2}_{q\bar{q}'},\ \mathcal{C}^v_{u\bar{d}g} = \frac{1}{15}N^{(1)\,2}_{q\bar{q}'g},\quad \mathcal{C}^v_{u\bar{d}gg} = \frac{1}{315}N^{(1)\,2}_{q\bar{q}'gg},\quad \mathcal{C}^v_{u\bar{d}\{s\bar{s}\}} = \frac{1}{5}N^{(1)\,2}_{q\bar{q}'s\bar{s}}, \\ &\mathcal{C}^g_{u\bar{d}g} = \frac{1}{3}N^{(1)\,2}_{q\bar{q}'g},\ \mathcal{C}^g_{u\bar{d}gg_1} = \frac{2}{105}N^{(1)\,2}_{q\bar{q}'gg},\ \mathcal{C}^g_{u\bar{d}gg_2} = \frac{1}{4410}N^{(1)\,2}_{q\bar{q}'gg}, \\ &\mathcal{C}^S_{u\bar{d}\{s\bar{s}\}} = \frac{1}{10}N^{(1)\,2}_{q\bar{q}'s\bar{s}}. \end{aligned} \tag{3.24}$$

The PDFs depend on the energy scale μ^2 of the physical process. The equations (3.21) - (3.23) are intended to be true at the model scale, as indicated by the explicit dependence on μ_0^2. The discussion about the scale evolution of the PDFs is postponed to Section (3.5.3).

The theoretical model, expressed in Eqs. (3.21) - (3.23) for the PDFs, automatically satisfies two fundamental properties: the *valence number* and the *momentum sum rule*:

$$\int_0^1 f^v_1(x)dx = 2, \tag{3.25}$$

$$\int_0^1 x\left(f^v_1(x)+f^g_1(x)+f^S_1(x)\right)dx = 1. \tag{3.26}$$

This is particularly useful when we implement the fitting code, since we do not need to impose these constraints as external auxiliary boundary conditions.

3.4 Fit procedure

The code used for the fit of the PDFs is the open-source tool xFitter [52]. Some of the technical details about the code can be found on the repository online and in Ref. [53]. In this Section we report only the basic concepts necessary to reproduce our results.

3.4.1 Data selection

The available experimental cross sections sensitive to the pion PDFs come from Drell-Yan (DY) experiments and from prompt photon production.

In our analysis, we considered the DY data from the NA10 [54] and E615 [55] experiments and the prompt photon production data of WA70 [56]. The DY data concern the scattering of negative pions off a tungsten target, with a pion beam energy of $E_\pi = 194$ and 286 GeV in the NA10 experiment, and $E_\pi = 252$ GeV in the E615 experiment. To distinguish between the two subsets relative to the different beam energies in the experiment NA10 we use the labels NA10-194 and NA10-286. In the WA70 experiment the data are relative to both positive and negative pion beams on a proton target, with $E_\pi = 280$ GeV. We refer to the subsets with positive or negative pions for WA70 with WA70$^{(+)}$ and WA70$^{(-)}$, respectively.

We applied some kinematic cuts to the original set of experimental data. First of all, we restrict the data to the mass region $(4.16, 7.68)$ GeV. By doing so, we exclude from the analysis those regions corresponding to the J/ψ [57,58] and Υ [59] resonances. The DY measurements are for the differential cross section $\dfrac{d^2\sigma}{d\sqrt{\tau}dx_F}$, expressed in nb/nucleus, and integrated over each $(\sqrt{\tau}, x_F)$ bin, where x_F and τ are defined, respectively, as:

$$x_F = x_0^\pi - x_0^W, \qquad \tau = \frac{m_{\mu\mu}}{\sqrt{s}}. \tag{3.27}$$

In Eq. (3.27), $x_0^{\pi(W)}$ is the minimum momentum fraction of the active parton in the pion (tungsten) to produce the lepton pair in the final state, $m_{\mu\mu}$ is the invariant mass of the muon pair and $\sqrt{s}$ is the center-of-mass energy of pion-nucleon system. A second cut is applied to the Feynman-x to avoid the lower edges of phase space, i.e., $x_F \geq 0$.

The number of data points that pass through the cuts is 70 for the NA10 set and 91 for the E615 set. By including also the prompt-photon data, our database consists of a total number of 260 points.

The prompt-photon cross sections are expressed in pb as $E\dfrac{d^3\sigma}{d^3q}$, and they are averaged over each $(x_F, \mathbf{q}_T)$ bin, where $\mathbf{q}_T$ is the transverse momentum of the emitted photon.

3.4.2 χ^2 definition

The usual definition of the minimization function is:

$$\chi^2 = \sum_{i=1}^{N_{\text{data}}} (E_i - t_i)\, V_{ij}^{-1} (E_j - t_j), \tag{3.28}$$

where V_{ij} is the covariance matrix, E_i is the experimental observable, t_i is the theoretical prediction and N_{data} is the number of data points.
There are, in general, different kinds of errors arising from the experimental measurements. The errors can be statistic or systematic, correlated or uncorrelated. The systematic uncertainties are responsible for the *systematic shifts*, that move the theoretical prediction by a quantity that takes into account the correlated error sources:

$$\tilde{t}_i = t_i \left(1 - \sum_{\alpha=1}^{N_{\text{corr}}} \lambda_{i\alpha}^{\text{corr}} b_\alpha \right), \tag{3.29}$$

where N_{corr} is the number of correlated sources. In Eq. (3.29), the b_α are the so called "nuisance parameters", that add to the number of fitting parameters, while their relative coefficients $\lambda_{i\alpha}^{\text{corr}}$ take into account the influence of the correlated error source α on the data point i and the nuisance parameter b_α.

By considering the systematic shifts, Eq. (3.28) can be rewritten as [53,60]:

$$\chi^2 = \sum_{i=1}^{N_{\text{data}}} \frac{\left(E_i - \tilde{t}_i\right)^2}{\left(\delta_i^{\text{syst}}\right)^2 + \left(\sqrt{\frac{\tilde{t}_i}{E_i}}\delta_i^{\text{stat}}\right)^2} + \sum_{\alpha=1}^{N_{\text{corr}}} b_\alpha^2, \tag{3.30}$$

where δ_i^{syst} and δ_i^{stat} are the systematic and statistic uncertainties, respectively. The optimal values of the nuisance parameters can be obtained analytically via:

$$\frac{\partial \chi^2}{\partial b_\alpha} = 0. \tag{3.31}$$

The part of the chi-squared generated from the nuisance parameters, i.e., $\chi^2_{\text{pen}} = \sum_{\alpha=1}^{N_{\text{corr}}} b_\alpha^2$, is refereed to as *penalty term*. The latter is an index of the quality of the fit, since it is expected to be small: if the nuisance parameters dominate the χ^2, this is an indication of a poor fit.

3.5 Fit results

The model parameters that enter in the PDF fit are only those contained in the set $\mathcal{A}^L$, as explained in Section 2.4.2.

The procedure of minimization of the χ^2-function takes into account the evolution of the PDFs at next-to-leading-order (NLO) from the initial model

scale μ_0 to the energy scale of the experimental data. Once the PDFs are evolved, the computation of the chi-squared is performed. Then, the parameters at the initial scale are modified and a new calculation of the function (3.30) is performed. This sequence is iterated until the minimum of the χ^2 is reached. The optimization routine employed for this purpose is MINUIT [61].

The initial scale μ_0 was fixed as the value that resulted in the lowest χ^2 after the minimization procedure.

The best fit of experimental cross sections sensitive to the PDFs yielded a reduced chi-squared of $\hat{\chi}^2/N_{d.o.f.} = 0.88$, where the number of degrees of freedom is $N_{d.o.f.} = 260 - 6 = 254$. Out of the total chi-squared, the amount due to the penalty term is $\chi^2_{\text{pen.}} = 3$, representing only the 1.3%. This is a signal of the goodness of the fit.

Moreover, the result remains remarkably consistent even with fluctuations in the initial scale: by varying μ_0 in the interval $\mu_0 \in [0.65, 1.05]$ the best-χ^2 changes of less then 1%.

In addition to μ_0, other quantities kept fixed in the fit are the factorization scale μ_F and the renormalization scale μ_R. Their values have been set to $\mu_F = \mu_R = 0.8$ GeV.

We report in the following the set of the best-fit longitudinal parameters:

$$\begin{aligned} &\gamma_q = 0.622, \qquad d_{q1} = -0.145, \qquad d_{g1} = 0.664 \\ &\alpha_1 = 0.842, \qquad \alpha_2 = 1.259, \qquad \alpha_3 = 0.329. \end{aligned} \tag{3.32}$$

Once obtained the best fit parameters, we performed an error analysis with the bootstrap method. The fit was repeated 1000 times by varying each time the experimental data with random gaussian shifts. The widths of the shifts take into account both the statistic and the systematic experimental uncertainties. In each single replica, the initial model scale is always μ_0; on the contrary, we modified the values of μ_F and μ_R replica by replica to keep under control the effects of the factorization and renormalization scale variations. In particular, the value of μ_F has been randomly generated from a uniform distribution in the range $[\mu_0/2, \mu_0]$, while μ_R was consequently varied in the range $[\mu_F, 2\mu_0]$, thus exploring the regions:

$$\begin{aligned} \frac{\mu_0}{2} \leq &\mu_F \leq \mu_0, \\ \frac{\mu_0}{2} \leq &\mu_R \leq 2\mu_0. \end{aligned} \tag{3.33}$$

We report here the average of the parameters from the 1000 replicas along with their respective standard deviations:

$$\begin{aligned} &\langle\gamma_q\rangle = 0.639, \quad \langle d_{q1}\rangle = -0.142, \quad \langle d_{g1}\rangle = 111.386, \\ &\langle\alpha_1\rangle = 0.816, \quad \langle\alpha_2\rangle = 1.364, \quad \langle\alpha_3\rangle = 0.554, \end{aligned} \tag{3.34}$$

$$\sigma_{\gamma_q} = 0.100, \qquad \sigma_{d_{q1}} = 0.047, \quad \sigma_{d_{g1}} = 221,$$

$$\sigma_{\alpha_1} = 0.018, \qquad \sigma_{\alpha_2} = 0.202, \qquad \sigma_{\alpha_3} = 0.082. \tag{3.35}$$

We observe that the parameter d_{g1} is not constrained at all. This is an indication that the data set is not sensitive enough to the gluon PDF. This emphasizes the necessity for new measurements, which could provide additional data points that have the potential to sensitively probe pion PDFs. Actually, there are various planned experiments that hold promising opportunities for further analysis of pion PDFs. Among these, at Jefferson Lab [62] and at the new-generation facility which is the Electron-Ion Collider (EIC) [63, 64] there are plans to study the pion PDFs by exploiting the Sullivan process [65], consisting of scattering off the pion in proton to pion fluctuations [66, 67]. Furthermore, the experiments from COMPASS++/AMBER will use high-energy and high-intensity pion beams to probe directly the partonic structure of the pion [68].

3.5.1 Observables

In addition to the total χ^2 calculated over the whole data set, we can consider the separate χ^2_s for the following five subsets:

	E615	NA10-194	NA10-286	WA70$^{(+)}$	WA70$^{(-)}$
χ^2_s	122.78	40.98	18.34	13.61	25.76
N^s_{data}	91	44	26	31	68
$\chi^2_s/N^s_{\text{data}}$	1.35	0.93	0.71	0.44	0.38

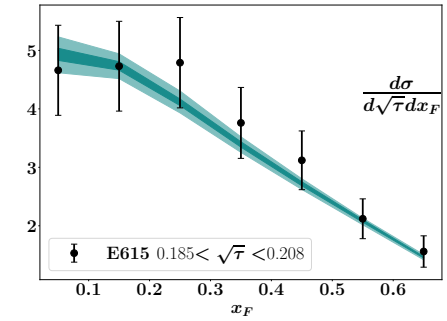
Table 3.1: Chi-squared contributions from the various subsets.

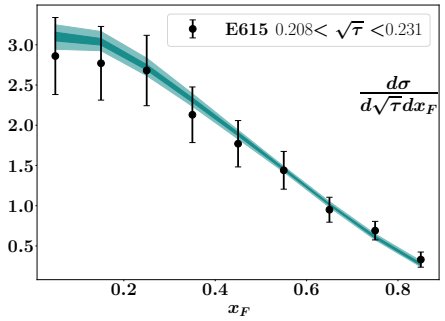
The last row of Tab. 3.1 shows the contribution to the total chi-squared from a specific subset, divided by the number N^s_{data} of data points in the considered subset. This represents an indicator of the average weighting of each single data point in the subset to the total chi-squared. The greatest contribution to the chi-squared comes from the set E615, containing the largest number of points that, on average, exhibit a poorer fit. This could potentially be attributed to the presence of certain data points in the subset E615 with very small errors (see Fig. 3.1), thus introducing relevant contributions to Eq. (3.30).

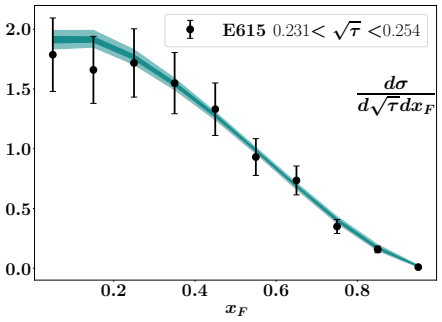
Apart from the quantitative result expressed by the χ^2, the goodness of the fit is qualitatively evident by observing the good agreement between the experimental data and our curves. The results are presented both divided in the different experimental sets and all together in a single figure (see Fig. 3.6).

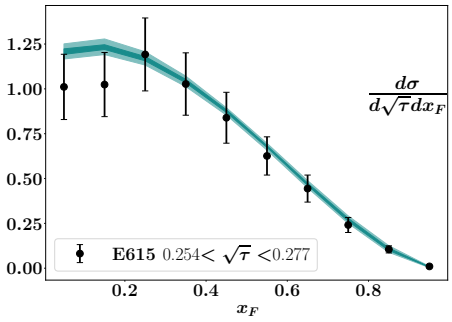
In each of the following plots, the darker internal band corresponds to the 68% of our replicas; the external lighter band refers to the fit of the 100% of the replicas.

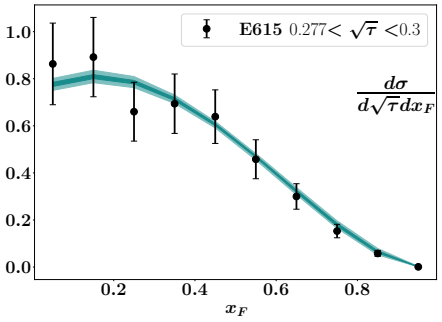
E615

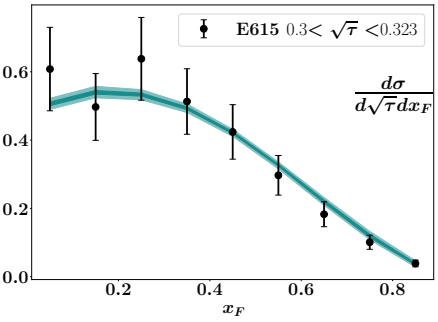

E615 0.185< $\sqrt{\tau}$ <0.208
$\frac{d\sigma}{d\sqrt{\tau}dx_F}$
x_F

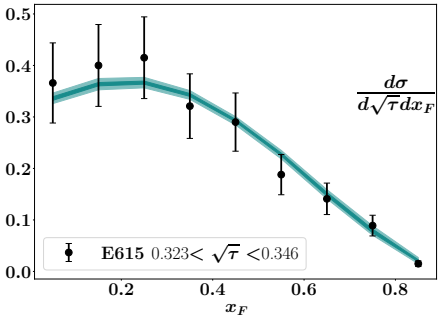

E615 0.208< $\sqrt{\tau}$ <0.231
$\frac{d\sigma}{d\sqrt{\tau}dx_F}$
x_F

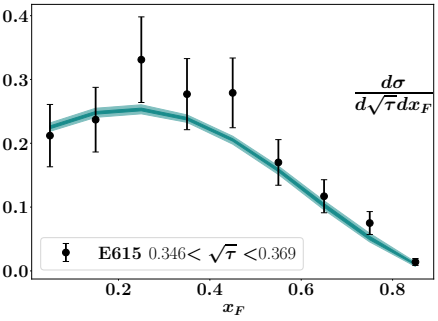

E615 0.231< $\sqrt{\tau}$ <0.254
$\frac{d\sigma}{d\sqrt{\tau}dx_F}$
x_F


E615 0.254< $\sqrt{\tau}$ <0.277
$\frac{d\sigma}{d\sqrt{\tau}dx_F}$
x_F


E615 0.277< $\sqrt{\tau}$ <0.3
$\frac{d\sigma}{d\sqrt{\tau}dx_F}$
x_F


E615 0.3< $\sqrt{\tau}$ <0.323
$\frac{d\sigma}{d\sqrt{\tau}dx_F}$
x_F


E615 0.323< $\sqrt{\tau}$ <0.346
$\frac{d\sigma}{d\sqrt{\tau}dx_F}$
x_F


E615 0.346< $\sqrt{\tau}$ <0.369
$\frac{d\sigma}{d\sqrt{\tau}dx_F}$
x_F

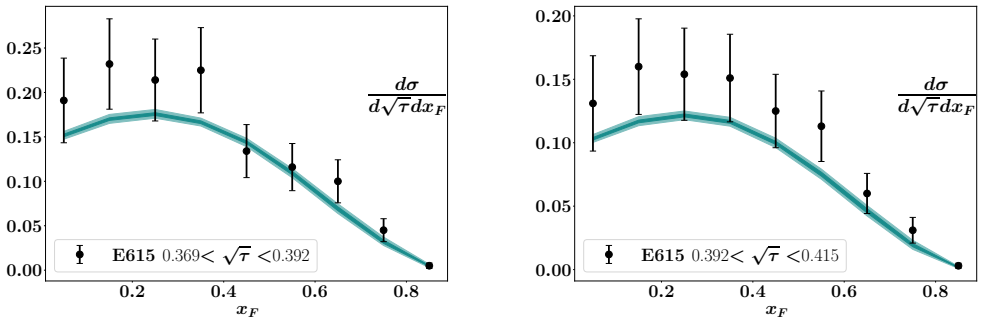
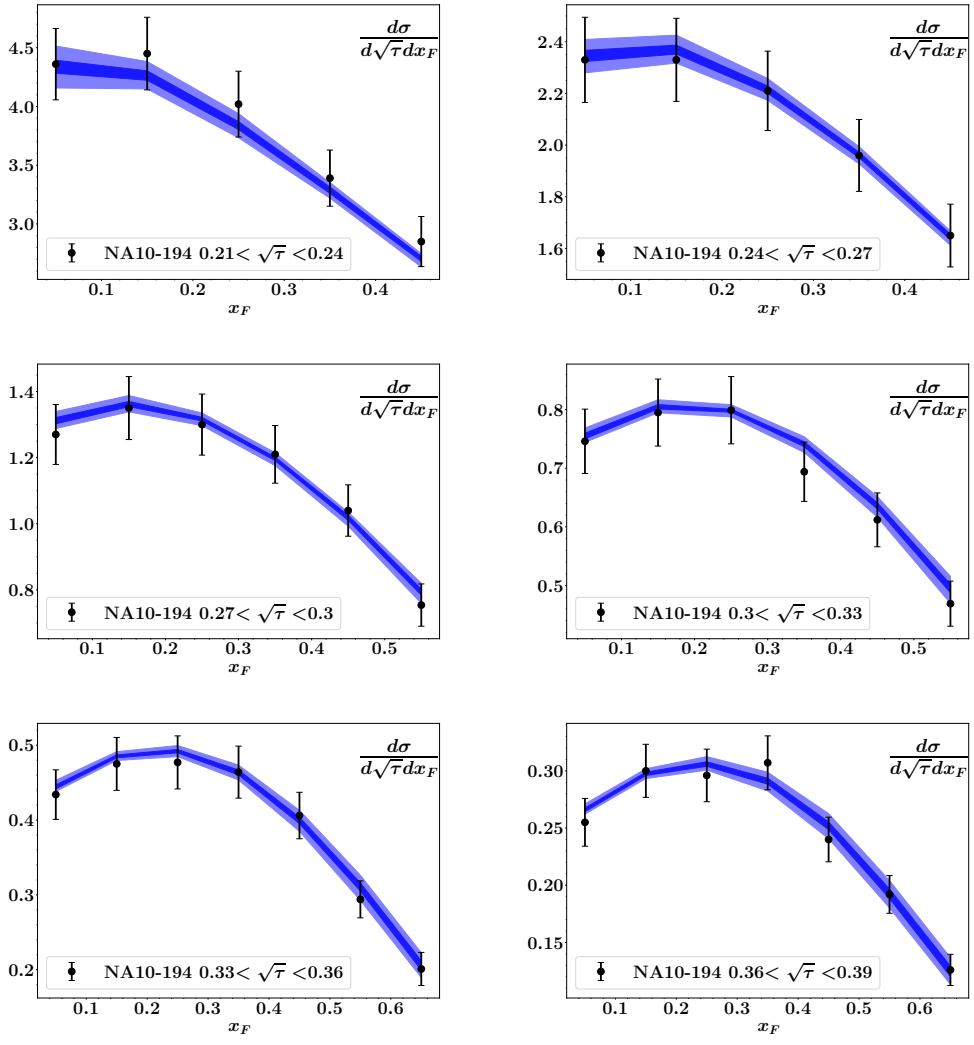


Figure 3.1: Comparison between our extractions (green bands) and the experimental cross sections (black points with error bars) for the DY data of the experiment E615 for different bins in $\sqrt{\tau} \in [0.185, 0.415]$.

NA10-194



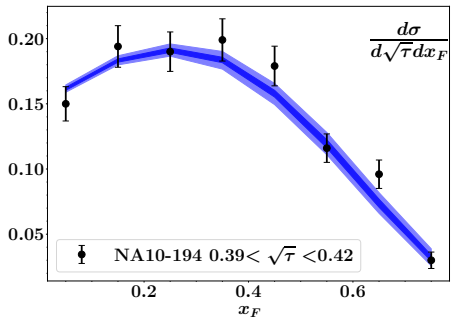
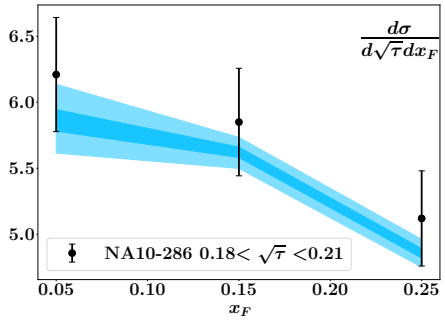
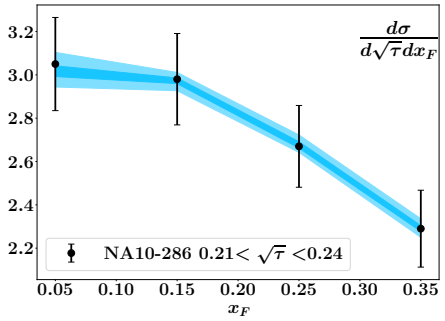
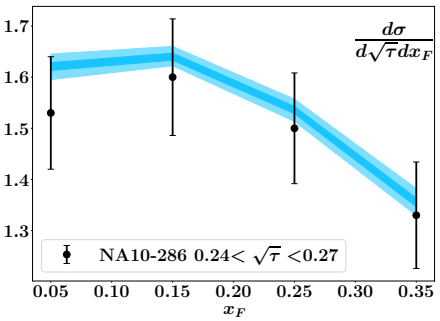


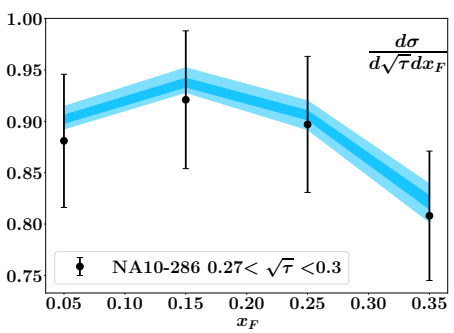
Figure 3.2: Comparison between our extractions (blue bands) and the experimental cross sections (black points with error bars) for the DY data of the experiment NA10 and a beam energy of 194 GeV for different bins in $\sqrt{\tau} \in [0.21, 0.42]$.

NA10-286









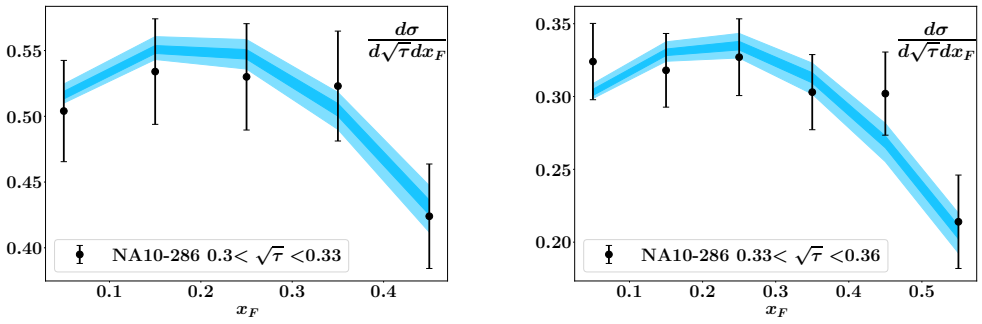
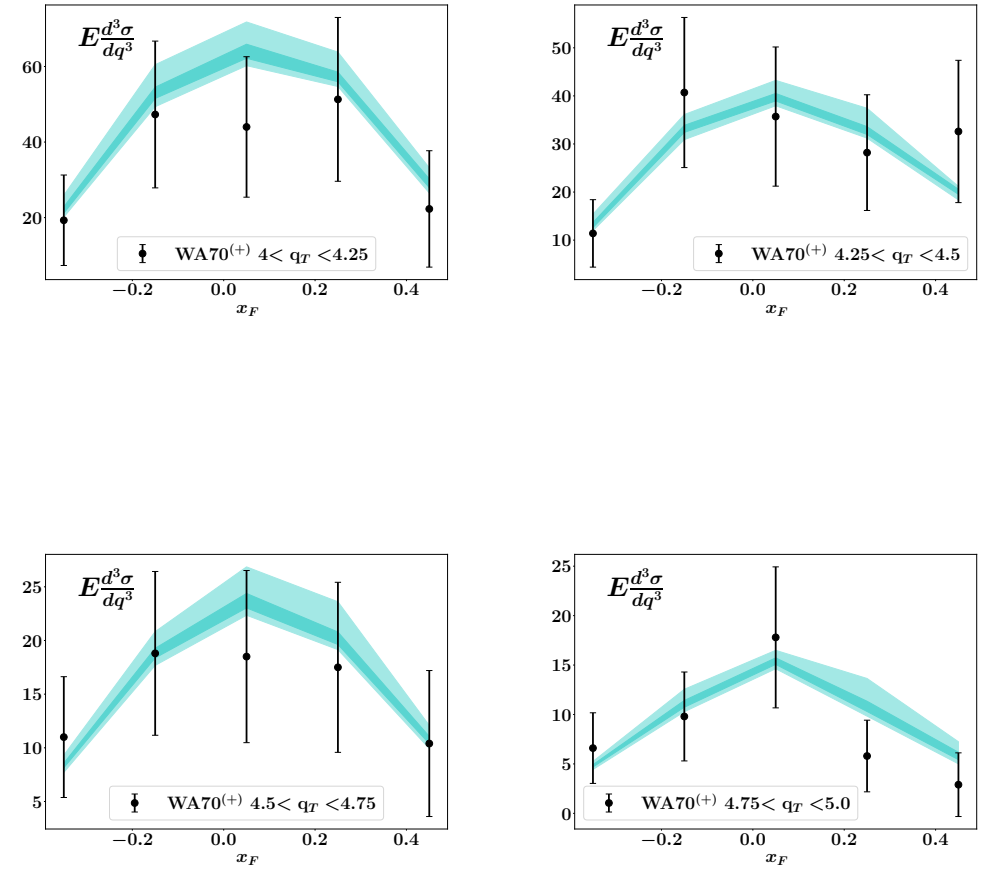


Figure 3.3: Comparison between our extractions (light blue bands) and the experimental cross sections (black points with error bars) for the DY data of the experiment NA10 and a beam energy of 286 GeV for different bins in $\sqrt{\tau} \in [0.18, 0.36]$.

WA70$^{(+)}$



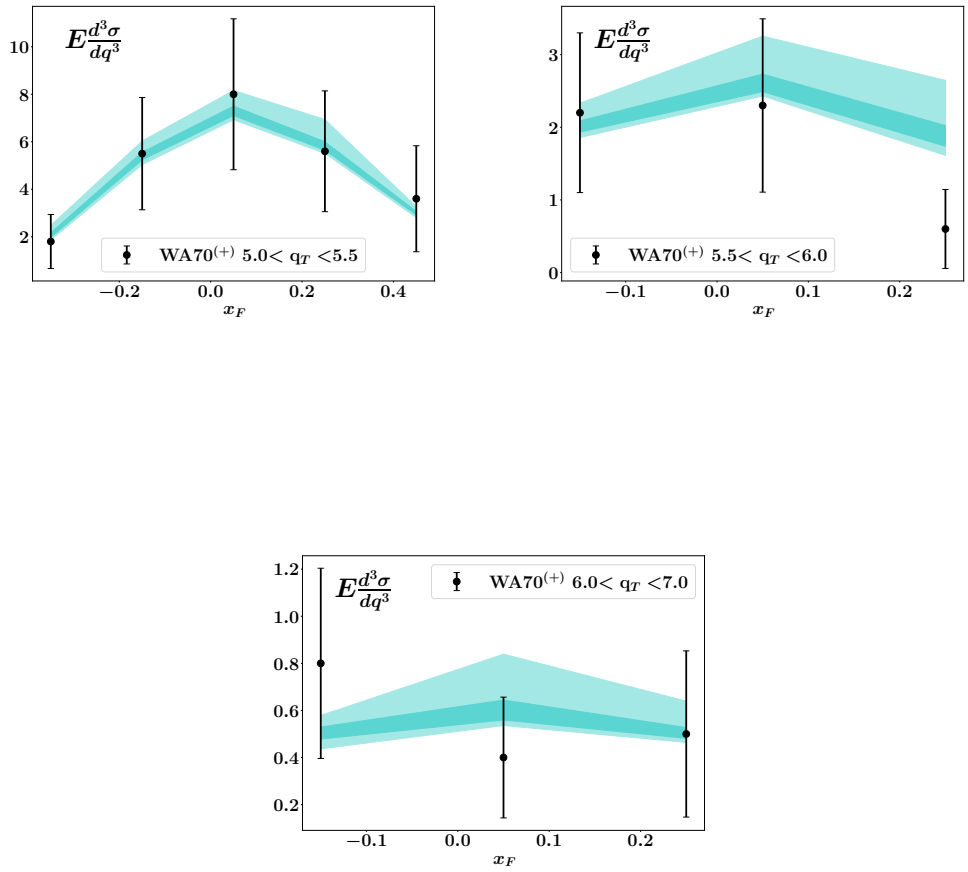
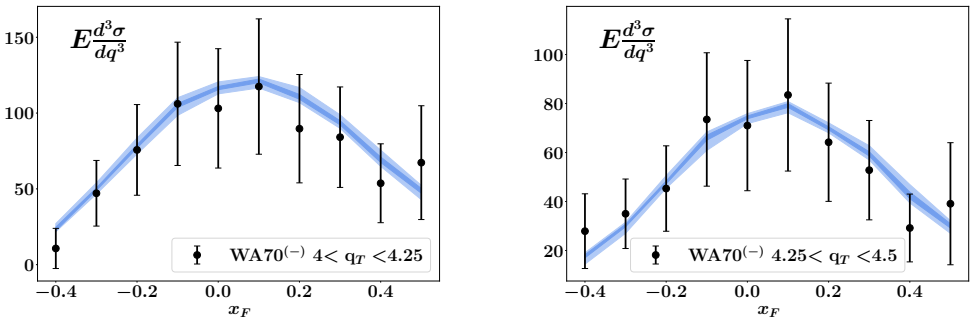


Figure 3.4: Comparison between our extractions (sea green bands) and the experimental cross sections (black points with error bars) for the prompt photon production data of the experiment WA70 with positive pions for different bins in $\mathbf{q}_T \in [4.00, 7.00]$ GeV.

WA70$^{(-)}$



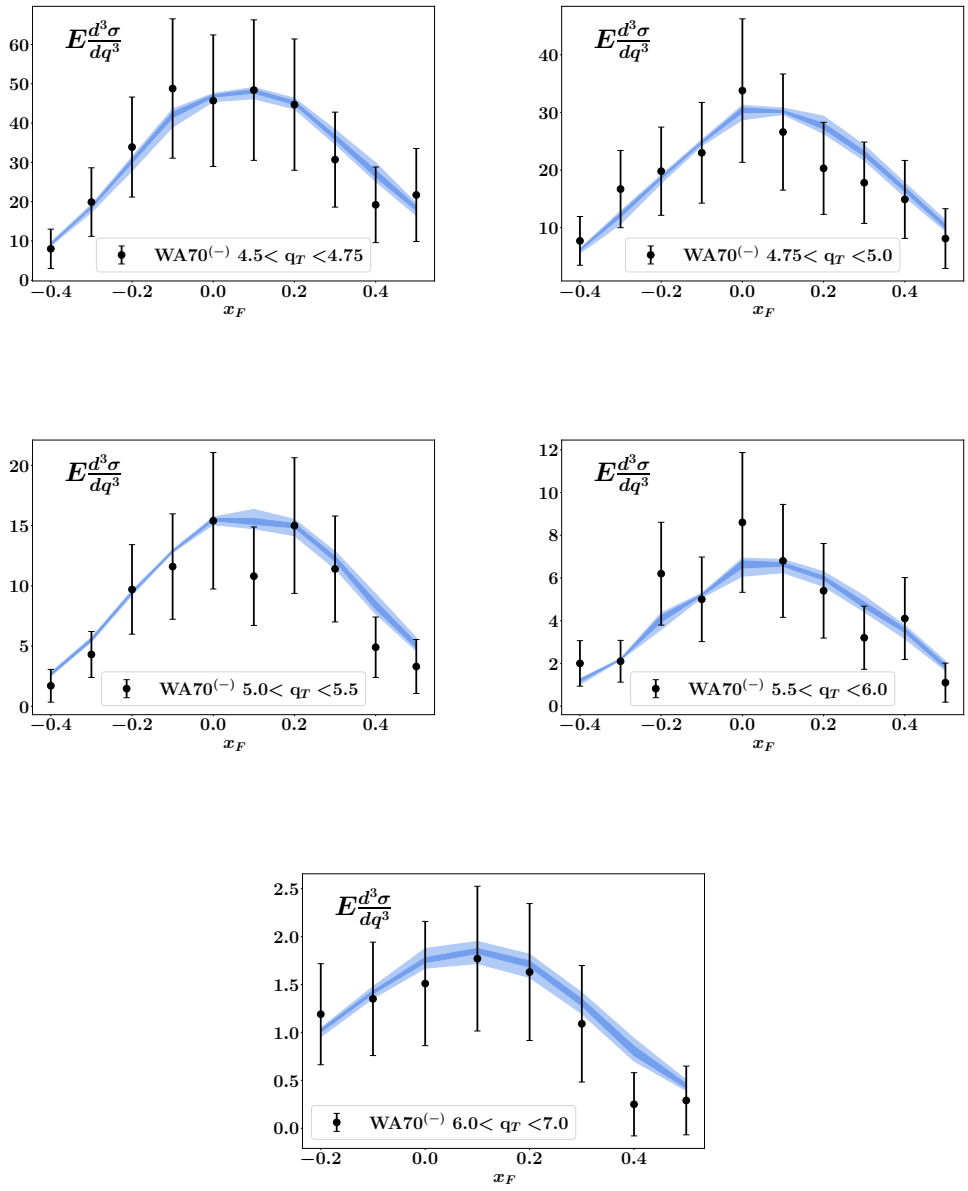


Figure 3.5: Comparison between our extractions (sky blue bands) and the experimental cross sections (black points with error bars) for the prompt photon production data of the experiment WA70 with negative pions for different bins in $\mathbf{q}_T \in [4.00, 7.00]$ GeV.

Summary of all the observables fitted

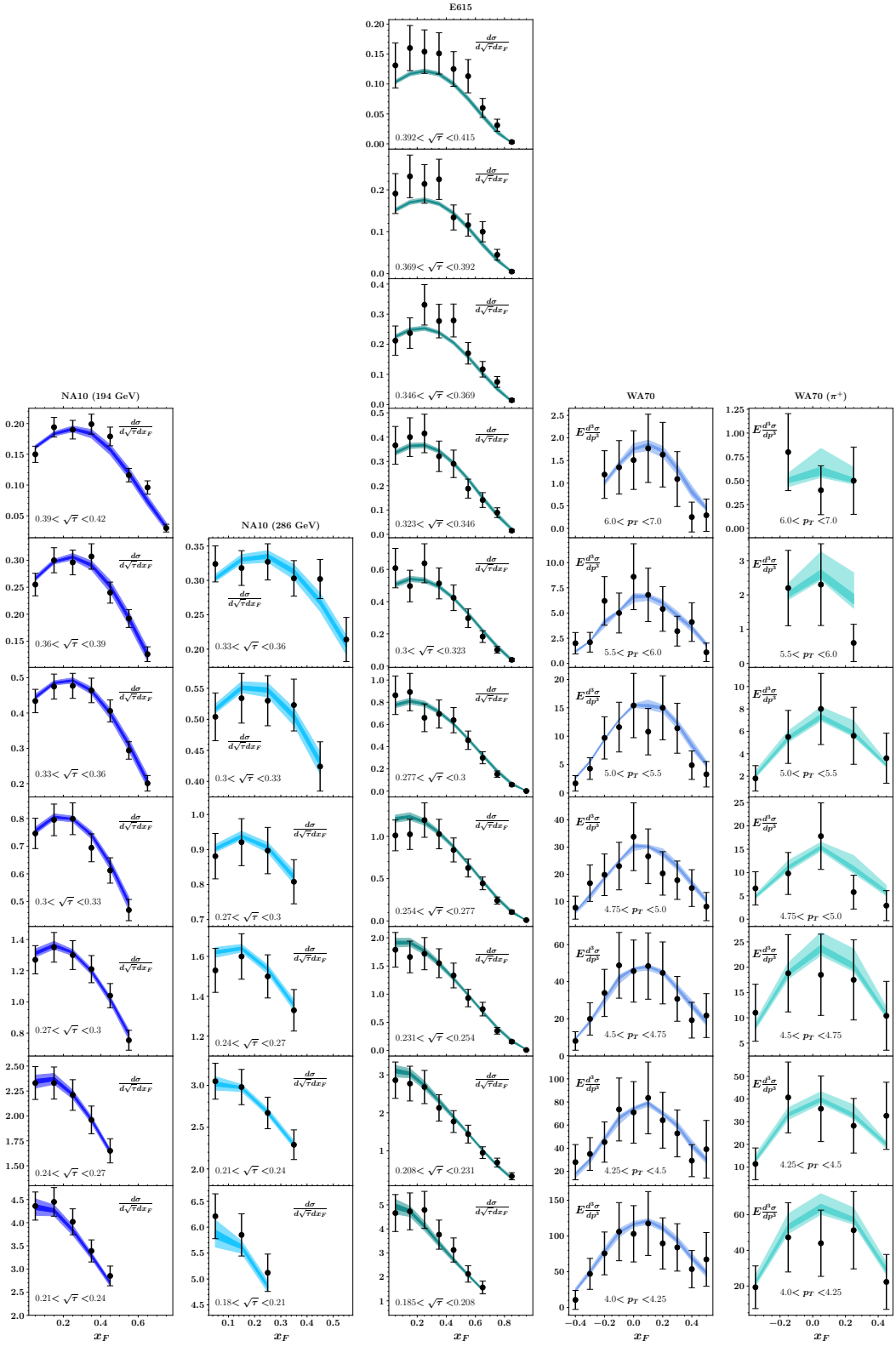


Figure 3.6: Comparison between the whole database of 260 points with the fit results of our model.

3.5.2 PDF results

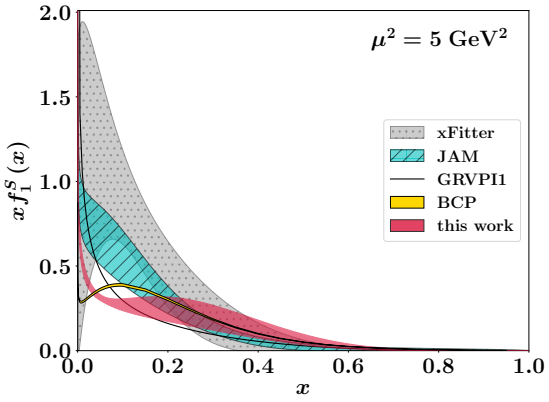
In the literature, the extraction of the PDF are typically presented by showing either the results of all the replicas or the uncertainty band at a certain confidence level. We opt for the second choice, and the construction of the 1-σ PDF band is carried out as follows:

1. a certain number (~ 1000) of values of x are chosen in the interval $[0, 1]$;
2. the set of all the 1000 PDF replicas are evaluated at each x of step 1.;
3. for each x, the mean $r(x)$ and the standard deviation $s(x)$ are computed;
4. the band $r(x) \pm s(x)$ is represented point by point in x.

In the following plots we present the results for our extraction of the PDFs. These results include both the 1-σ central band in dark red and the 3-σ external band in light red, which correspond to the 68% and 99.7% of the confidence level (CL), respectively. The results for the PDFs evolved at NLO are presented at the final scales of $\mu^2 = 5$ GeV^2 and $\mu^2 = 27$ GeV^2 in Fig. 3.7 and 3.8, respectively.

PDFs at 5 GeV^2

In Fig. 3.7 we compare our results at $\mu^2 = 5$ GeV^2 with the extractions of pion PDFs from other studies. The solid black curves represent the best-fit results of the GRVPI1 analysis [69] and the grey bands refer to the results of the xFitter collaboration [53]. All these analyses are based on the same experimental measurements, but with some variations in the database due to different kinematic cuts. The analysis from the JAM collaboration [70], depicted by light-blue bands, integrates both DY data and leading-neutron tagged electroproduction data. It also accounts for threshold resummation on DY cross-sections at next-to-leading log accuracy. A recent analysis, shown by the yellow bands and labelled with BCP, has been conducted using the statistical model [71]. This work extended the database considered in a previous work [72] to include J/ψ production data.



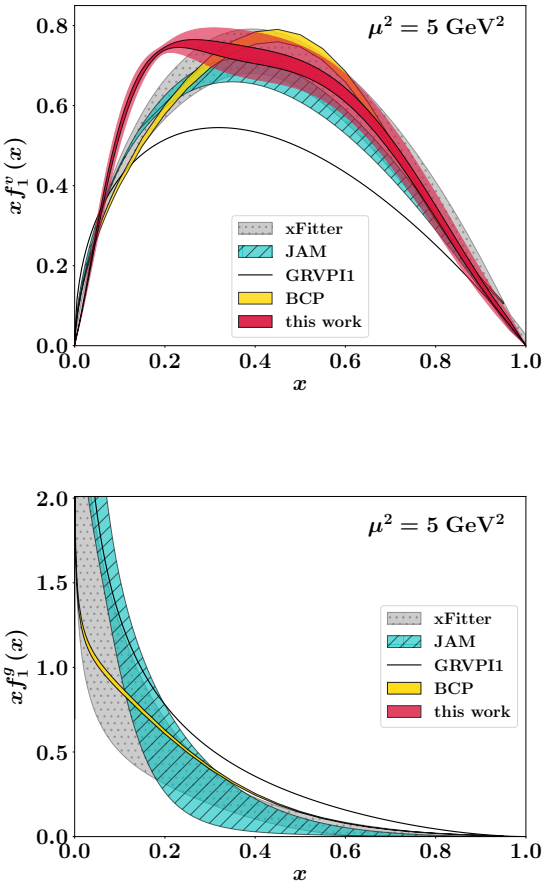


Figure 3.7: xf_1 as function of x for the total valence (upper panel), total sea (central panel) and gluon (lower panel) contributions at $\mu^2 = 5$ GeV^2. Our results (dark and light red bands corresponding to the 1σ and 3σ CL) are compared with the results from the JAM collaboration [70] (light blue bands), the analysis of xFitter collaboration [53] (grey bands), the BCP fit of Ref. [71] (yellow bands) and the GRVPI1 fit [69] (solid black curves).

Upon comparison, we notice that all the analyses yield compatible results, within the error bands. The best agreement is in the high-x region for the total valence and the total sea distributions, and at small x for the gluon contribution.

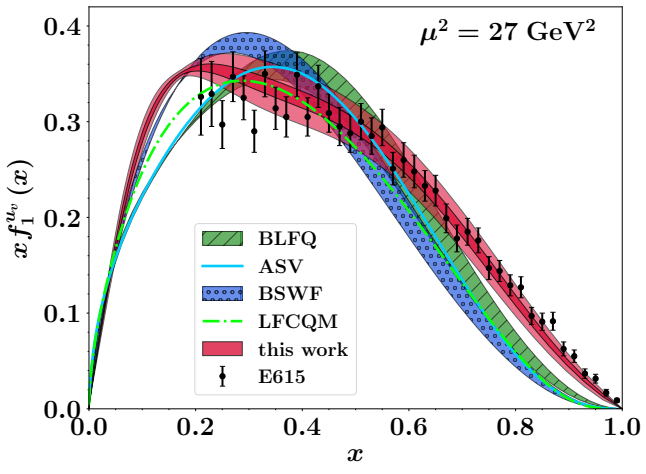
The difference in shape of our results for the valence PDF in the region $0.05 < x < 0.2$ can be attributed to strong correlations between the valence PDF at small x and the gluon PDF at large x. This characteristic is fundamentally tied to the LFWF approach with a finite and small number of Fock states included in the analysis. Let us consider, for example, the LFWA of the Fock state $q\bar{q}'gg$. This function contributes both to the total valence PDF in Eq. (3.16) when the active parton is the quark with momentum fraction $x_1 = x$, and to the gluon PDF in Eq. (3.20) when the active gluon has momentum fraction $x_3 = x$ or $x_4 = x$. Due to momentum conservation, we have

the constraint $\sum_{i=1}^{4} x_i = 1$, which implies that small values of x_1 correspond to high values of x_3 and x_4, and vice versa. Consequently, the behavior of the valence PDF at large x is closely intertwined with the behavior of the gluon PDF at small x, and vice versa. The same pattern applies to other Fock states that contribute to more than one PDF. These spurious correlations tend to decrease as the expansion in the Fock space includes a larger number of Fock components. However, extending the current formalism to higher-order Fock components could become quite intricate.

PDFs at 27 GeV^2

In Fig. 3.8, we report our results at $\mu^2 = 27$ GeV^2 in comparison with other studies and with the extraction of the E615 experiment [55].

The experimental data are obtained by neglecting threshold resummation effects [73–81], as well as we did in our analysis. However, in a seminal paper, Aicher, Schäfer, and Vogelsang (ASV) [82] found that the effects arising from threshold resummation are particularly important for the large-x asymptotic behavior of the valence quark contribution, as shown by the solid cyan curve in the plot of u_v in Fig. 3.8. A similar shape at large-x is reproduced very well from the light-front constituent quark model (LFCQM) predictions of Ref. [83] (green dashed-dotted curve), which considered only the first Fock component and applied NLO evolution to the relevant experimental scale. The same holds for the results of the basis light-front quantization (BLFQ) collaboration [84] within a light-front model including the $q\bar{q}'$ and $q\bar{q}'g$ Fock components as well as for the study of Ref. [85] with Bethe-Salpeter wave functions (BSWF) for the $q\bar{q}'$.



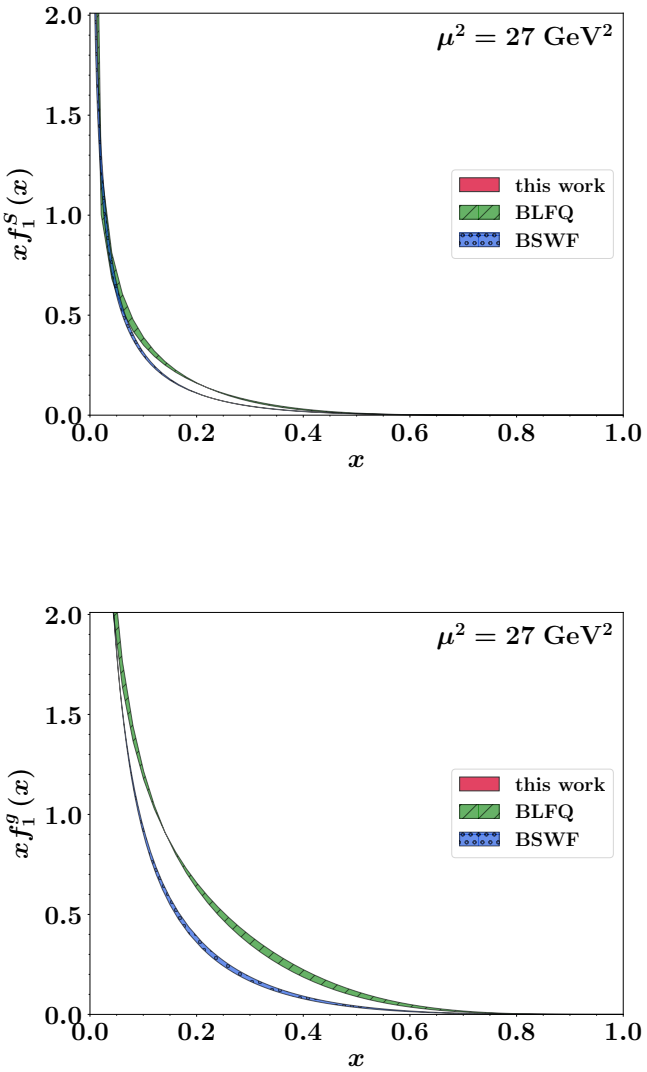


Figure 3.8: xf_1 as function of x for the u_v (upper panel), total sea (central panel) and gluon (lower panel) contributions at $\mu^2 = 27$ GeV^2. Our curves (dark and light red bands corresponding to the 1σ and 3σ CL) are compared with the BSWF results from Ref. [86] (blue bands) and the analysis from the BFLQ collaboration [84] (green bands). For the u-valence contribution we show also the ASV parametrization of Ref. [82] (cyan solid curve) and the results within the LFCQM of Ref. [83] (green dashed-dotted curve), while the data are from the E615 experiment [55] without threshold resummation effects.

From the previous plots, we observe that our results for the u-valence PDF is in very good agreement with the extraction of the E615 experiment [55], while they differ from the other theoretical parametrizations, due to the omission of threshold resummation effects. Moreover, by looking at the total sea and gluon distributions, it is interesting to notice the effects of includ-

ing the non-perturbative sea-quark and gluon PDFs at the initial scale. Our parametrizations take into account non-perturbative sea and gluon contributions at the initial hadronic scale. In contrast, the BSWF approach generates both the sea and gluon contributions solely from the scale evolution, and the BLFQ model includes only a dynamical gluon contribution at the initial scale and generates the sea PDF perturbatively. The difference is particularly evident in the range $x \geq 0.1$ where the contributions generated only perturbatively are lower than those incorporating non-perturbative contributions at the initial scale. Despite these differences, the first moments of the PDFs, defined as $\int dx\, x f_1(x)$, are well compatible within the error bars, as shown in Tab. 3.2. In the same table, we also collect the results of other studies at different scales.

We observe that, across the various scales, the central values for the gluon distribution are significantly lower in the case of xFitter as compared to JAM and our results. Nevertheless, they remain compatible within the error bars.

Comparing our results with the recent BCP extractions, a striking concordance becomes evident for the valence moments. However, our values for the sea contribution appear smaller, mainly because of the distinct behavior of the sea PDFs at $x \lesssim 0.1$, as shown, for example, in Fig. 3.7 for the results at $\mu^2 = 5$ GeV^2. We also notice that our results are in very good agreement with the valence contribution obtained in a recent study that exploits an approach based on the Bethe-Salpeter equation (BSE) [87]. Both of these studies yield higher values than the calculations within the BSWF and BLFQ approaches. This larger valence contribution in our model compared to the BSWF and BLFQ results is counterbalanced by smaller values for the moments of the gluon contributions.

The lattice calculations for the gluon contribution showed significant changes between the analysis of Ref. [88], which used quenched QCD and a large 800 MeV pion mass, and the study of Ref. [89], which employed clover fermion action and 450 MeV pion mass. Additionally, we present the results for the gluon contribution from a recent calculation [90]. This work provided, for the first time, the decomposition into gluon and quark contributions. They calculated the total $u + d$, s, and c contributions, from which it is not possible to reconstruct the separate valence and sea contributions. They obtained that the sum over all quark flavours is $\sum_q \langle x f_1^q(x) \rangle = 0.68 \pm 0.05^{+0.00}_{-0.03}$, while the sum of all contributions amounts to $1.20 \pm 0.13^{+0.00}_{-0.03}$, which is compatible with the expected value of 1 within a two-sigma range. Furthermore, there exists a new lattice study [91] which explored the x-dependence of the gluon PDF at $\mu^2 = 4$ GeV^2. However, they present the results for $x f_1^g(x)$ normalized to unity which prevents to deduce the value of the first moment.

The lattice calculations for the valence contribution obtained substantially smaller values than the phenomenological analyses, apart from [92], where they considered an ensemble of two degenerate light, a strange and a charm quark and reproduced a pion mass of 260 MeV.

	μ^2 GeV^2	$\langle x f_1^v \rangle$	$\langle x f_1^S \rangle$	$\langle x f_1^g \rangle$
JAM-Res [70]	1.61	0.53 ± 0.02	0.14 ± 0.04	0.34 ± 0.06
BLFQ [84]	1.69	0.536	0.069	0.395
JAM [93]	1.69	0.54 ± 0.01	0.16 ± 0.02	0.30 ± 0.02
xFitter [53]	1.69	0.55 ± 0.06	0.26 ± 0.15	0.19 ± 0.16
This work	1.69	0.58 ± 0.03	0.09 ± 0.04	0.33 ± 0.06
Latt1 [88]	4			$0.37 \pm 0.08 \pm 0.12$
Latt2 [89]	4			0.61 ± 0.09
Latt3 [94]	4	0.415 ± 0.0212		
Latt4 [95]	4	0.376 ± 0.112		
Latt5 [90]	4			$0.52 \pm 0.11^{+0.02}_{-0.00}$
BLFQ [84]	4	0.484	0.094	0.421
BSWF [86]	4	0.47 ± 0.02	0.11 ± 0.02	0.41 ± 0.02
xFitter [53]	4	0.50 ± 0.05	0.25 ± 0.13	0.25 ± 0.13
BSE [87]	4	0.52		
This work	4	0.52 ± 0.03	0.11 ± 0.03	0.37 ± 0.05
JAM [93]	5	0.48 ± 0.01	0.17 ± 0.01	0.35 ± 0.02
xFitter [53]	5	0.49 ± 0.05	0.25 ± 0.12	0.26 ± 0.13
BCP [71]	5	0.50 ± 0.01	0.19 ± 0.012	0.31 ± 0.002
This work	5	0.51 ± 0.03	0.12 ± 0.03	0.37 ± 0.05
BLFQ [84]	10	0.446	0.115	0.439
JAM [93]	10	0.44 ± 0.01	0.19 ± 0.01	0.37 ± 0.02
xFitter [53]	10	0.46 ± 0.02	0.22 ± 0.08	0.31 ± 0.06
BCP [71]	10	0.48 ± 0.08	0.21 ± 0.012	0.33 ± 0.015
This work	10	0.48 ± 0.03	0.13 ± 0.02	0.39 ± 0.05
Latt4 [95]	27	0.330 ± 0.018		
Latt6 [92]	27	0.58 ± 0.016		
BSWF [86]	27	0.41 ± 0.04	0.14 ± 0.02	0.45 ± 0.02
xFitter [53]	27	0.42 ± 0.04	0.25 ± 0.10	0.32 ± 0.10
BSE [87]	27	0.44		
This work	27	0.45 ± 0.02	0.15 ± 0.02	0.40 ± 0.04

Table 3.2: Results for the first moments of the total valence, total Sea and gluon distribution from our extractions (grey rows), in comparison with other phenomenological analyses, model calculations and lattice-QCD studies.

3.5.3 Evolution

The dependence of all the PDFs on the QCD scale is guided by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) [96] partial differential equation:

$$\mu^2 \frac{\partial \mathbf{f}\,(x,\mu)}{\partial \mu^2} = \mathbf{P}^{\mathrm{QCD}}\,(x, \alpha_S(\mu)) * \mathbf{f}\,(x,\mu)\,, \tag{3.36}$$

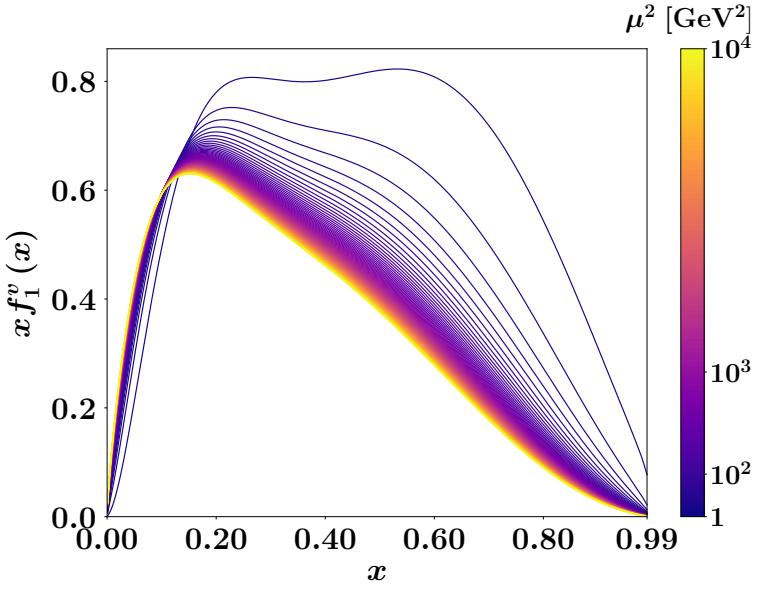
where $\mathbf{f}\,(x,\mu)$ is a vector containing the PDFs of all the partons (quarks with specific flavours and gluon) and $\mathbf{P}^{\mathrm{QCD}}\,(x,\alpha_S(\mu))$ is the QCD matrix of splitting functions. The operator $*$ in Eq. (3.36) stands for *convolution*. Given two functions D_1 and D_2, we define their convolution as:

$$(D_1 * D_2)\,(x) := \int_x^1 \frac{dz}{z} D_1(z) D_2(x/z). \tag{3.37}$$

The solution of the DGLAP equation provides an operator $\mathbf{\Gamma}^{\mathrm{QCD}}$ that evolves the PDFs $\mathbf{f}$ from the initial scale μ_0 to the final scale μ_1 according to

$$\mathbf{f}\,(x,\mu_1) = \mathbf{\Gamma}^{\mathrm{QCD}}(x|\mu_0,\mu_1) * \mathbf{f}\,(x,\mu_0)\,. \tag{3.38}$$

To better visualize the effects of the evolution, we report in Fig. 3.9 the results for the total valence, the total Sea and the gluon PDFs, multiplied by x, as function of x over the range $\mu \in [1, 100]$ GeV.



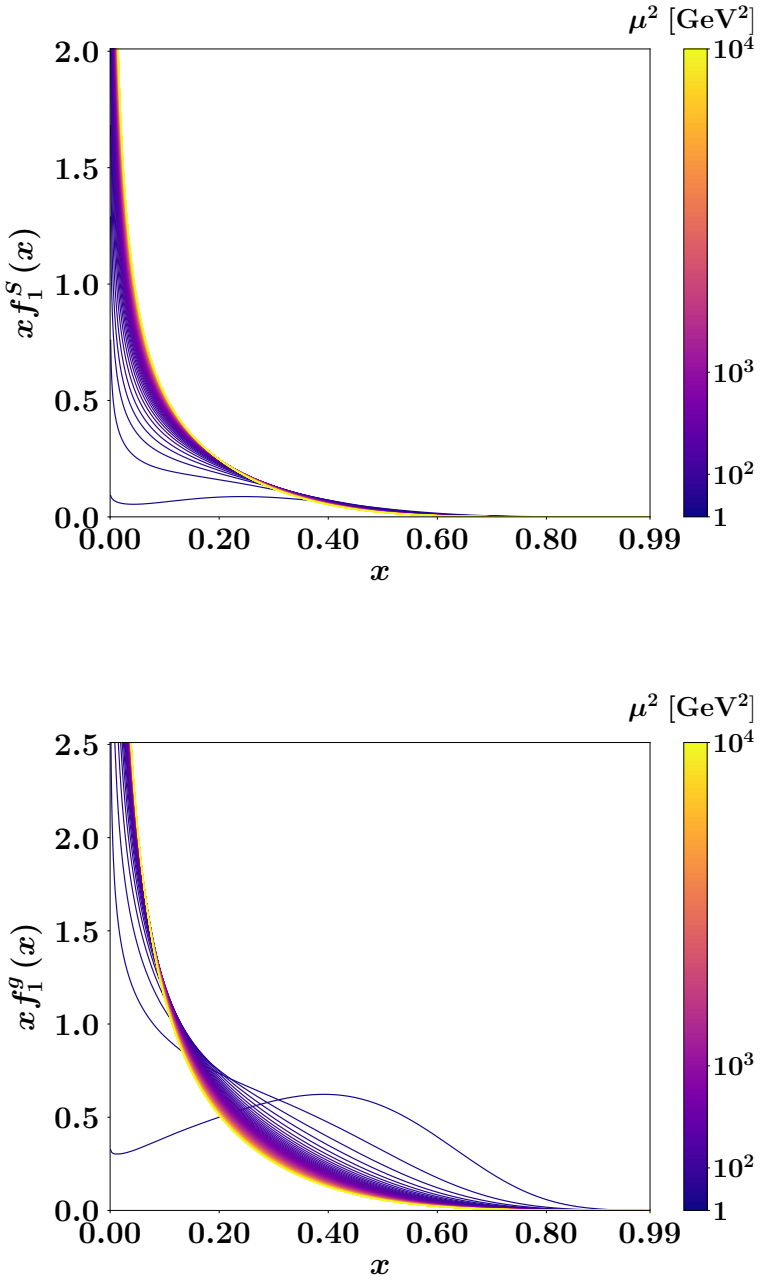


Figure 3.9: Evolution within the scale range $\mu^2 \in [1, 10^4]$ GeV^2 of the total valence, the total Sea and the gluon PDFs, multiplied by x, as function of x. The colour bar indicates the scale dependence, ranging from lower μ^2 values in dark blue to higher μ^2 values in yellow.

From Fig. 3.9 we distinguish a clear effect of the evolution: as the energy scale increases, the maximum value of the valence distribution decreases and shifts leftward. This implies that it is more likely to find a valence PDF with quarks and antiquarks carrying smaller fractions of momenta with respect to lower scales. Moreover, the area under the curve, corresponding to the valence

first moment, decreases as the μ increases. On the contrary, due to the momentum sum rule, the contributions of the gluon and sea PDFs increase. This implies that one can find more gluons and sea pairs carrying higher fraction of momentum inside the pions. This effect can be intuitively justified as the natural consequence of having more available energy to create from the QCD vacuum more sea pairs and gluons endowed with higher momenta.

3.5.4 Behaviour for $x \to 1$

As x approaches 1, the total valence quark distribution in the pion is expected to follow a power-law behavior given by

$$x\, f_1^v(x) \sim (1-x)^\beta\,. \tag{3.39}$$

The value of the asymptotic β is not precisely known from first principles in QCD, due to the non-perturbative nature of the theory and an ongoing debate persists in this regard. Perturbative QCD models generally predict a value of β around 2, while some non-perturbative models tend to favor smaller values of $\beta \approx 1$ [86, 97–114]. There is also the hypothesis that a β value not equal to 2 could potentially conflict with QCD itself [115].

To shed light on the problem of the pion valence behaviour at high x, future experimental programs have been recently proposed [67].

One possible non-perturbative approach to extract information about the asymptotic exponent is to perform fits to the experimental data by using phenomenological models or theory-based parametrizations which permit to extrapolate the best β that reproduces the valence PDF behaviour at high x.

Unfortunately, the coefficient β is not universally agreed within different non-perturbative parametrizations. In some cases the asymptotic exponent results $\beta \approx 1$, as in xFitter [53] and in Refs. [93, 116, 117]. In other parametrizations, such as the already mentioned ASV [82], BSWF [86] and other works [70, 118], the value tends to be higher, up to $\beta \geq 2$.

The procedure we use to recover the exponent β is to compute the following limit

$$\beta_{\text{eff}}(\mu) = \lim_{x\to 1} \frac{\partial \left|\log\left(x\, f_1^v(x,\mu)\right)\right|}{\partial \log(1-x)} \tag{3.40}$$

with $\beta_{\text{eff}}(\mu)$ being the asymptotic exponent at the scale μ.

The latter expression has been computed numerically by taking into account the entire set of our 1000 replicas. First, we explored the behaviour of β_{eff} for fixed value of μ. The results are shown in Fig. 3.10 for $\mu^2 = 27\,\text{GeV}^2$ and for the JAM [70] initial scale $\mu^2 = m_c^2 = (1.27)^2\,\text{GeV}^2$. In this figure, we can observe that the asymptotic exponents resulting from the various replicas form a quite stable band for the different values of $x > 0.75$, except for the region of very high $x \geq 0.93$, where the band fans out. This effect can be attributed to numerical extrapolations, considering that the phenomenological extraction covers the range up to $x \approx 0.94$.

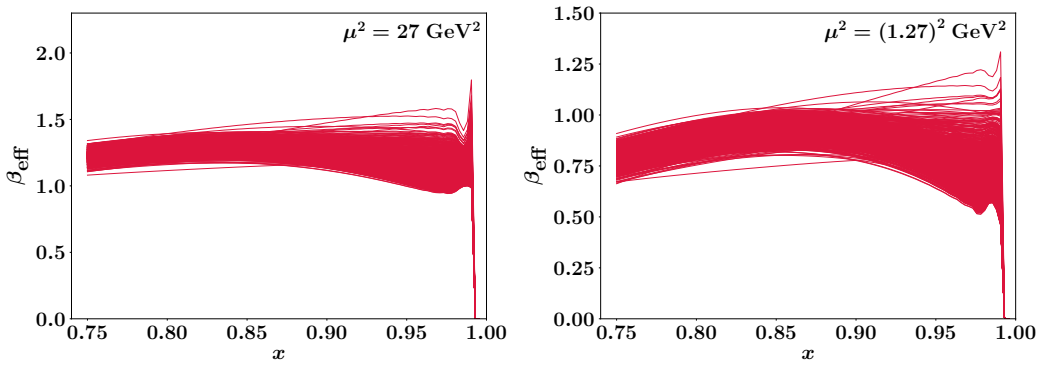


Figure 3.10: x-dependence of the asymptotic exponent for the 1000 replicas of our analysis at fixed $\mu^2 = 27\,\text{GeV}^2$ (left panel) and $\mu^2 = (1.27)^2$ GeV^2 (right panel).

To avoid such extrapolation effects, the study of the μ-dependence of $\beta_{\rm eff}(\mu)$ has been performed in the vicinity of the point $x = 0.92$. We report our results in Fig. 3.11, in comparison with those we computed using the replicas provided by the JAM study without threshold resummation [70].

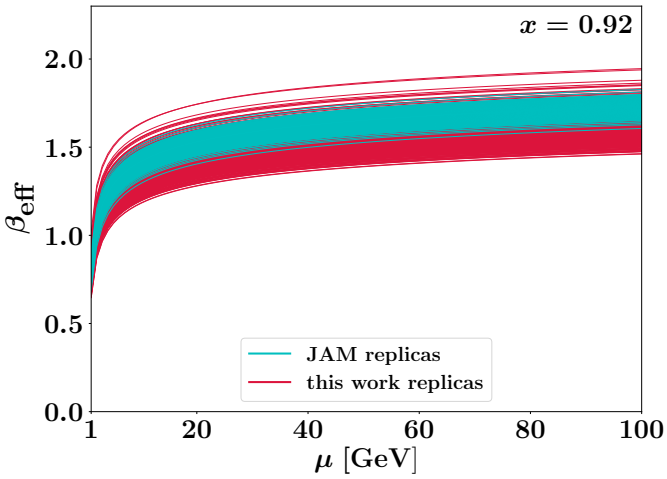


Figure 3.11: $\beta_{\rm eff}$ as function of μ, in comparison between all the replicas from our model (in red) and from JAM [70] (in light-blue), at fixed $x = 0.92$.

We can observe a very similar overall behaviour of the β exponent as function of the energy scale. As the μ scale increases, the asymptotic exponent increases as well, implying a steeper decrease of the valence PDF at high x, as it clear also from Fig. 3.9. We can appreciate that the two extractions in Fig. 3.11 are compatible inside the error bars, for all the values of the scale $\mu \in [1, 100]$ GeV.

The value of the β exponent extracted from this work at the JAM scale is

$$\beta_{\text{eff}}(1.27\,\text{GeV}) = 0.88 \pm 0.05, \tag{3.41}$$

resulting compatible with the mean value $\beta_{\text{eff}} \approx 1$ obtained by JAM without threshold resummation at the same scale.

In general, the inclusion of threshold resummation [119, 120] tends to increase the value of β_{eff}. This means that the net effect of threshold resummation is to reduce the valence PDF at very high x [82], as anticipated in Section 3.5.2. In Ref. [70] different approaches to treat threshold resummation are analyzed in detail, yielding very different results. These range from $\beta_{\text{eff}} \approx 1.2$ with the double Mellin approach, to $\beta_{\text{eff}} > 2$ with the NLL expansion and the cosine approach.

Chapter 4

Pion electromagnetic form factor

The main subject of this Chapter is the parametrization and the fit of the pion electromagnetic (e.m.) form factor (FF). The first part is devoted to the definition of the FF, the LFWA overlap representation and its parametrization. In the last part of the discussion, the fit procedure is analyzed in detail and the results of the fit are presented in comparison with available experimental data.

4.1 Matrix element for the FF

Studying the pion e.m. form factor provides various perspectives on the structure of the pion. It probes the charge distribution in the pion and serves as a valuable tool for exploring the transition toward the perturbative QCD regime in exclusive processes as energy increases [121, 122]. It has found successful descriptions across a range of light-front quark models [84, 123–134], and has seen significant progress in lattice calculations [135–149]. The e.m. FF is obtained from the off-forward matrix element of a local bi-quark operator, which can be seen as the GTMD correlator (2.123) integrated over the momentum variables x and $\boldsymbol{k}_\perp$ (cfr. Fig. 2.5), i.e.,

$$F_1(\Delta) = \frac{1}{2P^+} \langle \pi(p') | \, \bar{\Psi}^q(0) \gamma^+ \Psi^q(0) \, | \pi(p) \rangle \, . \tag{4.1}$$

As in the case of the PDFs, we consider positive pions, i.e., $F_1(\Delta) \equiv F_{1\pi^+}(\Delta)$. By charge conjugation, the FF for positive and negative charged pion are simply related by

$$F_{1\pi^+}(\Delta) = -F_{1\pi^-}(\Delta). \tag{4.2}$$

4.2 LFWA overlap representation for pion FF

Using the definition in Eq. (4.1), we can proceed to derive the LFWA overlap representation of the pion e.m. FF by inserting the Fock-state expansion for

the pion in the initial and final state. As we are dealing with a non-diagonal matrix element, the parton configurations in the LFWAs of the initial and final states differ in their kinematics. Consequently, we no longer arrive at the squared LFWAs, as in the case of the PDFs. This implies that the second equation of (2.158) is inapplicable when we parametrize the FF using our model. Consequently, the FF explicitly depends on the set of transverse parameters $\mathcal{A}^T$, as shown in Section 4.3.

The computation is derived in the *symmetric frame*, in which the initial and final momentum of the pion are, respectively

$$\begin{aligned} p &= P - \frac{\Delta}{2}, \\ p' &= P + \frac{\Delta}{2}. \end{aligned} \tag{4.3}$$

We take the momentum transfer in the transverse direction, i.e.,

$$\Delta \equiv (0, 0, \boldsymbol{\Delta}_\perp)\,, \tag{4.4}$$

and we define

$$Q^2 = -\Delta^2 = \boldsymbol{\Delta}_\perp^2. \tag{4.5}$$

The problem of pions in this context is that they cannot be used as fixed targets, because they decay very rapidly. Therefore, to get access to the pion e.m. form factor two main processes have been employed. Either the elastic scattering of pions off atomic electrons or the Sullivan process [65], wherein a virtual pion is firstly emitted from a nucleon, and then it interacts with an electron producing a real pion in the final state (see Fig. 4.1). If we consider, for example, a Sullivan process, we notice that Q^2 and Δ coincide, respectively, with the *virtuality* and the momentum of the photon exchanged in the interaction between the pion and the electromagnetic probe.

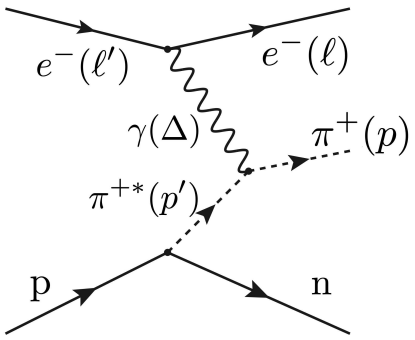


Figure 4.1: Diagram for the Sullivan process, where the virtual photon emitted from a lepton probe interacts with the pion cloud of the proton target. The photon has momentum $\Delta = \ell' - \ell = p' - p$ and virtuality $Q^2 = -\Delta^2$.

The complete expression for the LFWA overlap representation of the e.m. FF can be split, as in the case of the PDFs, in the different contributions

arising from the four distinct Fock components (2.134) - (2.137):

$$F_1(Q^2) = F_{1,u\bar{d}}(Q^2) + F_{1,u\bar{d}g}(Q^2) + F_{1,u\bar{d}gg}(Q^2) + \sum_{\{s\bar{s}\}} F_{1,u\bar{d}\{s\bar{s}\}}(Q^2), \tag{4.6}$$

where

$$F_{1,u\bar{d}}(Q^2) = 2\int \mathrm{d}[1]\mathrm{d}[2]\sqrt{x_1x_2}\,\psi^{*(1)}_{q\bar{q}}\,(x_1,x_2,\boldsymbol{k}'_{\perp 1},\boldsymbol{k}'_{\perp 2})\,\psi^{(1)}_{q\bar{q}}(x_1,x_2,\boldsymbol{k}_{\perp 1},\boldsymbol{k}_{\perp 2}), \tag{4.7}$$

$$\begin{aligned} F_{1,u\bar{d}g}(Q^2) = 2\int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3]\sqrt{x_1x_2x_3} \\ \times \psi^{*(1)}_{q\bar{q}g}\,(x_1,x_2,x_3,\boldsymbol{k}'_{\perp 1},\boldsymbol{k}'_{\perp 2},\boldsymbol{k}'_{\perp 3})\,\psi^{(1)}_{q\bar{q}g}(x_1,x_2,x_3,\boldsymbol{k}_{\perp 1},\boldsymbol{k}_{\perp 2},\boldsymbol{k}_{\perp 3}), \end{aligned} \tag{4.8}$$

$$\begin{aligned} F_{1,u\bar{d}gg}(Q^2) = 4\int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3]\mathrm{d}[4]\sqrt{x_1x_2x_3x_4} \\ \times \Big[\psi^{*(1)}_{q\bar{q}gg}(x_1,x_2,x_3,x_4,\boldsymbol{k}'_{\perp 1},\boldsymbol{k}'_{\perp 2},\boldsymbol{k}'_{\perp 3},\boldsymbol{k}'_{\perp 4}) \\ \times \psi^{(1)}_{q\bar{q}gg}(x_1,x_2,x_3,x_4,\boldsymbol{k}_{\perp 1},\boldsymbol{k}_{\perp 2},\boldsymbol{k}_{\perp 3},\boldsymbol{k}_{\perp 4}) \\ + \psi^{*(1)}_{q\bar{q}gg}(x_1,x_2,x_4,x_3,\boldsymbol{k}'_{\perp 1},\boldsymbol{k}'_{\perp 2},\boldsymbol{k}'_{\perp 4},\boldsymbol{k}'_{\perp 3}) \\ \times \psi^{(1)}_{q\bar{q}gg}(x_1,x_2,x_3,x_4,\boldsymbol{k}_{\perp 1},\boldsymbol{k}_{\perp 2},\boldsymbol{k}_{\perp 3},\boldsymbol{k}_{\perp 4}) \\ + \psi^{*(2)}_{q\bar{q}gg}(x_1,x_2,x_3,x_4,\boldsymbol{k}'_{\perp 1},\boldsymbol{k}'_{\perp 2},\boldsymbol{k}'_{\perp 2},\boldsymbol{k}'_{\perp 4}) \\ \times \psi^{(2)}_{q\bar{q}gg}(x_1,x_2,x_3,x_4,\boldsymbol{k}_{\perp 1},\boldsymbol{k}_{\perp 2},\boldsymbol{k}_{\perp 3},\boldsymbol{k}_{\perp 4}) \\ - \psi^{*(2)}_{q\bar{q}gg}(x_1,x_2,x_4,x_3,\boldsymbol{k}'_{\perp 1},\boldsymbol{k}'_{\perp 2},\boldsymbol{k}'_{\perp 4},\boldsymbol{k}'_{\perp 3}) \\ \times \psi^{(2)}_{q\bar{q}gg}(x_1,x_2,x_3,x_4,\boldsymbol{k}_{\perp 1},\boldsymbol{k}_{\perp 2},\boldsymbol{k}_{\perp 3},\boldsymbol{k}_{\perp 4})\Big], \end{aligned} \tag{4.9}$$

$$\begin{aligned} F_{1,u\bar{d}\{s\bar{s}\}}(Q^2) = 4\int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3]\mathrm{d}[4]\sqrt{x_1x_2x_3x_4} \\ \times \Big[\psi^{*(1)}_{q\bar{q}gg}(x_1,x_2,x_3,x_4,\boldsymbol{k}'_{\perp 1},\boldsymbol{k}'_{\perp 2},\boldsymbol{k}'_{\perp 3},\boldsymbol{k}'_{\perp 4}) \\ \times \psi^{(1)}_{q\bar{q}gg}(x_1,x_2,x_3,x_4,\boldsymbol{k}_{\perp 1},\boldsymbol{k}_{\perp 2},\boldsymbol{k}_{\perp 3},\boldsymbol{k}_{\perp 4}) \\ + \psi^{*(2)}_{q\bar{q}gg}(x_1,x_2,x_3,x_4,\boldsymbol{k}'_{\perp 1},\boldsymbol{k}'_{\perp 2},\boldsymbol{k}'_{\perp 3},\boldsymbol{k}'_{\perp 4}) \\ \times \psi^{(2)}_{q\bar{q}gg}(x_1,x_2,x_3,x_4,\boldsymbol{k}_{\perp 1},\boldsymbol{k}_{\perp 2},\boldsymbol{k}_{\perp 3},\boldsymbol{k}_{\perp 4}) \\ + \psi^{*(3)}_{q\bar{q}gg}(x_1,x_2,x_3,x_4,\boldsymbol{k}'_{\perp 1},\boldsymbol{k}'_{\perp 2},\boldsymbol{k}'_{\perp 3},\boldsymbol{k}'_{\perp 4}) \\ \times \psi^{(3)}_{q\bar{q}gg}(x_1,x_2,x_3,x_4,\boldsymbol{k}_{\perp 1},\boldsymbol{k}_{\perp 2},\boldsymbol{k}_{\perp 3},\boldsymbol{k}_{\perp 4})\Big]. \end{aligned} \tag{4.10}$$

The "primed" variables in the previous equations are:

$$\begin{cases} \boldsymbol{k}'_{\perp 1} = \boldsymbol{k}_{\perp 1} + (1-x_1)\boldsymbol{\Delta}_\perp \\ \boldsymbol{k}'_{\perp i} = \boldsymbol{k}_{\perp i} - x_i\boldsymbol{\Delta}_\perp \quad \text{for } i \neq 1 \end{cases}. \tag{4.11}$$

4.3 Explicit model

If we now insert for the LFWAs the explicit parametrization introduced in Section 2.4.2 and compute all the integrals, we find the complete expression for the FF in terms of the model parameters. Nevertheless, the final result is extremely long, non-trivial and quite uninformative.

We have chosen to present here only the unintegrated form of the various Fock-state contributions (4.7) - (4.10):

$$F_{1,u\bar{d}}(Q^2) = \frac{1}{2}\mathcal{C}^v_{u\bar{d}} \int_0^1 dx \exp\left[-a^2_{q\bar{q}'}Q^2\frac{\bar{x}}{2x}\right] \times (x\bar{x})^{2\gamma_q-1}\left[1+d_{q1}(1+2\gamma_q)\left(1+\gamma_q(x-\bar{x})^2-6x\bar{x}\right)\right]^2, \tag{4.12}$$

$$F_{1,u\bar{d}g}(Q^2) = \frac{1}{2}\mathcal{C}^v_{u\bar{d}g} \int_0^1 dx \exp\left[-a^2_{q\bar{q}'g}Q^2\frac{\bar{x}}{2x}\right] \times x\bar{x}^5\left(3+18xd_{g1}-10\bar{x}d_{g1}+13d^2_{g1}+14xd^2_{g1}(x-4\bar{x})\right), \tag{4.13}$$

$$F_{1,u\bar{d}gg}(Q^2) = \frac{1}{2}\mathcal{C}^v_{u\bar{d}gg} \int_0^1 dx \ \exp\left[-a^2_{q\bar{q}'gg}Q^2\frac{\bar{x}}{2x}\right] x\bar{x}^9, \tag{4.14}$$

$$F_{1,u\bar{d}\{s\bar{s}\}}(Q^2) = \frac{1}{2}\mathcal{C}^v_{u\bar{d}s\bar{s}} \int_0^1 dx \exp\left[-a^2_{q\bar{q}'s\bar{s}}Q^2\frac{\bar{x}}{2x}\right] x\bar{x}^5. \tag{4.15}$$

The expressions in Eqs. (4.12) - (4.15) have been obtained using the following simplifications for the transverse-dependent part of the LFWAs

$$\begin{aligned}\Omega^{(1)}_{4,q\bar{q}'gg} &= \Omega^{(2)}_{4,q\bar{q}'gg},\\ \Omega^{(1)}_{4,q\bar{q}'s\bar{s}} &= \Omega^{(2)}_{4,q\bar{q}'s\bar{s}} = \Omega^{(3)}_{4,q\bar{q}'s\bar{s}}\end{aligned} \tag{4.16}$$

which correspond to the following relations between the fitting parameters

$$\begin{aligned}a_{q\bar{q}'gg} &= a^{(1)}_{q\bar{q}'gg} = a^{(2)}_{q\bar{q}'gg},\\ a_{q\bar{q}'s\bar{s}} &= a^{(1)}_{q\bar{q}'s\bar{s}} = a^{(2)}_{q\bar{q}'s\bar{s}} = a^{(3)}_{q\bar{q}'s\bar{s}}.\end{aligned} \tag{4.17}$$

These assumptions have been verified by performing various fits in which the number of independent LFWAs associated to the same Fock state was varied. The result is that the fit is not able to distinguish between a configuration in which the conditions (4.17) are enforced and a parametrization with distinct parameters.

We observe that the valence sum rule for the PDF, given in the first line of Eq. (3.25), is reflected on the e.m. FF by the condition:

$$F_1(Q^2=0) = 1, \tag{4.18}$$

which is satisfied by construction using our parametrization.

4.4 Fit procedure

As it is clear from Eqs. (4.12) - (4.15), the e.m. FF depends both on the collinear set of parameters $\mathcal{A}^L$ and on the transverse set of parameters $\mathcal{A}^T$.

Our strategy is to use the best fit parameters obtained from the fit of the PDFs to fix $\mathcal{A}^L$, and then determining $\mathcal{A}^T$ by fitting the available experimental data of the pion form factor. The tricky part consists in propagating the uncertainties entering in the collinear part of the LFWAs, which were fitted to the PDFs, to the FFs. We describe the method used in Section 4.4.2.

4.4.1 Data set

The experimental data for the pion e.m. FF at low momentum transfer ($0.015\ \text{GeV}^2 \leq Q^2 \leq 0.253\ \text{GeV}^2$) come from scattering experiments of pions off atomic electrons conducted at Fermilab [150, 151] and CERN [152, 153]. These data were also used to extract the *charge radius*.

As the momentum transfer increases, up to $Q^2 \sim 10\ \text{GeV}^2$, the pion form factor can be extracted by exploiting the Sullivan mechanism. This process has been studied at Cornell [154–156], DESY [157, 158], and JLab [159–163].

There are proposals from JLab of future measurements [62] that are going to explore the intermediate region, while the kinematics at higher momentum transfer, up to $Q^2 = 30\ \text{GeV}^2$, will potentially be accessed by the future electron-ion collider (EIC) [63].

The total data set is composed of 100 experimental points, that we collect in the following table divided in subsets according to the experiments and their Q^2 range.

Ref.	Q^2 range [GeV^2]	N_{data}
Dal81 [150]	0.0317 - 0.0705	20
Dal82 [151]	0.039 - 0.092	14
Ame86 [152]	0.015 - 0.253	45
Beb78 [156]	1.18 - 9.77	5
Ack78 [158]	0.35	1
Vol01 [159]	0.6 - 1.60	5
Hor06 [161]	1.6 - 2.45	2
Hub08 [163]	0.6 - 2.45	8

4.4.2 χ^2 definition

For the propagation of the experimental uncertainties we use again the bootstrap method. For an in-depth understanding of the procedure and its application to extractions involving multiple data sets with systematic errors, we refer to Ref. [164].

In our particular case, the situation is as follows: we have a set of parameters, $\mathcal{A}^L$, that have been determined from the PDF fit with their own statistics generated by the first bootstrap. These parameters enter in the analytic form

of the form factors, which depend also on the transverse set $\mathcal{A}^T$ to be statistically sampled. Reference [164] discusses exactly how the bootstrap technique is valuable for propagating uncertainties associated with model parameters that are not directly treated as free fitting variables.

The 1000 replicas of the PDFs are associated to an equal number of best-fit parameter vectors $\mathcal{A}_i^L$, with $i = 1, \ldots, 10^3$. We want to generate 10^4 replicas for the e.m. FF, with the aim to obtain statistical meaningful results. For each bootstrap cycle in the FF fit we uniformly extract a set of PDF parameters and generate a replica of the experimental FF data. We then perform the fit of just the free parameters $\mathcal{A}_{m,i}^T$, where $m = 1, \ldots, 10^4$ refers to the FF replica cycle and i keeps trace on the extracted $\mathcal{A}_i^L$ set.

By considering the presence of N_{set} multiple subsets, corresponding to the 8 rows in Tab. (4.4.1), and the existence of $N_{\text{data}j}$ associated to the subset j, the chi-squared definition for the m-th replica is:

$$\chi_m^2 = \sum_{j=1}^{N_{\text{set}}} \sum_{l=1}^{N_{\text{data}_j}} \frac{(E_{mjl} - F_1(Q^2; \mathcal{A}_i^L)_{jl})^2}{\sigma_{mjl}^2}, \tag{4.19}$$

where $F_1(Q^2; \mathcal{A}_i^L)_{jl}$ is the prediction for the pion e.m. form factor computed using as input the m-th uniformly sampled vector of collinear parameters $\mathcal{A}_i^L$. In Eq. (4.19), the bootstrap quantities are defined as

$$\begin{aligned} E_{mjl} &= (1 + \delta_{ml})\,(E_{jl} + \sigma_{mjl}) \\ \sigma_{mjl} &= \gamma_{mjl}\sigma_{jl}, \end{aligned} \tag{4.20}$$

where E_{jl} is the l-th experimental datum of the j-th experimental subset with relative error σ_{jl}, δ_{ml} is a random value extracted from the distribution for the systematic error of set i, and γ_{mjl} is a random number sampled from a normal Gaussian distribution.

From the minimization of all the m functions in Eq. (4.19) we obtain the vector of 10^4 transverse parameters, that can be used to recreate the plot of the form factor with the corresponding error bars.

4.5 Fit results

We recall that the set of transverse parameters $\mathcal{A}^T$ contains four elements, one for each Fock state. However, from some preliminary tests we performed, we noticed that the form-factor fit is not sensitive to the value of the parameter $a_{q\bar{q}'g}$. This can be traced back to the nearly negligible norm associated to the $q\bar{q}'g$ Fock state, especially in contrast to the norms of other Fock state components, as shown in Fig. 4.2.

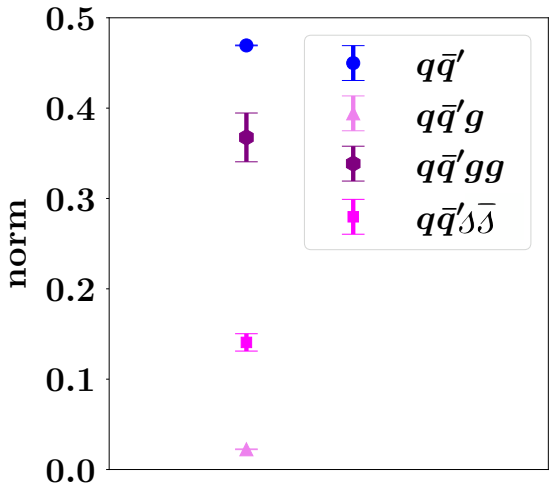


Figure 4.2: The norms of the different Fock states generated from the fit of the PDFs.

The norms of the Fock states depend by construction only on the collinear parameters fitted to the PDF, since we impose the normalization in Eq. (2.158) for the transverse-momentum dependent part of the LFWAs.

We saw in Section 3.5 that the experimental data set for the PDF fit were not sensitive enough to the gluon contribution generated by the $q\bar{q}'g$ component and this eventually guided the fit towards a configuration where the associated norm becomes relatively small.

For these reasons we decided to exclude the parameter $a_{q\bar{q}'g}$ from the fit by fixing it arbitrarily to 1 GeV^{-1}. The choice of this value is totally irrelevant, since in the end the corresponding contribution to the form factor is negligible.

The remaining 3 transverse best fit parameters result (in units of GeV^{-2}):

$$\langle a^2_{q\bar{q}'}\rangle = 1.559, \quad \langle a^2_{q\bar{q}'gg}\rangle = 0.509, \quad \langle a^2_{q\bar{q}'s\bar{s}}\rangle = 0.796. \tag{4.21}$$

We can use these results to extract the multidimensional probability distribution and to construct the correlation coefficients for all pairs of parameters, that we report in Tab. 4.1. We also include the correlation matrix in Fig.4.3 in order to directly visualize all the correlations between the parameters.

	$a^2_{q\bar{q}'}$	$a^2_{q\bar{q}'gg}$	$a^2_{q\bar{q}'s\bar{s}}$
$a^2_{q\bar{q}'}$	0.356	-0.593	-0.656
$a^2_{q\bar{q}'gg}$	-0.593	0.080	0.484
$a^2_{q\bar{q}'s\bar{s}}$	-0.656	0.484	0.309

	d_{1g}	d_{1q}	γ_q	α_1	α_2	α_3
d_{1g}	220.545	-0.260	-0.253	0.277	-0.495	0.002
d_{1q}	-0.260	0.047	0.982	-0.108	0.048	0.245
γ_q	-0.253	0.982	0.100	-0.083	0.079	0.195
α_1	0.277	-0.108	-0.083	0.018	-0.635	-0.605
α_2	-0.495	0.048	0.079	-0.635	0.202	0.054
α_3	0.002	0.245	0.195	-0.605	0.054	0.082

	d_{1g}	d_{1q}	γ_q	α_1	α_2	α_3
$a^2_{q\bar{q}'}$	0.067	-0.050	-0.056	0.056	-0.088	-0.046
$a^2_{q\bar{q}'gg}$	-0.381	0.069	0.099	-0.288	0.599	-0.161
$a^2_{q\bar{q}'s\bar{s}}$	-0.181	0.076	0.081	-0.177	0.273	-0.011

Table 4.1: Correlation coefficients for all the fitted parameters. The diagonal elements in the first two tables coincide with the standard deviations; all the other values are the linear correlation coefficients for the corresponding pairs of parameters.

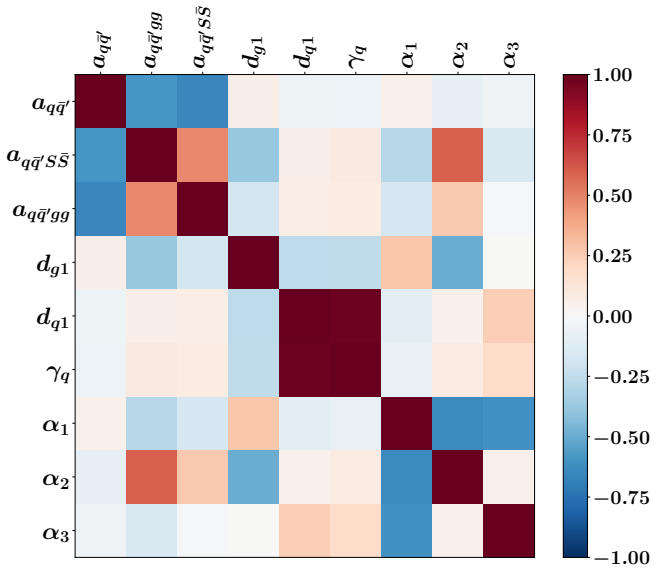


Figure 4.3: Correlation matrix of all the fitted parameters to the PDFs and form factor.

From Tab. 4.1 and Fig. 4.3 we observe strong correlations between the pairs γ_q-d_{q1}, α_1-α_2, α_1-α_3, and between all the transverse parameters.

We come now to the crux of the problem, consisting in finding the *goodness-of-fit distribution*, with the calculation of the confidence level (CL). When we

include the systematic uncertainties of the experimental data and the set of non-fitted parameters, the statistical distribution of the minimization function (4.19) might deviate significantly from the conventional chi-squared distribution. The procedure to adopt is to perform a second bootstrap, by assuming the ideal situation in which the model predictions obtained from the best parameters fitted to the real data (not varied by the first bootstrap) are "correct". This is obtained by replacing E_{jl} with $F_1\left(Q^2; \mathcal{A}^L_{\text{rep}\,0}, \mathcal{A}^T_{\text{rep}\,0}\right)_{jl}$ in the bootstrap quantities (4.20), where $\mathcal{A}^{L/T}_{\text{rep}\,0}$ refers to the set of the collinear/transverse parameters of the "replica-0", representing the fits performed on the real experimental data. By accumulating a substantial statistics of chi-squared values resulting from this second bootstrap, we can reconstruct the goodness-of-fit distribution and compute the p-value associated to the fit results using the two-sided χ^2 test. This analysis allows us to deduce the confidence interval (CI) for the reduced chi-squared $\hat{\chi}^2/N^{\text{FF}}_{d.o.f.}$, where $N^{\text{FF}}_{d.o.f.} = 97$:

$$\hat{\chi}^2/N^{\text{FF}}_{d.o.f.} = 1.194, \quad \text{CI } 68\% = [0.890, 1.204], \quad \text{CI } 99\% = [0.682, 1.593]. \tag{4.22}$$

We observe that the reduced $\hat{\chi}^2$ falls within the 68% CI, as expected.

The results are shown in Fig. 4.4, where we plot the square of the e.m. FF as function of Q^2. Similar to the approach taken for the PDF fit, we have chosen to show only the bands corresponding to 68% (inner dark blue) and 99.7% (outer light blue) uncertainty. They have been obtained with the procedure explained in Section 3.5.2. We stress that the two bands incorporate the error propagation of the PDF parameters, representing therefore more than just the experimental uncertainty on the form factor.

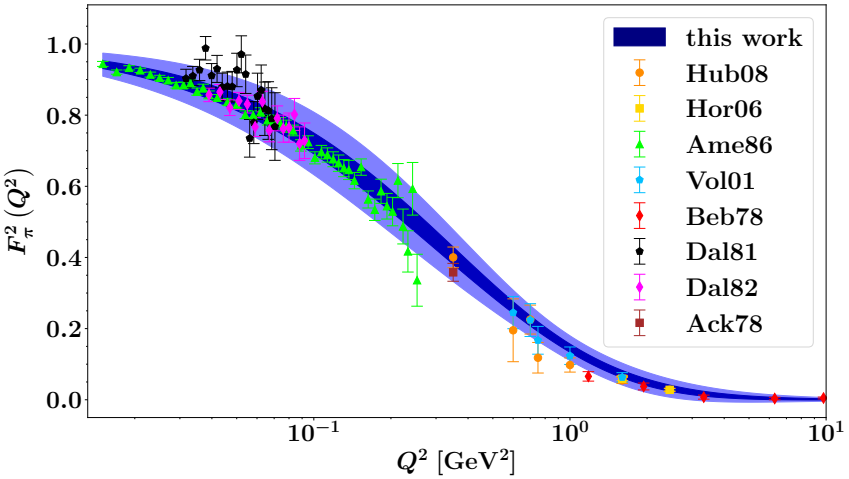


Figure 4.4: Fit results for the square of the pion electromagnetic form factor as function of Q^2. The dark (light) blue band shows the 68% (99.7%) of the replicas. The experimental data correspond to Hub08 [163], Hor06 [161], Ame86 [153], Vol01 [159], Beb78 [156], Dal82 [150, 151], Ack78 [158].

4.6 DA results

Having obtained the whole sets of parameters for the LFWAs from the PDF and e.m. FF, we can now analyze the results for the pion DA in Eq. (2.193). Using the parametrization for $\phi_{q\bar{q}'}(x, 1-x)$ in Eq. (2.164) , we obtain

$$\phi_\pi(x) = \frac{\sqrt{6}}{a_2 \pi f_\pi} n_\pi N^{(1)}_{q\bar{q}'}(x - x^2)^{\gamma_q} \left(1 + d_{q1} C_2^{(\gamma_q + 1/2)}(2x-1)\right), \tag{4.23}$$

where n_π ensures unit normalization for the value of the pion decay constant of

$$f_\pi = 133 \text{ MeV}. \tag{4.24}$$

Using the best fit values of the collinear parameters d_{q1}, γ_q, n_1 and of the transverse parameter a_2 we obtain the results at the scale μ_0 shown in Fig. 4.5. The black band corresponds to the uncertainty band obtained from the 10000 replicas of the e.m. FF fit.

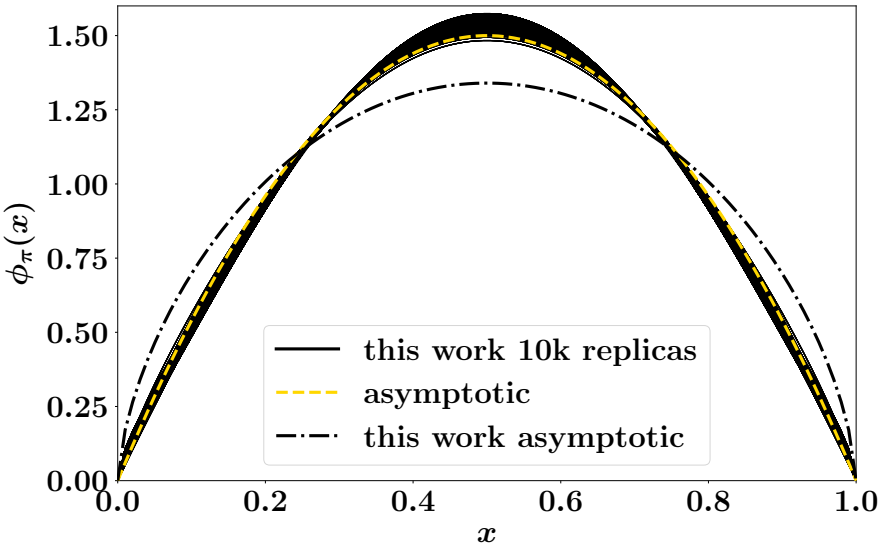


Figure 4.5: Pion DA at twist-2 computed at the initial scale μ_0 for the 10000 replicas of our model (continuous black curves), in comparison with the asymptotic DA with fixed exponent $\gamma_q = 1$ (dashed yellow curve) and with variable exponent (dashed dotted black curve).

The yellow dashed curve refers to the asymptotic DA, i.e.,

$$\phi_{\pi,\,\text{asympt}}(x) = 6x(1-x). \tag{4.25}$$

The asymptotic DA is obtained from the $q\bar{q}'$ state LFWA without any next-to-asymptotic term for the collinear part ($d_{q1} = 0$) and with the fixed exponent $\gamma_q = 1$. This configuration coincides with our initial parametrization attempt

in Eq. (2.160), that we abandoned in favour of the model with the variable exponent and the first next-to-asymptotic contribution, as explained in Section 2.4.2 (see (2.164)).

The presence of the variable exponent γ_q modifies the asymptotic contribution to the DA. We can see the effects of the variable exponent γ_q by comparing the asymptotic DA in Eq. (4.25) shown by the yellow curve with the dashed-dotted curve obtained from the single minimization fit of our model, by imposing $d_{q1} = 0$ and using the best fit value of the exponent γ_q. We notice the difference in shape, with the asymptotic DA having a narrower width balanced by a higher value of the peak in such a way to guarantee the normalization of Eq. (2.129). A similar behaviour has been observed in many other works that make use of variable exponents [165–169]. Nevertheless, it is noteworthy that with the inclusion of the first next-to-asymptotic contribution, the final result for the DA with variable exponent returns to be compatible with the shape of $\phi_{\pi,\,\mathrm{asympt}}(x)$ with fixed exponent.

The two fits for the longitudinal and transverse momentum dependence of the LFWFs have yielded quite satisfactory outcomes, indicating the efficacy of the LFWAs-based approach.

Next step is to trying to combine this information with the analysis of the pion TMDs and GPDs that depend on both $\mathcal{A}^L$ and $\mathcal{A}^T$. Regarding TMDs, our primary objective is to discern if the results of the two previous fits, which are related to distinct but complementary processes, accurately predict - or not - the data related to the TMDs. A positive outcome would yield a significant result, whereas a opposite scenario would prompt the need for a concurrent extraction of FF, PDFs and TMDs. For GPDs, since at present do not exist experimental data, we will limit ourselves to compare the predictions from our model with those from other existing works, providing a comprehensive assessment of the predictive power of our model.

Chapter 5

Pion transverse-momentum dependent parton distributions

Transverse-Momentum Dependent Parton Distributions are more complex objects, if compared to the PDFs and to the FFs. They describe the distribution of the partons inside a hadron in three-dimensional momentum space. While the scarcity of available data prevents us from refining our understanding of the one-dimensional structure of the pion, even less is known about its three-dimensional structure. At present, there are various model calculations for pion TMDs that are discussed in Refs. [83,170–176], but very few phenomenological fits have been performed.

The starting point for the extraction of the TMDs from experimental data are the TMD factorization theorems, which are more challenging with respect to those of the collinear distributions.

In the case of the proton, the factorization theorems have been proven for Semi-Inclusive Deep-Inelastic Scattering (SIDIS), for Drell-Yan (DY) lepton-pair production in hadronic collisions, and for semi-inclusive electron-positron annihilation [16]. Moreover, very accurate analyses of proton unpolarized TMDs have been recently published [177–183].

In the case of the pion, only two studies have been released in Refs. [184, 185]. They concern the extraction of the pion TMDs from the available experimental data that, at present, come only from DY experiments.

This Chapter is structured as follows: we start by introducing the different types of leading-twist TMDs in a hadron and the formalism of the DY process, which allows us to express the DY cross section in terms of TMDs. We then show the logic at the basis of the Nanga Parbat[1] framework, an open source QCD software developed in Pavia for fitting hadron TMDs. We proceed with the fit of pion TMDs following two approaches: a pure phenomenological fit and a fit based on the parametrization of the pion TMDs derived from the model developed in the previous Sections of this work. In this context, we

[1] https://github.com/MapCollaboration/NangaParbat

present the LFWA overlap representation for pion TMDs and cross-check the fitted parameters against those obtained from the e.m. form factor and PDF fits. In the final part we include the comparison of the two fits, both at the level of the fitted observables and of the extracted TMDs.

5.1 TMDs at leading-twist

At leading-twist, various TMDs can be defined, on the basis of the quark and hadron polarization [186]. In this work, the symbols f, g and h are used for TMDs corresponding to an unpolarized, longitudinally polarized, and transversely polarized quark, respectively. In cases where the distribution does not have a collinear counterpart and the parent hadron possesses a non-zero polarization, the TMDs are labelled by introducing a subscript, L or T, that stands for a longitudinally polarized or a transversely polarized hadron, respectively. If the TMD enters in the decomposition of the correlator with a pre-factor depending on $\boldsymbol{k}_\perp$ with an open index, then a superscript $\perp$ is added. The different types of TMDs are presented in Tab. 5.1.

	f	g	h
U	f_1		$h_1^\perp$
L		g_1	$h_{1L}^\perp$
T	$f_{1T}^\perp$	g_{1T}	h_1, $h_{1T}^\perp$

Table 5.1: Existing types of TMDs at leading-twist.

The TMDs on the diagonal have a collinear counterpart, except for the *pretzelocity* $h_{1T}^\perp$ which only exists as $(x, \boldsymbol{k}_\perp)$-dependent distribution, like all the out-off-diagonal distributions. Recalling that the pion is a spin-less meson, the only distribution we address in this Chapter is the unpolarized TMD for an unpolarized quark, $f_1(x, \boldsymbol{k}_\perp)$.

5.2 Drell-Yan formalism

In the Drell-Yan process

$$h_A(p_A) + h_B(p_B) \to \gamma^*(q) + X \to \ell(l') + \bar{\ell}(l) + X, \tag{5.1}$$

two hadrons h_A and h_B of momentum p_A and p_B, respectively, and center of mass energy squared $s = (p_A + p_B)^2$ produce a neutral vector boson γ^*/Z with four-momentum q and large invariant mass $Q = \sqrt{q^2}$. The vector boson then decays into a lepton - antilepton pair $(\ell\bar{\ell})$ with final momenta constrained by the momentum conservation, $q = l + l'$. According to the parton model assumptions, the active particles involved in the DY process are two partons,

each located within one of the two hadrons. We make the hypothesis that the two partons are a quark of h_A and an antiquark of h_B.

The two hadron momenta identify the longitudinal direction z and, consequently, the transverse momentum of the vector boson $\boldsymbol{q_T}$, while the pseudo-rapidity of the vector boson is

$$y = \frac{1}{2} \ln \left(\frac{q^+}{q^-} \right). \tag{5.2}$$

Since the pions are spinless particles, we are interested in unpolarized cross-sections integrated over the azimuthal angle of the exchanged boson that lead, in the end, to the extraction of the pion unpolarized TMD.

This cross section can be written in terms of two structure functions F^1_{UU}, F^2_{UU} [187,188]. In the limit in which the mass of the incoming hadrons M_h and the $\boldsymbol{q}_T$ are in the kinematic conditions $M_h^2 \ll Q^2$ and $\boldsymbol{q}_T^2 \ll Q^2$, the structure function F^2_{UU} is suppressed and the cross section can be expressed as

$$\frac{d\sigma^{DY}}{d|\boldsymbol{q}_T|\, dy\, dQ} \simeq \frac{16\pi^2 \alpha_{em}^2}{9Q^3} |\boldsymbol{q}_T| F^1_{UU}(x_A, x_B, |\boldsymbol{q}_T|, Q), \tag{5.3}$$

where α_{em} is the electromagnetic coupling, $x_{A,B} = \dfrac{Q}{\sqrt{s}} e^{\pm y}$ and the structure function is defined as follows

$$\begin{aligned} F^1_{UU}(x_A, x_B, |\boldsymbol{q}_T|, Q) &= x_A\, x_B\, H^{DY}(Q; \mu_R) \sum_a c_a(Q^2) \\ &\times \int d^2\boldsymbol{k}_{\perp A} d^2\boldsymbol{k}_{\perp B} f_1^a \left(x_A, \boldsymbol{k}_{\perp A}^2; \mu_R, \zeta_A\right) f_1^{\bar{a}} \left(x_B, \boldsymbol{k}_{\perp B}^2; \mu_R, \zeta_B\right) \\ &\times \delta^{(2)} \left(\boldsymbol{k}_{\perp A} + \boldsymbol{k}_{\perp B} - \boldsymbol{q}_T\right) \\ &= x_A\, x_B\, H^{DY}(Q; \mu_R) \sum_a c_a(Q^2) \\ &\times \int_0^{+\infty} d|\boldsymbol{b}_T| |\boldsymbol{b}_T| J_0(|\boldsymbol{q}_T||\boldsymbol{b}_T|) \hat{f}_1^a \left(x_A, \boldsymbol{b}_T^2; \mu_R, \zeta_A\right) \hat{f}_1^{\bar{a}} \left(x_B, \boldsymbol{b}_T^2; \mu_R, \zeta_B\right), \end{aligned} \tag{5.4}$$

where $\boldsymbol{k}_{\perp A}$ and $\boldsymbol{k}_{\perp B}$ are the transverse momenta of the active partons, J_0 is the cylindrical Bessel function of order $n = 0$, while $\hat{f}_1^a$ is the Fourier transform of the unpolarized TMD for flavour a from the transverse-momentum space $\boldsymbol{q}_T$ to the position space $\boldsymbol{b}_T$. For the Fourier transform definition we follow the convention of Ref. [189]. In Eq. (5.4), the hard factor $H^{DY}(Q; \mu_R)$ can be computed order by order in α_S and it reduces to the identity at leading order. It depends on the renormalization scale μ_R and on the hard scale Q and it encodes the virtual part of the hard scattering. The index a runs over the flavour of the active quarks with electroweak charge $c_a(Q^2)$ [182].

We notice the dependence of the unpolarized TMD on the renormalization scale μ_R and on the rapidity scale ζ. This dependence arises from the necessity

to address ultraviolet and rapidity divergences in the operation definition of the TMDs [16, 190, 191]. The TMDs obey the following evolution equations from the two initial scales μ_{Ri} and ζ_i to the two final scales μ_{Rf} and ζ_f

$$\hat{f}_1^a\left(x, \boldsymbol{b}_T^2; \mu_{Rf}, \zeta_f\right) = \hat{f}_1^a\left(x, \boldsymbol{b}_T^2; \mu_{Ri}, \zeta_i\right) \exp\left\{\int_{\mu_{Ri}}^{\mu_{Rf}} \frac{d\mu_R}{\mu_R} \gamma(\mu_R, \zeta_f)\right\} \left(\frac{\zeta_f}{\zeta_i}\right)^{K(|\boldsymbol{b}_T|, \mu_{Ri})/2}. \quad (5.5)$$

Equation (5.5) contains the anomalous dimension K of the Collins-Soper evolution in $\sqrt{\zeta}$ and the anomalous dimension $\gamma(\mu_R, \zeta)$ for the Renormalization Group evolution in μ_R. The latter takes the following form [180]

$$\gamma(\mu_R, \zeta) = \gamma_F\left(\alpha_s(\mu_R)\right) - \gamma_K\left(\alpha_s(\mu_R)\right) \ln \frac{\sqrt{\zeta}}{\mu_R}, \quad (5.6)$$

where γ_K is the cusp anomalous dimension and $\gamma_F\left(\alpha_s(\mu_R)\right) = \gamma\left(\mu_R, \mu_R^2\right)$ is the boundary condition [180]. The Collins-Soper kernel, instead, is the anomalous dimension for the evolution in the rapidity scale ζ [192].

Working in the $\overline{\text{MS}}$ renormalization scheme the following choice of the initial scales

$$\mu_{Ri} = \sqrt{\zeta_i} = \mu_b(|\boldsymbol{b}_T|) = 2\frac{e^{-\gamma_E}}{|\boldsymbol{b}_T|}, \quad (5.7)$$

where γ_E is the Euler constant, simplifies the calculations, by avoiding the insurgence of large logarithms. Nevertheless, another problem related to the value of α_S arises. In fact, in Eq. (5.6), α_S is computed at the value of the renormalization scale and μ_R in Eq. (5.7) could, in principle, become larger than Q or even hit the Landau pole for small values of $|\boldsymbol{b}_T|$, which is the kinematic region in which the TMD evolution is perturbative meaningful.

To overcome this inconvenience, we introduce an *ad hoc* prescription in which $|\boldsymbol{b}_T|$ is replaced by b_* defined as follows

$$\begin{cases} \mu_{b_*} = 2\dfrac{e^{-\gamma_E}}{b_*}, \\ b_*(|\boldsymbol{b}_T|, b_{\min}, b_{\max}) = b_{\max}\left(\dfrac{1 - e^{-|\boldsymbol{b}_T|^4/b_{\max}^4}}{1 - e^{-|\boldsymbol{b}_T|^4/b_{\min}^4}}\right)^{\frac{1}{4}}, \end{cases} \quad (5.8)$$

where $b_{\min} = 2\dfrac{e^{-\gamma_E}}{\mu_{Rf}}$ and $b_{\max} = 2e^{-\gamma_E} \simeq 1.123\ \text{GeV}^{-1}$. The net effect of this prescription is to saturate b_* at $b_{\max}$ for large $|\boldsymbol{b}_T|$, thus preventing μ_{b_*} from entering in the non-perturbative regime. On the opposite limit, at small $|\boldsymbol{b}_T|$, b_* approaches the value of $b_{\min}$, preventing μ_{b*} from becoming larger than the hard scale Q.

The b_* prescription introduces also power corrections scaling as $\left(\dfrac{\Lambda_{\text{QCD}}}{|\boldsymbol{q}_T|}\right)^k$, with $k > 0$, to be taken into account by a non-perturbative function when

$|\boldsymbol{q}_T| \sim \Lambda_{\rm QCD}$ and sub-leading power corrections scaling as $\left(\frac{|\boldsymbol{q}_T|}{Q}\right)^n$, with $n > 0$. This has the advantage of facilitating the potential matching of the TMD formula, which is valid for $|\boldsymbol{q}_T| \ll Q$, with the fixed-order calculation, which is valid for $|\boldsymbol{q}_T| \simeq Q$ [193, 194].

Accordingly, the TMD can be rewritten in the following form [180]

$$\begin{aligned}\hat{f}_1(x, \boldsymbol{b}_T^2; \mu_R, \zeta) &= \left[\frac{\hat{f}_1(x, \boldsymbol{b}_T^2; \mu_R, \zeta)}{\hat{f}_1(x, b_*(\boldsymbol{b}_T^2); \mu_R, \zeta)}\right] \hat{f}_1(x, b_*(\boldsymbol{b}_T^2); \mu_R, \zeta) \\ &\equiv f_{1NP}(x, \boldsymbol{b}_T^2; \zeta)\, \hat{f}_1(x, b_*(\boldsymbol{b}_T^2); \mu_R, \zeta)\ , \end{aligned} \tag{5.9}$$

where the second factor is perturbative calculable, while the first one effectively defines the non-perturbative part f_{1NP} of the TMD.

The perturbative function actually depends on a set of collinear parton distribution functions and matching coefficients C, that are perturbative calculable, i.e.,

$$\begin{aligned}&\hat{f}_1(x, b_*(\boldsymbol{b}_T^2); \mu_R, \zeta) \\ &\quad = [C \otimes f_1]\left(x, b_*; \mu_{b_*}, \mu_{b_*}^2\right) \exp\left\{\int_{\mu_{b_*}}^{\mu_{Rf}} \frac{d\mu_R}{\mu_R} \gamma(\mu_R, \zeta_f)\right\} \left(\frac{\zeta_f}{\mu_{b_*}^2}\right)^{K(b_*, \mu_{b_*})/2} .\end{aligned} \tag{5.10}$$

5.2.1 Nanga Parbat framework

The open-source software of Nanga Parbat[2], developed by the MAP collaboration, is a fitting framework aimed at the determination of the non-perturbative component of TMD distributions.

The software works in two main different executive steps, based on Eq. (5.9): firstly, all the perturbative physics is computed; then, the fit of the non-perturbative part f_{1NP} is performed.

In the specific case of pions, the DY process takes place between a pion and a nucleus, that, in first approximation, is described as a collection of free nucleons. Therefore, the cross section of Eq. (5.3) involves the two unpolarized TMDs $\hat{f}_{1p}^a$ and $\hat{f}_{1\pi}^a$ of a quark a in the proton and in the pion, respectively.

For the proton, we have chosen to use the recent results of the global extraction of the MAP Collaboration [182] at next-to-next-to-next-leading-logarithm ($\rm N^3LL$) accuracy, adopting the same $b_*(\boldsymbol{b}_T^2)$ prescription and the same parametrization of the non-perturbative part f_{1NP}^p.

As for the pion TMD, we opted for two distint approaches.

1. Pure phenomenological approach

A first approach is purely phenomenological, and it provides directly a model for the parametrization of the non-perturbative function of the pion f_{1NP}^π.

[2] https://github.com/MapCollaboration/NangaParbat

2. LFWA approach

A second approach consists in adapting the model developed in this work (Section (2.4.2)) to the Nanga Parbat framework. The first step is to compute the pion TMDs in terms of overlap of LFWAs. Once the TMDs for the quarks are parametrized, we need to reproduce Eq. (5.9). The non-perturbative function for the LFWA model is defined as the model TMD divided by the model PDF of Chapter (3), while in Eq. (5.10) we insert the LHAPDF grids generated by the model of the PDFs. Unfortunately, the initial scale of the model implemented in Section 2.4.2 needs to be increased to the value of 1 GeV and both the collinear fit and the form factor fit need to be performed again. This necessity arises by the choice to use the same b_* prescription of the proton TMD: if we were to retain the old value of the initial scale (0.85 GeV), it would fall below the lowest permissible value of $\mu_{b_{\max}}$, rendering the b_* prescription meaningless.

5.2.2 Data selection

The experimental data included in the analysis are the same for the two approaches we undertake.

To our knowledge, the only two existing experiments that provide observables sensitive to pion TMDs are E615 [195] and E537 [196], which involved negative pion beams colliding with tungsten targets. The kinematic coverage of the complete sets of these data in the x-Q^2 plane is shown in Fig. 5.1.

The TMD factorization theorems are valid in the kinematic regime where $|\boldsymbol{q}_T| \ll Q$. For this reason, we apply to the data sets the following cut

$$\frac{|\boldsymbol{q}_T|}{Q} < 0.3 + \frac{0.6}{Q} \,. \tag{5.11}$$

The manner in which we have applied this cut differs from what has been done in the proton TMD extraction. This difference stems from the limited data sets at our disposal. However, this adjustment represents a practical compromise between the necessity of including more data and the need to stay within the bounds of TMD factorization applicability. We also truncate those bins of E615 corresponding to an invariant mass around the Υ resonance, i.e., 9.00 GeV $< Q <$ 11.70 GeV.

In Tab. 5.2, we present the data set statistics, distinguishing between the number of data before (N_{dat}) and after (N_{surv}) applying the kinematic cut, divided by experiment. We also include there information regarding the definition of the observable, the center-of-mass energy and the x_F range.

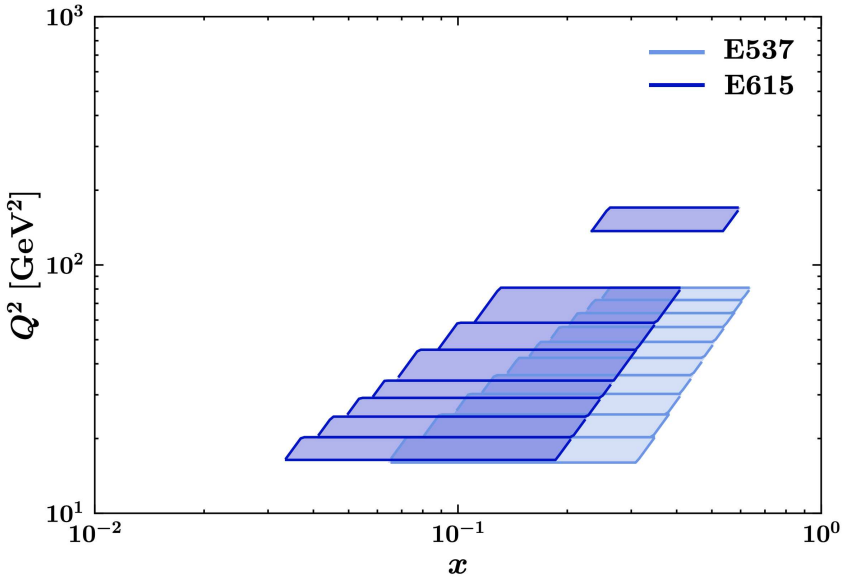


Figure 5.1: The x-Q^2 coverage of E615 [195] (blue) and E537 [196] (light blue) data.

Experiment	N_{dat}	N_{surv}	Observable	$\sqrt{s}$ [GeV]	x_F range
E615	155	74	$\frac{d^2\sigma}{dQd\lvert\mathbf{q}_T\rvert}$	21.8	0.0 - 1.0
E537	150	64	$\frac{d^2\sigma}{dQd\lvert\mathbf{q}_T\rvert^2}$	15.3	0.0 - 1.0

Table 5.2: Useful kinematic information of the data sets: number of data before (N_{dat}) and after (N_{surv}) the cut, observable definition, center of mass energy and x_F range.

The statistical uncertainties associated to the experimental data are quite large for both E615 (16%) and E537 (8%) and we treat them as uncorrelated. On the contrary, we treat the systematic uncertainties as fully correlated.

We notice that for small values of $\boldsymbol{b}_T$, i.e. $|\boldsymbol{b}_T| \ll \frac{1}{\Lambda_{\text{QCD}}}$, the TMDs match with the related collinear PDFs and the errors on the collinear distributions must be taken into account and propagated to the TMDs.

The collinear set that we choose to associate to proton TMDs are the unpolarized PDFs from MMHT2014 [197], consistently with Ref. [182]. For the pion the choice depends on the approach: for the pure phenomenological method we use the xFitter20 set [53], while for the LFWA model we insert the PDFs of Chapter 3 evolved from the initial scale $\mu_0 = 1$ GeV.

For the propagation of the uncertainties of the proton PDFs to the proton TMDs as well as of the pion PDFs to the proton TMDs within the pure phenomenological approach, we use the Hessian method [198–200], by consid-

ering the 80% of the uncertainties as fully correlated and the remaining 60% as uncorrelated, as done in Ref. [182]. Instead, for the propagation of the uncertainties of the 1000 PDF replicas of the pion PDFs obtained within the LFWA approach we exploit the same approach used for the error propagation to the FFs and described in Section 4.4.2.

For the renormalization and rapidity scales we choose $\mu_R = \sqrt{\zeta} = Q$, for both the phenomenological and the LFWA approaches.

5.3 Pure phenomenological fit

In the pure phenomenological approach, the parametrization for the non-perturbative function of the pion is the following

$$f_{1NP}^{\pi}(x, \boldsymbol{b}_T^2; \zeta) = \exp\left\{-g_{1\pi}(x)\frac{\boldsymbol{b}_T^2}{4}\right\}\left[\frac{\zeta}{Q_0}\right]^{g_K(\boldsymbol{b}_T^2)/2}, \tag{5.12}$$

where $g_K(\boldsymbol{b}_T^2)$ is the non-perturbative part of the Collins-Soper kernel

$$\begin{cases} K(|\boldsymbol{b}_T|, \mu_{b_*}) = K(b_*, \mu_{b*}) + g_K(|\boldsymbol{b}_T|) \\ g_K(\boldsymbol{b}_T^2) = -g_2\dfrac{\boldsymbol{b}_T^2}{2} \end{cases}, \tag{5.13}$$

where the parameter g_2 is fixed to the fitted value in Ref. [182]. Q_0 is the initial scale of 1 GeV, while the function $g_{1\pi}(x)$, which introduces the dependence on the fraction of collinear momentum x, is defined as

$$g_{1\pi}(x) = N_{1\pi}\frac{x^{\sigma_\pi}(1-x)^{\alpha_\pi^2}}{\tilde{x}^{\sigma_\pi}(1-\tilde{x})^{\alpha_\pi^2}}, \tag{5.14}$$

with $\tilde{x} = 0.1$. The total number of free fitting parameters is 3: $N_{1\pi}, \sigma_\pi$ and α_π. Various tests with alternative functional forms and varying the number of free parameters have been performed. However, the best-fit result has been obtained with the analytic form in Eq. (5.14).

5.3.1 Fit Results

As of now, the collinear sets of pion PDFs, such as xFitter20, are available at next-to-leading order (NLO) accuracy. All the other ingredients for the extraction of pion TMDs are at N^{3}LL accuracy. Consequently, the final result will be categorized as what has been defined as "N^{3}LL$^-$" accuracy (similarly to Ref. [182]).

The error analysis of the non-perturbative function has been studied with the bootstrap method by generating 200 Monte Carlo replicas of experimental data. Given that the number of replicas for the extraction of proton quark TMDs in Ref. [182] is 250, we performed the fit of the non-perturbative pion

TMD by keeping the correspondence between the i-th replica in the pion with the same number of replica in the proton. From the final fit results we cut the outliers replicas, i.e., those replicas whose χ^2 falls in the highest or lowest 5% percentile.

The chi-squared definition is analogous to Eq. (3.30) and we can decompose the χ^2_i of of the i-th replica as

$$\chi^2_i = \chi^2_{i\,\mathrm{unc}} + \chi^2_{i\,\mathrm{pen}}, \tag{5.15}$$

where $\chi^2_{i\,\mathrm{unc}}$ is the contribution of uncorrelated uncertainties and $\chi^2_{i\,\mathrm{pen}}$ is the penalty term corresponding to the correlated uncertainties.

The replica 0 refers to the single minimization of the fit by using the original (non-fluctuated) data. The results of the best fit of the replica 0 are reported in Tab. 5.3 for the two experimental data sets included in the analysis.

Experiments	N_{surv}	$\chi^2_{0\,\mathrm{unc}}/N_{\mathrm{surv}}$	$\chi^2_{0\,\mathrm{pen}}/N_{\mathrm{surv}}$	$\chi^2_0/N_{\mathrm{surv}}$
E537	58	1.00	0.58	1.58
E615	74	0.31	1.22	1.53
Total	132	0.63	0.92	1.55

Table 5.3: Values for the total (χ^2_0), uncorrelated ($\chi^2_{0\,\mathrm{unc}}$) and correlated ($\chi^2_{0\,\mathrm{pen}}$) chi-squared, divided by the number of points surviving the kinematic cuts (N_{surv}).

The high contribution of the global penalty term ($\chi^2_{0\,\mathrm{pen}}/N_{\mathrm{surv}} = 0.92$) in Tab. 5.3 suggests that the comparison between data and theory is strongly affected by normalization errors. The small values of the partial uncorrelated chi-squared are indications of very good agreement in the shape between the theory and the experimental data, which is almost perfect in the data set of E615. However, there is a systematic disagreement between data and theory in the normalization which is responsible of the increase in the penalty terms.

We argue that the origin of the large values of the penalty resides in the correlated systematic uncertainties of the experimental data ($\sim 16\%$), and not in the relatively small uncertainties associated with the collinear PDFs, that both for the proton and for the pion are not larger than 5-8%. This conclusion is compatible with the findings of Ref. [185], which suggested that a potential explanation for the same issue observed in the E615 data could be attributed to an incorrect normalization of the experimental data.

The best fit parameters of the replicas are listed below, together with their standard deviations:

$$\begin{aligned} \langle N_{1\pi}\rangle &= 0.47, \quad \langle \sigma_\pi\rangle = 4.50, \quad \langle \alpha_{1\pi}\rangle = 4.40, \\ \sigma_{N_{1\pi}} &= 0.12, \quad \sigma_{\sigma_\pi} = 2.25, \quad \sigma_{\alpha_{1\pi}} = 1.34. \end{aligned} \tag{5.16}$$

The best fit parameters are not very well constrained, as we can see from the high values of the uncertainties. This issue does not change by varying the

functional form of the non-perturbative function and the number of parameters, as previously anticipated. Moreover, in agreement with Ref. [185], we also find strong correlations among the three parameters, as it is clear from Fig. 5.2. We conclude that the current set of experimental data is not very sensitive to these degrees of freedom and more data are needed to better constrain them. The plots of the results for the observables and the TMDs are shown

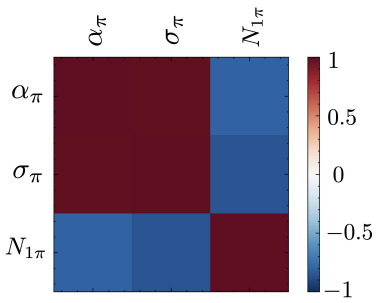


Figure 5.2: Correlation matrix of the three fitted parameters in the pure phenomenological approach.

and commented in Section 5.5, together with the results from the LFWAs approach.

5.4 LFWA framework

In this Section we exploit the formalism developed in Chapter 2 to recover the representation of the unpolarized pion TMDs in terms of overlaps of LFWAs. This allows us to use the model of Section 2.4.2 to obtain a new parametrization for $f_1(x, \boldsymbol{k}_\perp)$.

5.4.1 Correlator for the TMDs

The TMDs parametrize the quark-quark correlator (2.123) in the limit of $\Delta = 0$. At leading-twist with $\Gamma = \gamma^+$, we obtain the definition of the quark TMD f_1 as follows

$$f_1^q(x, \boldsymbol{k}_\perp) = \int \frac{d\zeta^- d^2\boldsymbol{\zeta}_\perp}{2(2\pi)^3} e^{i(k^+\zeta^- - \boldsymbol{\zeta}_\perp \cdot \mathbf{k}_\perp)} \langle \pi(p)| \, \bar{\Psi}^q(0) \mathcal{U}_{(0,\zeta)} \gamma^+ \Psi^q(\zeta) \, |\pi(p)\rangle \Big|_{\zeta^+=0} . \tag{5.17}$$

As explained in Section 2.2, the contribution of the gauge link in Eq. (5.17) simplifies to the identity, when it is incorporated within a complex phase of the LFWAs. This follows from the expression of the unpolarized T-even TMD which involves the squared module of the LFWAs (see Section 5.4.2).
Thanks to the presence of the Fock states with one and two gluons in the model developed in Chapter. 2 (see Eqs. (2.135) - (2.136)), we can also define

the gluon TMD, which takes the following form

$$f_1^g(x,\boldsymbol{k}_\perp) = \frac{1}{xp^+}\int \frac{d\zeta^- d^2\boldsymbol{\zeta}_\perp}{(2\pi)^3} e^{i\left(k^+\zeta^- - \boldsymbol{\zeta}_\perp\cdot\mathbf{k}_\perp\right)} \langle \pi(p)|G^{+i}{}_a(0)G_{+i\,a}(\zeta)|\pi(p)\rangle\Big|_{\zeta^+=0}. \tag{5.18}$$

Since the experimental data for pion TMDs are obtained from negative pions, in this Chapter we refer to the distributions of the quarks in a π^-, differently from what we did for the other distribution functions. However, the charge symmetry equations (3.3) for the PDFs also hold for the TMDs, and it is straightforward to recover the same distributions for the π^+ or π^0.

5.4.2 LFWA overlap representation for pion TMDs

Following the approach taken for the PDFs and the FFs, we proceed by inserting the Fock-space expansion of the pion state of Eqs. (2.134) - (2.137) in the definition of the TMD correlators (5.17) - (5.18). The result is the LFWA overlap representation of the TMDs.

Analogously to the case of the pion PDFs, the $\Delta = 0$ condition leads to expressions that involve only the squares of the LFWAs.

Separating the contributions of each distinct Fock component we obtain the following results for the gluon TMDs

$$f_1^g\,(x,\boldsymbol{k}_\perp) = f^g_{1,d\bar{u}g}\,(x,\boldsymbol{k}_\perp) + f^g_{1,d\bar{u}gg}\,(x,\boldsymbol{k}_\perp) \tag{5.19}$$

where

$$f^g_{1,d\bar{u}g}(x,\boldsymbol{k}_\perp) = 2\int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3]\sqrt{x_1x_2x_3}\,\delta(x-x_3)\delta^{(2)}(\boldsymbol{k}_\perp-\boldsymbol{k}_{\perp 3})|\psi^{(1)}_{d\bar{u}g}(1,2,3)|^2, \tag{5.20}$$

$$\begin{aligned} f^g_{1,d\bar{u}gg}(x,\boldsymbol{k}_\perp) = 16\int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3]\mathrm{d}[4]\sqrt{x_1x_2x_3x_4}\,\delta(x-x_3)\delta^{(2)}(\boldsymbol{k}_\perp-\boldsymbol{k}_{\perp 3}) \\ \times\left[|\psi^{(1)}_{d\bar{u}gg}(1,2,3,4)|^2 + |\psi^{(2)}_{d\bar{u}gg}(1,2,3,4)|^2\right]. \end{aligned} \tag{5.21}$$

It is evident that with an integral over $\boldsymbol{k}_\perp$ of Eq. (5.19) we recover Eq. (3.13).

We also present the LFWA overlap representation for the quark (and antiquark) TMDs at the initial scale:

$$\begin{aligned} f_1^u(x,\boldsymbol{k}_\perp) &= f_1^{\bar{d}}(x,\boldsymbol{k}_\perp) = f_1^s(x,\boldsymbol{k}_\perp) = f_1^{\bar{s}}(x,\boldsymbol{k}_\perp) \\ &= \frac{2}{3}\int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3]\mathrm{d}[4]\sqrt{x_1x_2x_3x_4}\,\delta(x-x_3)\delta^{(2)}\left(\boldsymbol{k}_\perp-\boldsymbol{k}_{\perp 3}\right) \\ &\quad\times\left[|\psi^{(1)}_{d\bar{u}s\bar{s}}(1,2,3,4)|^2 + |\psi^{(2)}_{d\bar{u}s\bar{s}}(1,2,3,4)|^2 + \frac{1}{2}|\psi^{(3)}_{d\bar{u}s\bar{s}}(1,2,3,4)|^2\right], \end{aligned} \tag{5.22}$$

$$
\begin{aligned}
f_1^d(x,\boldsymbol{k}_\perp) &= f_1^{\bar{u}}(x,\boldsymbol{k}_\perp) \\
&= 2\int \mathrm{d}[1]\mathrm{d}[2]\sqrt{x_1x_2}\delta(x-x_1)\,(\boldsymbol{k}_\perp-\boldsymbol{k}_{\perp 1})\,|\psi^{(1)}_{d\bar{u}}(1,2)|^2 \\
&\quad + 2\int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3]\sqrt{x_1x_2x_3}\delta(x-x_1)\,(\boldsymbol{k}_\perp-\boldsymbol{k}_{\perp 1})\,|\psi^{(1)}_{d\bar{u}g}(1,2,3)|^2 \\
&\quad + 8\int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3]\mathrm{d}[4]\sqrt{x_1x_2x_3x_4}\delta(x-x_1)\,(\boldsymbol{k}_\perp-\boldsymbol{k}_{\perp 1}) \\
&\qquad \times\left[|\psi^{(1)}_{d\bar{u}gg}(1,2,3,4)|^2+|\psi^{(2)}_{d\bar{u}gg}(1,2,3,4)|^2\right], \\
&\quad + 4\int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3]\mathrm{d}[4]\sqrt{x_1x_2x_3x_4}\delta(x-x_1)\delta^{(2)}\,(\boldsymbol{k}_\perp-\boldsymbol{k}_{\perp 1}) \\
&\qquad \times\left[|\psi^{(1)}_{d\bar{u}s\bar{s}}(1,2,3,4)|^2+|\psi^{(2)}_{d\bar{u}s\bar{s}}(1,2,3,4)|^2+\frac{1}{2}|\psi^{(3)}_{d\bar{u}s\bar{s}}(1,2,3,4)|^2\right] \\
&\quad + \frac{2}{3}\int \mathrm{d}[1]\mathrm{d}[2]\mathrm{d}[3]\mathrm{d}[4]\sqrt{x_1x_2x_3x_4}\delta(x-x_3)\delta^{(2)}\,(\boldsymbol{k}_\perp-\boldsymbol{k}_{\perp 3}) \\
&\qquad \times\left[|\psi^{(1)}_{d\bar{u}s\bar{s}}(1,2,3,4)|^2+|\psi^{(2)}_{d\bar{u}s\bar{s}}(1,2,3,4)|^2+\frac{1}{2}|\psi^{(3)}_{d\bar{u}s\bar{s}}(1,2,3,4)|^2\right].
\end{aligned}
\tag{5.23}
$$

When integrated over $\boldsymbol{k}_\perp$, these TMDs yield the quark (and antiquark) PDFs that, if combined appropriately, reproduce the valence and total Sea PDFs in Eqs. (3.11) and (3.12), respectively.

5.4.3 Explicit model for pion TMDs

By inserting the model developed in Section 2.4.2 in Eqs. (5.22) - (5.23), we find the explicit expressions for the quark and gluon unpolarized TMDs depending on x, $\boldsymbol{k}_\perp$ and all the model parameters.

However, the Nanga Parbat software requires the parametrization in the space of $\boldsymbol{b}_T$, according to Eq. (5.9). Therefore, the results must be Fourier-transformed in the $\boldsymbol{b}_T$ space. In the following, we report the final results for the parametrizations of the pion TMDs at the initial scale $Q_0 = 1$ GeV after performing the Fourier transforms:

$$
\begin{aligned}
&\hat{f}_1^d\left(x,\boldsymbol{b}_T^2;Q_0,Q_0^2\right) \equiv \hat{f}_1^{\bar{u}}\left(x,\boldsymbol{b}_T^2;Q_0,Q_0^2\right) \\
&\quad = \frac{C^v_{u\bar{d}}}{2}(x\bar{x})^{2\gamma_q-1}\left[1+d_{q1}(1+2\gamma_q)\left(1+\gamma_q(x-\bar{x})^2-6x\bar{x}\right)\right]^2\exp\left\{-\frac{\boldsymbol{b}_T^2\,x\bar{x}}{8\,a^2_{q\bar{q}'}}\right\} \\
&\qquad + \frac{C^v_{u\bar{d}g}}{2}x\bar{x}^5\left[3+18xd_{g1}-10\bar{x}d_{g1}+13d_{g1}^2+14xd_{g1}^2(x-4\bar{x})\right] \\
&\qquad\quad \times\exp\left\{-\frac{\boldsymbol{b}_T^2\,x\bar{x}}{8\,a^2_{q\bar{q}'g}}\right\}
\end{aligned}
$$

$$
\begin{aligned}
&+ x\bar{x}^9 \left[\frac{N^{(1)2}_{q\bar{q}'gg}}{1260} \exp\left\{ -\frac{\boldsymbol{b}_T^2\, x\bar{x}}{8\, a^{(1)2}_{q\bar{q}'gg}} \right\} + \frac{N^{(2)2}_{q\bar{q}'gg}}{540} \exp\left\{ -\frac{\boldsymbol{b}_T^2\, x\bar{x}}{8\, a^{(2)2}_{q\bar{q}'gg}} \right\} \right] \\
&+ x\bar{x}^5 \left[\frac{2\, N^{(1)2}_{q\bar{q}'s\bar{s}}}{15} \exp\left\{ -\frac{\boldsymbol{b}_T^2\, x\bar{x}}{8\, a^{(1)2}_{q\bar{q}'s\bar{s}}} \right\} + \frac{2\, N^{(2)2}_{q\bar{q}'s\bar{s}}}{15} \exp\left\{ -\frac{\boldsymbol{b}_T^2\, x\bar{x}}{8\, a^{(2)2}_{q\bar{q}'s\bar{s}}} \right\} \right. \\
&\qquad \left. + \frac{N^{(3)2}_{q\bar{q}'s\bar{s}}}{15} \exp\left\{ -\frac{\boldsymbol{b}_T^2\, x\bar{x}}{8\, a^{(3)2}_{q\bar{q}'s\bar{s}}} \right\} \right],
\end{aligned}
\tag{5.24}
$$

$$
\begin{aligned}
\hat{f}_1^g\left(x, \boldsymbol{b}_T^2; Q_0, Q_0^2\right) &= \mathcal{C}^g_{q\bar{q}'g} (x\bar{x})^3 \left[1 + d_{g1}(3 - 7x)\right]^2 \exp\left\{ -\frac{\boldsymbol{b}_T^2\, x\bar{x}}{8\, a^{2}_{q\bar{q}'g}} \right\} \\
&\quad + \mathcal{C}^{g(1)}_{q\bar{q}'gg}\, x^3\bar{x}^7 \exp\left\{ -\frac{\boldsymbol{b}_T^2\, x\bar{x}}{8\, a^{(1)2}_{q\bar{q}'gg}} \right\} \\
&\quad + \frac{N^{(2)2}_{q\bar{q}'gg}}{1890} x\bar{x}^5 (5 - 20x - 6x^2 + 52x^3 + 95x^4) \\
&\qquad \times \exp\left\{ -\frac{\boldsymbol{b}_T^2\, x\bar{x}}{8\, a^{(1)2}_{q\bar{q}'gg}} \right\},
\end{aligned}
\tag{5.25}
$$

$$
\begin{aligned}
&\hat{f}_1^u\left(x, \boldsymbol{b}_T^2; Q_0, Q_0^2\right) \equiv \hat{f}_1^{\bar{d}/s/\bar{s}}\left(x, \boldsymbol{b}_T^2; Q_0, Q_0^2\right) \\
&\quad = \frac{14}{3} \sin^2(\alpha_1) \sin^2(\alpha_2) \sin^2(\alpha_3) x\bar{x}^5 \\
&\qquad \times \left[\exp\left\{ -\frac{\boldsymbol{b}_T^2\, x\bar{x}}{8\, a^{(1)2}_{q\bar{q}'\delta\bar{\delta}}} \right\} + \exp\left\{ -\frac{\boldsymbol{b}_T^2\, x\bar{x}}{8\, a^{(2)2}_{q\bar{q}'\delta\bar{\delta}}} \right\} + \exp\left\{ -\frac{\boldsymbol{b}_T^2\, x\bar{x}}{8\, a^{(3)2}_{q\bar{q}'\delta\bar{\delta}}} \right\} \right].
\end{aligned}
\tag{5.26}
$$

To obtain the fitting function f_{1NP} utilized in the minimization routine of the software, we need to divide the previous expressions for the TMDs by the analytic expressions of the PDFs.

Although the gluon TMD is not included in the current fitting framework of Nanga Parbat, we presented its expression in Eq. (5.25) for sake of completeness.

We notice that Eqs. (5.24) - (5.26) depend on all the parameters of the model implemented in Section 2.4.2. In particular, they involve all the transverse parameters of the set $\mathcal{A}^T$. In other words, Eqs. (4.16) - (4.17) have not been applied to maintain a more general expression for the TMDs.

5.4.4 Predictions of TMDs from FF parameters

Raising the initial scale from 0.85 GeV to 1 GeV leads to a revised set of best fit collinear and transverse parameters. These parameters have been determined under identical conditions to the fits of PDFs and FFs described in the preceding Chapters, except for the change in the initial scale.

The mean values and the standard deviations of the parameters from the fit with the new initial scale are listed below:

$$\begin{aligned} \langle \gamma_q \rangle &= 0.670, \quad \langle d_{q1} \rangle = -0.139, \quad \langle d_{g1} \rangle = 21.188, \\ \langle \alpha_1 \rangle &= 0.856, \quad \langle \alpha_2 \rangle = 1.787, \quad \langle \alpha_3 \rangle = 2.608, \end{aligned} \tag{5.27}$$

$$\begin{aligned} \sigma_{\gamma_q} &= 0.104, \quad \sigma_{d_{q1}} = 0.048, \quad \sigma_{d_{g1}} = 117, \\ \sigma_{\alpha_1} &= 0.022, \quad \sigma_{\alpha_2} = 0.175, \quad \sigma_{\alpha_3} = 0.096. \end{aligned} \tag{5.28}$$

$$\begin{aligned} \langle a_{q\bar{q}'} \rangle &= 12.307, \quad \langle a_{q\bar{q}'gg} \rangle = 0.832, \quad \langle a_{q\bar{q}'s\bar{s}} \rangle = 0.995 \\ \sigma_{a_{q\bar{q}'}} &= 11.02, \quad \sigma_{a_{q\bar{q}'gg}} = 1.467, \quad \sigma_{a_{q\bar{q}'s\bar{s}}} = 0.009. \end{aligned} \tag{5.29}$$

In order to further test the formalism of LFWAs, we used the values of the parameters of Eqs. (5.27) - (5.29) to predict the parametrized non-perturbative functions and then the DY observable in Eq. (5.3). We adopted the same kinematics of the available data set, and then compared our theoretical predictions with the experimental points by computing the chi-squared.

Unfortunately, the results are quite poor: using the parameters in Eqs. (5.27) - (5.29), we obtain a reduced χ^2_{pred} of 3.95 and with the parameters of the replica 0 we achieve $\chi^2_{\text{pred},0} = 2.99$. This outcome is not surprising, given that the form factors are associated with a completely different kinematics with respect to the TMDs, and it is not guaranteed that the parameters optimized for the FFs will also effectively reproduce the TMDs.

5.4.5 Fit of transverse parameters

The challenge to find a comprehensive set of model parameters that well describe in a unified matter the pion PDFs, the TMDs and the FFs, may be addressed through a global extraction strategy. Such an approach involves a simultaneous fitting of observables sensitive to distinct distribution functions and the data for the FFs. This is a challenging task, but we expect significant progress in this direction from various phenomenological groups in the near future.

At present, such a global extraction is not feasible due to the scarcity and poor quality of the available data. Therefore, we have chosen to proceed in exactly the same way as we did for the form factors: we keep fixed the set $\mathcal{A}^L$ during the TMD fit, maintaining the values of the 1000 replicas of the PDF fit, and we perform the fit to the TMDs with the transverse set $\mathcal{A}^T$ alone.

By doing so, the fit will provide us with a new set of best fit transverse parameters, that can be compared with the corresponding set obtained from the fit of FFs. From now on, we use two different notations to refer to these two transverse sets: $\mathcal{A}^T_{FF}$, associated to the fit of FFs, and $\mathcal{A}^T_{TMD}$, referring to the fit of TMDs.

A first interesting question is to understand if the fit of the TMDs is sensitive to all the parameters of the set $\mathcal{A}^T$, or there are some parameters that can be fixed to arbitrary values, like $a_{q\bar{q}'g}$ in $\mathcal{A}^T_{FF}$, and others that can be fixed equal to each other, as we assumed in Eq. (4.17).

The general expression of the model in Eqs. (5.24) - (5.26), containing all the parameters of the transverse set, helps to answer these questions. From some preliminary tests, we noticed that the assumption in Eq. (4.17) holds also for the TMDs, i.e., the fit with the current experimental data is not able to distinguish between $a^{(1)}_{q\bar{q}'gg}$ - $a^{(2)}_{q\bar{q}'gg}$, as well as between $a^{(1)}_{q\bar{q}'s\bar{s}}$ - $a^{(2)}_{q\bar{q}'s\bar{s}}$ - $a^{(3)}_{q\bar{q}'s\bar{s}}$. For what concerns the remaining transverse parameters, we found that the TMD data set is sensitive to all of them, even if some parameters are not well constrained in the end, due to the limited amount of available data.

Thanks to these considerations, the number of fitted parameters is just 4: $a_{q\bar{q}'}, a_{q\bar{q}'g}, a_{q\bar{q}'gg}, a_{q\bar{q}'s\bar{s}}$ (the same as for the FFs, but with the addition of $a_{q\bar{q}'g}$).

Before presenting the results of the fit, we show the explicit expressions for the non-perturbative fitting functions f^q_{1NP}, that assume very simple forms once the degeneracy condition (4.17) is applied:

$$\begin{aligned} f^d_{1NP}\left(x, \boldsymbol{b}_T^2; Q_0^2\right) &\equiv f^{\bar{u}}_{1NP}\left(x, \boldsymbol{b}_T^2; Q_0^2\right) \\ &= \frac{1}{2}\left[\exp\left\{-\frac{\boldsymbol{b}_T^2\, x\bar{x}}{8\, a^2_{q\bar{q}'}}\right\} + \exp\left\{-\frac{\boldsymbol{b}_T^2\, x\bar{x}}{8\, a^2_{q\bar{q}'g}}\right\} + \exp\left\{-\frac{\boldsymbol{b}_T^2\, x\bar{x}}{8\, a^2_{q\bar{q}'gg}}\right\}\right] \\ &\quad + \frac{2}{3}\exp\left\{-\frac{\boldsymbol{b}_T^2\, x\bar{x}}{8\, a^{(2)2}_{q\bar{q}'s\bar{s}}}\right\}, \end{aligned} \tag{5.30}$$

$$f^u_{1NP}\left(x, \boldsymbol{b}_T^2; Q_0^2\right) \equiv f^{\bar{d}/s/\bar{s}}_{1NP}\left(x, \boldsymbol{b}_T^2; Q_0^2\right) = \frac{1}{6}\exp\left\{-\frac{\boldsymbol{b}_T^2\, x\bar{x}}{8\, a^{(2)2}_{q\bar{q}'s\bar{s}}}\right\}. \tag{5.31}$$

The error analysis has been carried out with the bootstrap method, with the generation of 150 Monte Carlo replicas for the non-perturbative functions (5.30) - (5.31). Each TMD replica was associated with a corresponding PDF replica. Similar to the pure phenomenological approach, we cut the outliers corresponding to the highest or lowest 5% percentile in χ^2. The number of final replicas is, thus, 117.

We report in the following the equivalent of Tab. 5.3, with the results for the fit in the LFWA approach. The considerations valid for the pure phenomenological fit hold also for the fit based on LFWAs. In particular, from Tab. 5.4 we observe that the largest contribution to the chi-squared from single minimization for the E615 data is due to the penalty term, indicating the presence of substantial normalization errors. Moreover, analogously to the other fit, the sets of E537 have a poorer fit compared to those from E615.

The best fit parameters of $\mathcal{A}^T_{TMD}$ with their standard deviations are as follows

$$\langle a_{q\bar{q}'}\rangle = 0.91, \quad \langle a_{q\bar{q}'g}\rangle = 0.52, \quad \langle a_{q\bar{q}'gg}\rangle = 0.31, \quad \langle a_{q\bar{q}'s\bar{s}}\rangle = 0.42,$$

Experiments	N_{surv}	$\chi^2_{0\,\text{unc}}/N_{\text{surv}}$	$\chi^2_{0\,\text{pen}}/N_{\text{surv}}$	χ^2_0/N_{surv}
E537	58	1.27	1.00	2.27
E615	74	0.29	1.42	1.71
Total	132	0.72	1.23	1.95

Table 5.4: Values for the total, uncorrelated and correlated chi-squared of the replica 0 divided by the number of points surviving the kinematic cuts for the parametrization in the LFWA approach.

$$\sigma_{a_{q\bar{q}'}} = 0.51, \qquad \sigma_{a_{q\bar{q}'}} = 0.16, \qquad \sigma_{a_{q\bar{q}'gg}} = 0.02, \qquad \sigma_{a_{q\bar{q}'s\bar{s}}} = 0.03, \tag{5.32}$$

with the correlation matrix of the parameters given in Fig. 5.3. The correla-

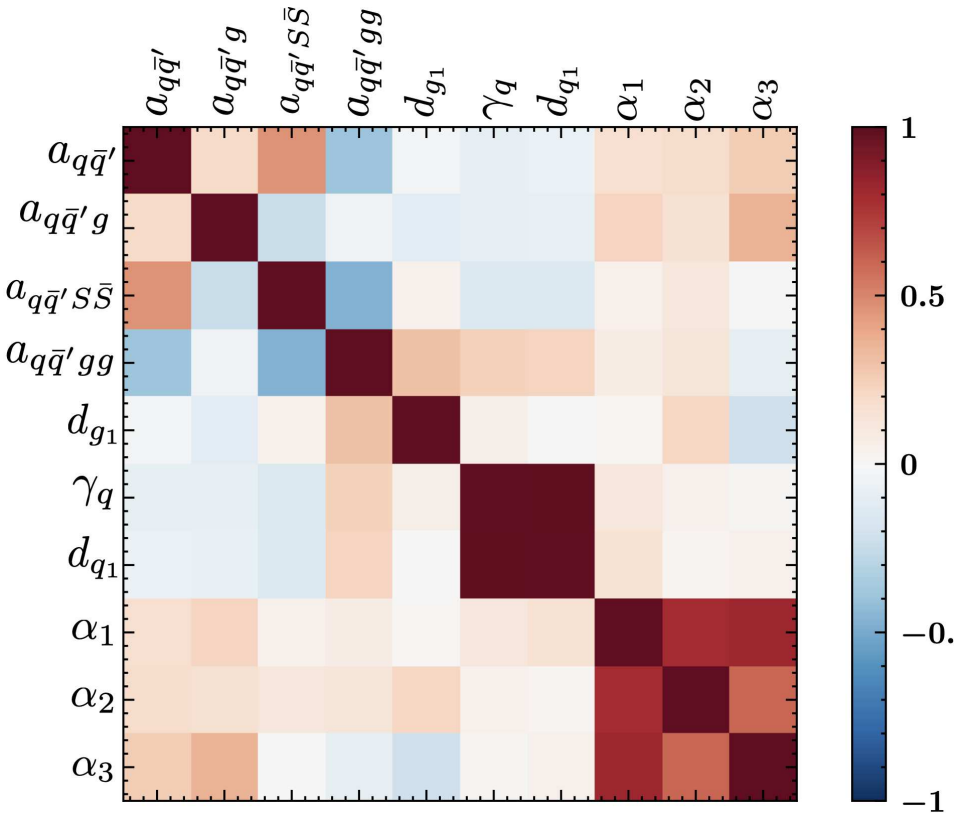


Figure 5.3: Correlation matrix of all the model parameters in the approach based on LFWAs.

tion matrix 5.3 contains all the parameters of the model, both $\mathcal{A}^L$ and $\mathcal{A}^T_{TMD}$. We can compare it with the same matrix built for the form factors fit 4.3 and notice, in general, a similar pattern: strong correlations among some collinear parameters and weaker correlations - in module - among the transverse parameters. Nevertheless, we remind that the two matrices pertain to two fits differing in both the number of replicas included and in the initial model scale. Therefore, this comparative analysis of correlation matrices remain at a qualitative level.

The results of the fit for the observables and the extracted TMDs are reported and analyzed in Section 5.5.

5.4.6 Comparison of pion DAs

We now want to perform another rigorous consistency check to evaluate the effectiveness of the approach based on the LFWA formalism.

Equation (4.23) provides the analytic expression for the pion DA at twist-2. Our goal is now to compute and plot the DA by inserting the two sets of best fit parameters coming from:

1. the fit of pion FFs with 10000 replicas for $\mathcal{A}^T_{FF}$ and the longitudinal parameters $\mathcal{A}^L$ from the fit of pion PDFs evolved from the initial scale of 1 GeV;

2. the fit of pion TMDs with 117 replicas for $\mathcal{A}^T_{TMD}$ and the longitudinal parameters $\mathcal{A}^L$ from the fit of pion PDFs, evolved from the initial scale of 1 GeV.

The results are shown in Fig 5.4, where we plotted all the single replicas. The

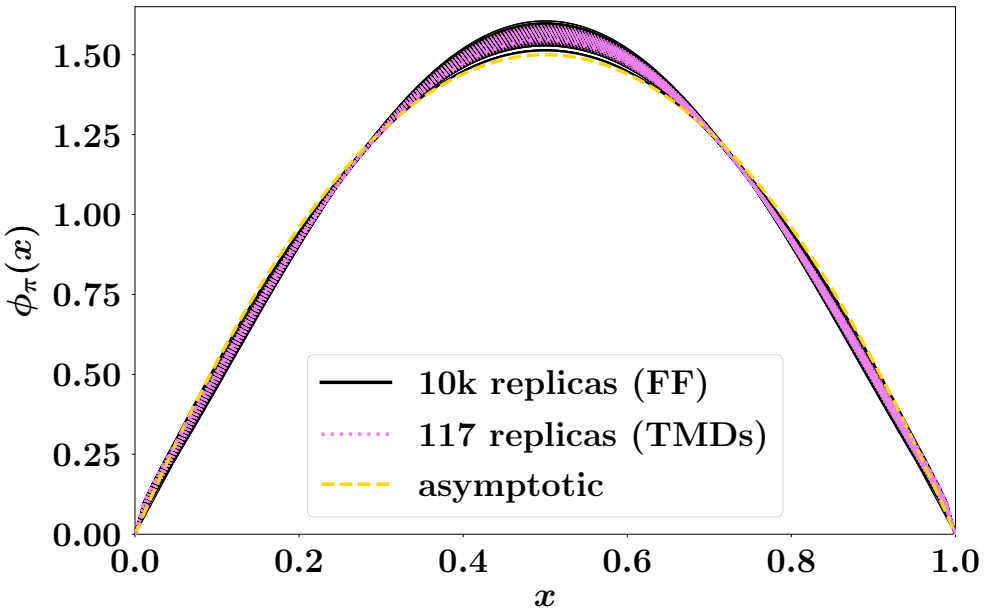


Figure 5.4: Comparison between the DAs obtained with the parameters fitted to the FFs (black band) and to the TMDs (dashed pink band).

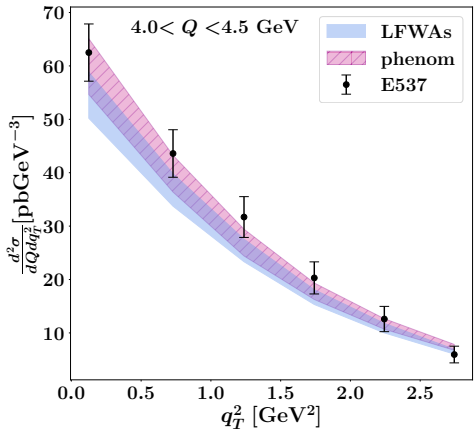
consistency between the predictions from the two sets of parameters provides a valuable test for validating the LFWA approach and emphasizes the versatility and robustness of our model in describing different aspects of the pion internal structure.

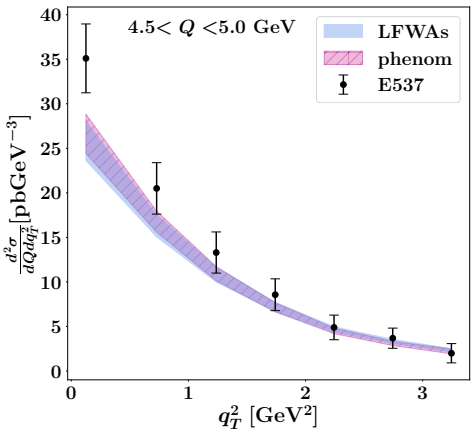
5.5 Comparison between the two TMD fits

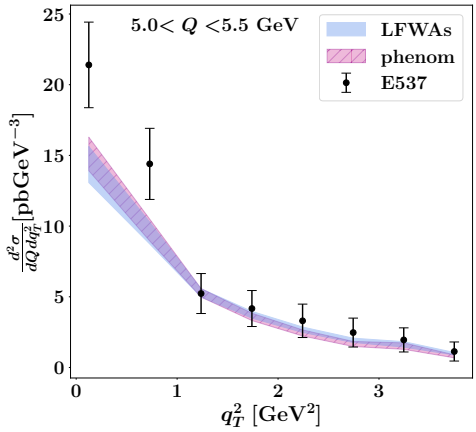
This Section is devoted to the comparison between the two fits of pion TMDs. In the following figures we label with "phenom" the results corresponding to the pure phenomenological approach of Section 5.3 and with "LFWAs" the ones relative to the parametrization outlined in Section 5.4.

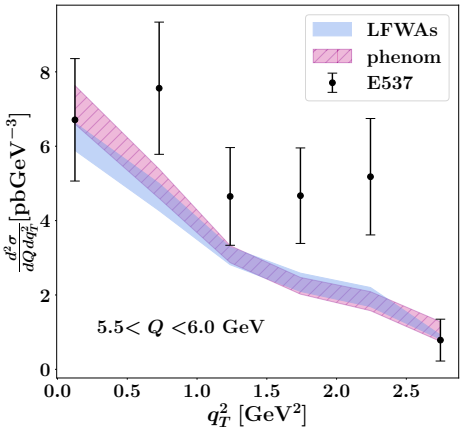
5.5.1 Comparison of DY Observables

E537









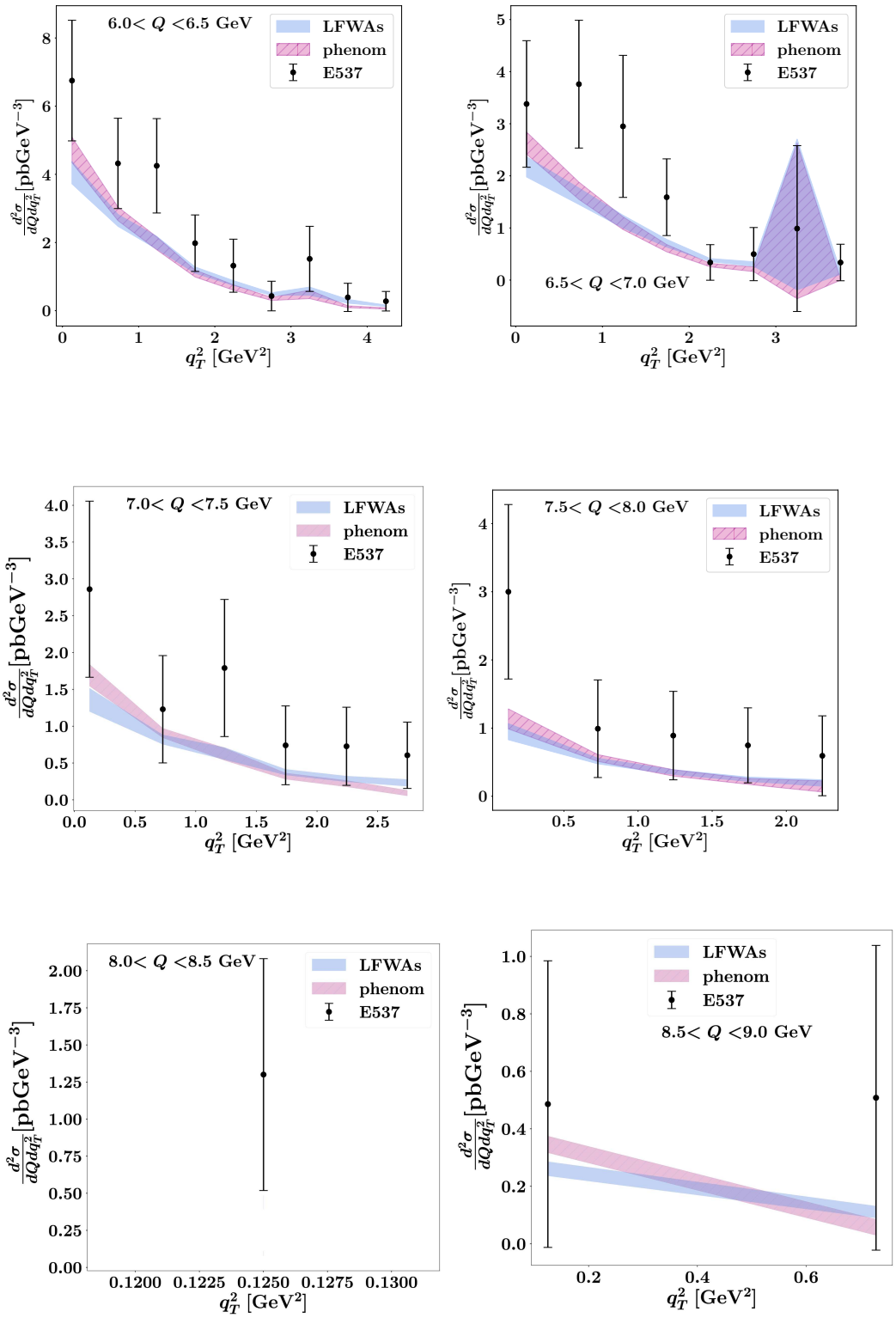
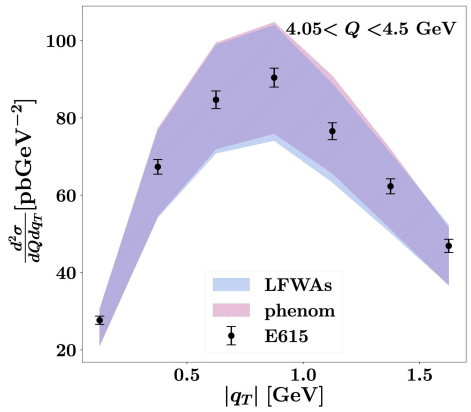
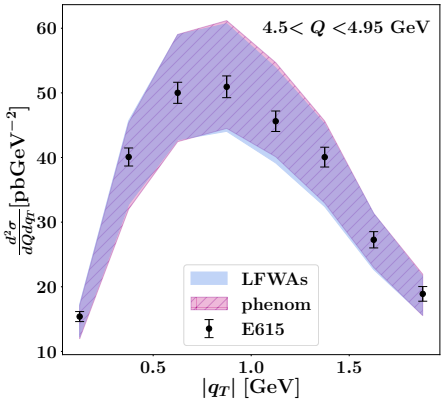
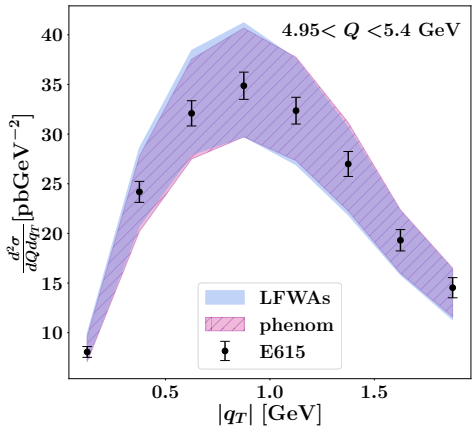


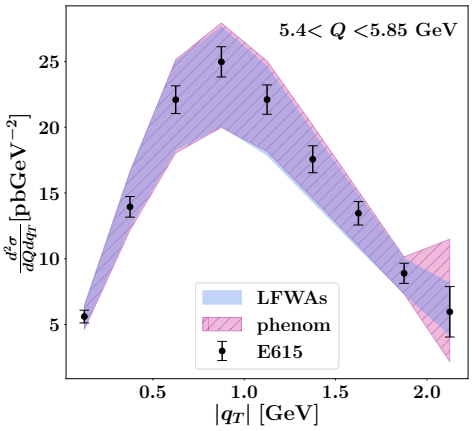
Figure 5.5: Comparison between the experimental DY data of E537 (black points), the pure phenomenological approach (pink bands) and the model based on LFWAs (light blue bands) for different bins in $Q \in [4.0, 9.00]$ GeV.

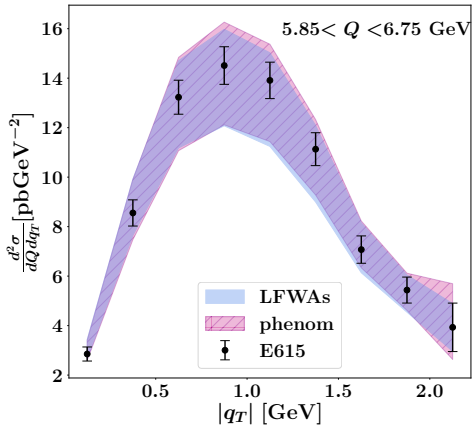
E615

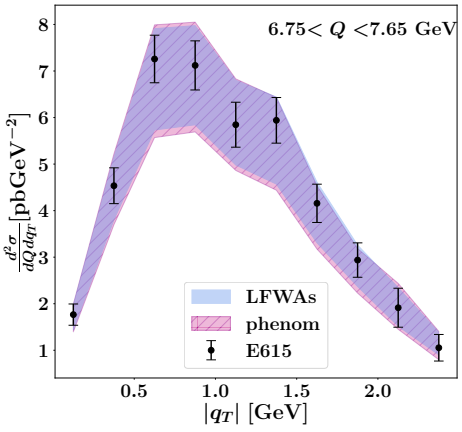

4.05< Q <4.5 GeV
LFWAs
phenom
E615
$\frac{d^2\sigma}{dQdq_T}$[pbGeV^{-2}]
$|q_T|$ [GeV]


4.5< Q <4.95 GeV
LFWAs
phenom
E615
$\frac{d^2\sigma}{dQdq_T}$[pbGeV^{-2}]
$|q_T|$ [GeV]


4.95< Q <5.4 GeV
LFWAs
phenom
E615
$\frac{d^2\sigma}{dQdq_T}$[pbGeV^{-2}]
$|q_T|$ [GeV]


5.4< Q <5.85 GeV
LFWAs
phenom
E615
$\frac{d^2\sigma}{dQdq_T}$[pbGeV^{-2}]
$|q_T|$ [GeV]


5.85< Q <6.75 GeV
LFWAs
phenom
E615
$\frac{d^2\sigma}{dQdq_T}$[pbGeV^{-2}]
$|q_T|$ [GeV]


6.75< Q <7.65 GeV
LFWAs
phenom
E615
$\frac{d^2\sigma}{dQdq_T}$[pbGeV^{-2}]
$|q_T|$ [GeV]

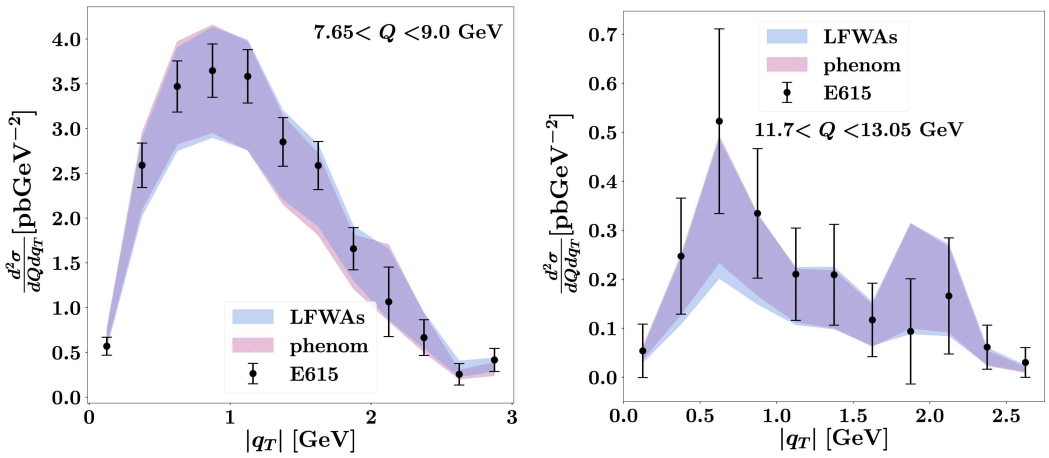


Figure 5.6: Comparison between the experimental DY data of E615 (black points), the pure phenomenological approach (pink bands) and the model based on LFWAs (light blue bands) for different bins in $Q \in [4.05, 13.05]$ GeV.

The error bands in the figures throughout this Section correspond to the 68% CL, and have been obtained by excluding the largest and the smallest 16% of the replicas, for both the pure phenomenological fit (pink bands) and the model based on the pion LFWAs (light blue bands).

In the various panels of Figs. 5.5 - 5.6, we plot the experimental differential DY cross sections for different bins in Q, as function of the transverse momentum of the virtual vector boson $|\boldsymbol{q}_T|$ (for E615) or its square $\boldsymbol{q}_T^2$ (for E537). As the chi-squared in Tabs. 5.3 - 5.4 already suggested for the replica 0, we can observe that the general shapes of all the experimental data are quite well reproduced by both fits. For both the models, the overall quality of the fits of the E537 data is worse compared to that for the E615 data set. By looking at the plots, we observe that the points at low $\boldsymbol{q}_T^2$ are generally poorly described. This issue is particularly evident for the model based on LFWAs, where the uncorrelated chi-squared of the replica 0 for E537 is five times greater than its contribution for E615.

Furthermore, the contribution due to the penalty chi-squared is greater for E615 compared to E537 in both approaches. This signals the presence of larger systematic uncertainties associated with the E615 data set, reflected in the wider bands.

In summary, both fits provide a good overall description of the data, although the presence of normalization errors significantly impacts the final results.

5.5.2 Comparison of pion TMDs

The TMDs extracted by the two fits are shown in Fig. 5.7. The results are for the TMD of the d quark in the π^-, plotted at the two scales $Q = 2$ GeV (upper panel) and $Q = 10$ GeV (lower panel), as function of the quark

transverse momentum $|\boldsymbol{k}_\perp|$ for three different values of $x = 0.1$, $x = 0.2$ and $x = 0.3$.

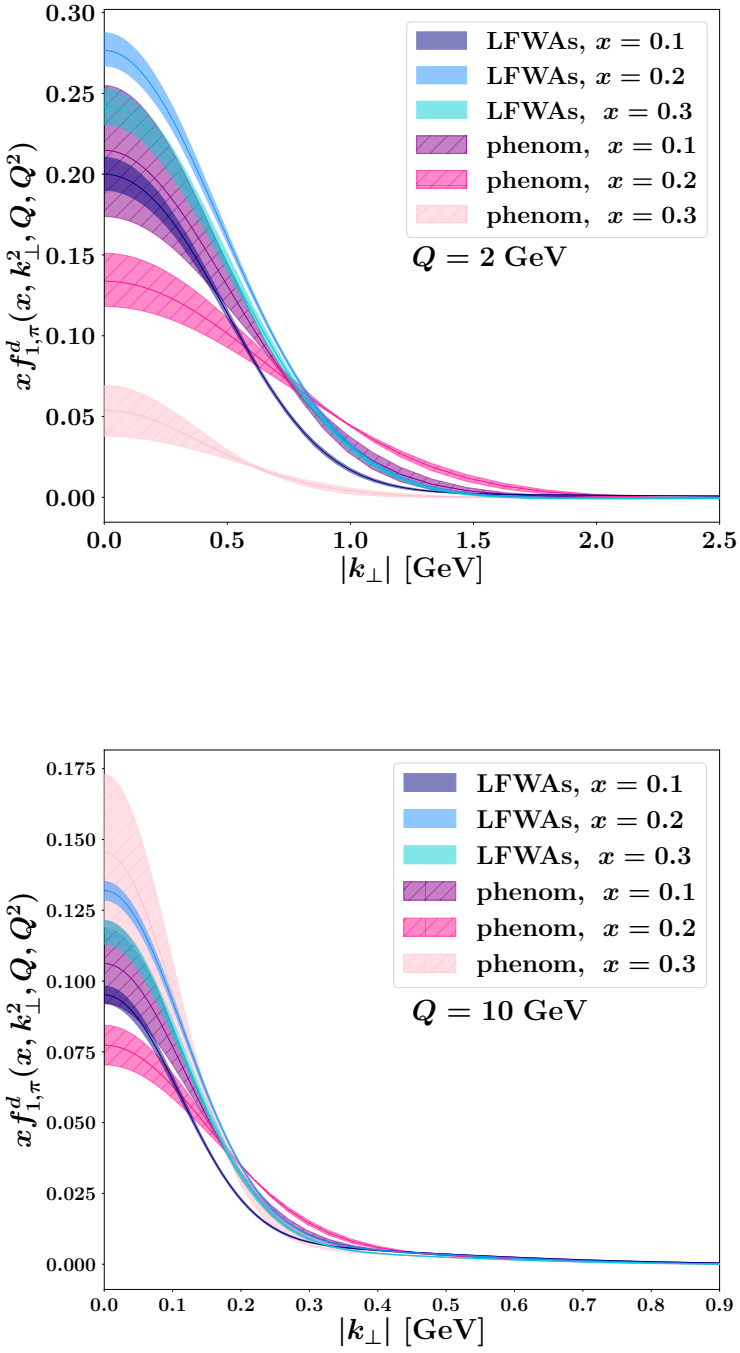


Figure 5.7: Comparison of the TMDs within the pure phenomenological and the LFWA approach, as function of the quark transverse momentum $|\boldsymbol{k}_\perp|$ at $x = 0.1$, $x = 0.2$ and $x = 0.3$ for $Q = 2$ GeV (upper panel) and $Q = 10$ GeV (lower panel).

For $Q = 10$ GeV, the values of x were selected to correspond to the region covered by the experimental data (see Fig. 5.1). The results at $Q = 2$ GeV have been plotted at the same values of x to highlight the effects of the evolution.

In both the plots, we can observe that the error bands in the phenomenological work are larger in comparison to the model employing the LFWA approach. This difference can likely be attributed to the different sets of pion collinear PDFs that have been inserted in Eq. (5.9). The error bands of xFitter are, in fact, larger than those of the PDFs extracted in Chapter (3), especially for $x \leq 0.3$, as we can see from the plots in Fig. 3.7.

5.5.3 Predictions for future experiments

The COMPASS Collaboration has recently released data for (un)polarized azimuthal asymmetries in the (polarized) pion–induced Drell–Yan processes [201], and will probably release also data for the unpolarized cross sections in the near future. The new COMPASS data will extend the experimental coverage of Fig. 5.1 and will be used, once published, for future extractions of pion TMDs.

We can, for now, compute the predictions in the same kinematic region of Ref. [201], by using the two models developed in this Chapter. Our theoretical predictions are plotted in Fig. 5.8 as functions of $|\boldsymbol{q}_T|$. The left panel refers

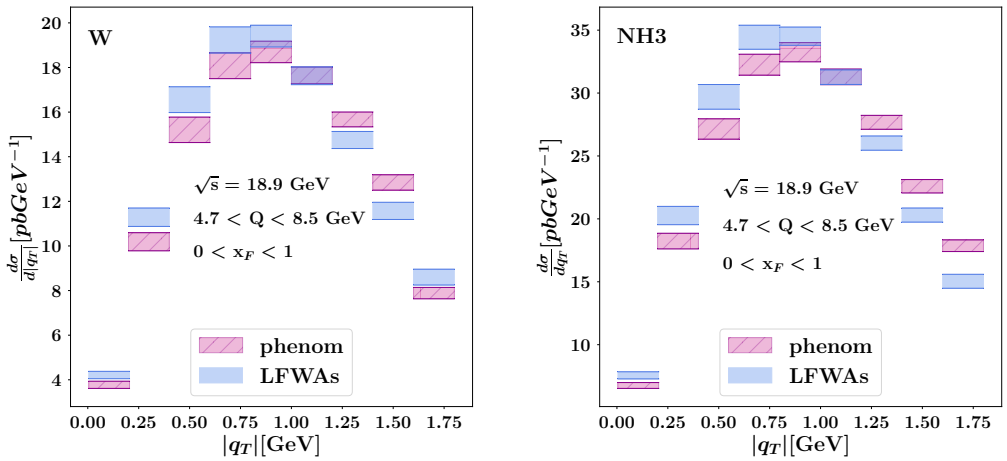


Figure 5.8: Theoretical predictions for future COMPASS data within the phenomenological approach and the model based on the LFWA parametrization.

to the tungsten nucleus, while the right panel to the ammonia molecule. The large uncertainties in the theoretical predictions stem from the wide bands in the fits of the DY data, that, in their turn, are a result of the limited quantity and precision of the existing experimental measurements.

The upcoming COMPASS data could help to better constrain the TMDs and to shed light on the normalization issue between theory and data.

We conclude by emphasizing that our analysis incorporates the entire available data set of DY cross sections for the first time, differently from existing works in literature [184, 185], which performed the fit with only the sets of E615.

Chapter 6

Pion generalized parton distributions

After the investigation of the PDFs, the e.m. FFs and the TMDs, we are now in the position to study the Generalized Parton Distributions of the pion.

The GPDs represent a relatively recent area of study, introduced in the latter years of the 20-th century. They have been the subject of extensive theoretical and experimental research for nucleons [48, 202–213].

The GPDs can be used to describe multiple exclusive processes, as, for example, Deeply Virtual Meson Production [214–217] (DVMP) or the Deeply Virtual Compton Scattering [203, 218–221] (DVCS). Moreover, they have gained substantial relevance due to their connection with the energy-momentum tensor form factors [203, 222–226] and the encoded information on the three-dimensional reconstruction of the internal structure of the hadrons in a mixed coordinate-momentum space [28, 29, 227].

Despite their intriguing properties, determining GPDs presents significant challenges in both the extraction from experimental data and non-perturbative calculations (see, e.g., Refs. [217, 221] for a recent overview on GPD phenomenology.)

On the experimental side, achieving precise measurements of GPD-related exclusive processes is difficult due to the need to preserve the integrity of the hadronic target, requiring high luminosity.

On the theoretical side, various approaches have been used to compute GPDs. Among them, we mention the Lattice QCD simulations, which permit to determine the Mellin moments of GPDs from first principles [228–234]. Recently, significant progress has been made towards new methods in lattice computation to access the the x - and t - dependence of GPDs by using the quasi-distribution approach and pseudo-distribution approach [235–243]. Another possibility is represented by the parametrization of GPDs with ansätze to be used in phenomenological analyses. In this context, the hadronic matrix elements can be parametrized in terms of double distributions modelled by assuming factorized t-dependence determined by form factors [208, 244].

Alternatively, the so-called dual representation has been proposed, which is constrained by data on the first Mellin moments or from simultaneous fits of data and lattice calculations of the higher Mellin moments [245–248]. A third approach involves the direct calculations of GPDs. We mention the bag model [249], the chiral quark-soliton models [250–253], the Nambu-Jona-Lasinio model [254], the models exploiting a non-relativistic approach [255] and a light-front Hamiltonian approach [256], the light-cone constituent models [257–260] and the meson-cloud models [261].

The pion has been a central focus in the studies of GPDs due to its unique role within the SM. Many theoretical studies have explored pion GPDs within different approaches [51, 66, 115, 125, 259, 262–278] and recent efforts have evaluated the pion DA and PDFs using continuum Schwinger methods (CSMs) [85, 86, 115, 173, 264, 265].

If the measurements of nucleon GPDs are already difficult to realize, extracting pion GPDs is even more challenging, due to the unstable nature of pions and - at present - no experimental data is available. Thus, theoretical predictions are the only source of information regarding pion GPDs.

For the future, new experimental updates have been proposed, with the aim of measuring more data of hadron GPDs. The upcoming experiments with 12 GeV at JLab and COMPASS-II at CERN are expected to significantly enhance precision. There is a comprehensive plan for conducting a series of DVCS and DVMP experiments at Hall A, Hall B and Hall C [63]. The new data have the potential of improving the current knowledge of the hadronic structure.

6.1 Correlator for the GPDs

The GPDs represent an extension of the collinear PDFs and contain information encoded in the FFs as well. In fact, under two different kinematic limits, we recover the PDFs and the FFs from the GPDs, as reported in Section 6.2 and shown in Fig. 2.5.

The variables on which the GPDs depend are $\bar{x}$ and Δ, where $\bar{x}$ is the fraction of average collinear momentum carried by the active parton, defined as

$$\bar{x} = \frac{\bar{k}^+}{P^+}, \qquad \bar{k} = \frac{k + k'}{2}, \tag{6.1}$$

while k and k' are the momenta of the active parton in the initial and in the final state, respectively. Generally, the GPDs are expressed as functions of $(\bar{x}, \xi, t)$, where

$$t = (p' - p)^2 = \Delta^2, \tag{6.2}$$

$$\xi = \frac{(p - p')^+}{(p + p')^+} = -\frac{\Delta^+}{2P^+}, \tag{6.3}$$

where p (p') is the four-momentum of the initial (final) hadron. The skewness, ξ, describes the change in the plus momentum of the hadron and t is one of the standard Mandelstam variables denoting the squared four-momentum transferred to the hadron.

In the symmetric frame, the four-momenta of the incoming and outgoing hadron can be written as

$$p = \left[(1+\xi)\,\bar{P}^+, \frac{M^2 + \boldsymbol{\Delta}_\perp^2/4}{2\,(1+\xi)\,\bar{P}^+}, -\frac{\boldsymbol{\Delta}_\perp}{2}\right], \tag{6.4}$$

$$p' = \left[(1-\xi)\,\bar{P}^+, \frac{M^2 + \boldsymbol{\Delta}_\perp^2/4}{2\,(1-\xi)\,\bar{P}^+}, \frac{\boldsymbol{\Delta}_\perp}{2}\right], \tag{6.5}$$

and the square of the momentum transfer reads

$$t = -\frac{4\xi^2 M^2 + \boldsymbol{\Delta}_\perp^2}{1-\xi^2}. \tag{6.6}$$

From Eq. (6.6) the minimum value of $-t$ results

$$-t \geq -t_{min} = \frac{4\xi^2 M^2}{1-\xi^2}, \tag{6.7}$$

for fixed skewness or, alternatively,

$$-\sqrt{\frac{-t}{4M^2 - t}} \leq \xi \leq \sqrt{\frac{-t}{4M^2 - t}}. \tag{6.8}$$

for fixed t.

At leading-twist with $\Gamma = \gamma^+$, the quark and gluon GPDs are labelled as H^q and H^g respectively, and their definitions in terms of matrix elements are the following

$$H^q\,(\bar{x},\xi,t) = \frac{1}{2}\int \frac{d\zeta^-}{2\pi} e^{i\bar{k}^+\zeta^-}\,\langle \pi(p')|\,\bar{\Psi}^q\,(-\zeta/2)\,\mathcal{U}_{(-\zeta/2,\zeta/2)}\gamma^+\Psi^q\,(\zeta/2)\,|\pi(p)\rangle\,\Big|_{\substack{\zeta^+=0\\ \boldsymbol{\zeta}_\perp=\mathbf{0}}}, \tag{6.9}$$

$$H^g\,(\bar{x},\xi,t) = -\frac{g_{\perp ii'}}{P^+}\int \frac{d\zeta^-}{2\pi} e^{i\bar{k}^+\zeta^-}\,\langle \pi(p')|G^{+i'}{}_a(-\zeta/2)G_a^{+i}\,(\zeta/2)|\pi(p)\rangle\Big|_{\substack{\zeta^+=0\\ \boldsymbol{\zeta}_\perp=\mathbf{0}}}, \tag{6.10}$$

where the dependence on the renormalization scale is understood and the link operator reduces to the identity due to the choice of the light-cone gauge (2.64).

6.2 Properties of the GPDs

Isospin invariance

In this Section we refer to the pion GPDs for a π^+, by using the notations $H^{q/g}_{\pi^+} \equiv H^{q/g}_{\pi}$. For negative and neutral pions, the isospin invariance implies:

$$\begin{aligned} H^{u}_{\pi^+}(\bar{x},\xi,t) &= H^{d}_{\pi^-}(\bar{x},\xi,t)\,, \\ H^{d}_{\pi^+}(\bar{x},\xi,t) &= H^{u}_{\pi^-}(\bar{x},\xi,t)\,, \\ H^{u}_{\pi^+}(\bar{x},\xi,t) + H^{d}_{\pi^+}(\bar{x},\xi,t) &= 2H^{u}_{\pi^0}(\bar{x},\xi,t) = 2H^{d}_{\pi^0}(\bar{x},\xi,t)\,, \end{aligned} \tag{6.11}$$

while the gluon distribution is invariant in the three pion states.

Time-reversal invariance

The range in which the GPDs are defined is

$$\begin{cases} \bar{x} \in [-1,1] \\ \xi \in [-1,1] \end{cases} . \tag{6.12}$$

However, thanks to time-reversal invariance, we can restrict ourselves to consider only the region $\xi > 0$. In fact, it holds

$$H^{q/g}_{\pi}(\bar{x},\xi,t) = H^{q/g}_{\pi}(\bar{x},-\xi,t)\,. \tag{6.13}$$

The range in $\bar{x}$ can be divided into three different regions, each with its own different partonic interpretation, as we will discuss in Section 6.3. We refer to these regions as DGLAP1 when $\xi \leq \bar{x} \leq 1$, DGLAP2 when $-1 \leq \bar{x} \leq -\xi$ and ERBL when $-\xi \leq \bar{x} \leq \xi$.

Reduction to other distributions

The *forward limit* consists in imposing $t = \xi = 0$ or, equivalently, $\Delta = 0$. This corresponds, as illustrated in Fig. 2.5, to following the light-blue line from the black point of the GPDs to its endpoint at the PDFs. The forward limit, in fact, implies

$$\begin{aligned} H^{q}_{\pi}(\bar{x},0,0) &= f^{q}_{1}(x)\Theta(x) - f^{\bar{q}}_{1}(-x)\Theta(-x), \\ H^{g}_{\pi}(\bar{x},0,0) &= x f^{g}_{1}(x)\Theta(x) - x f^{g}_{1}(-x)\Theta(-x). \end{aligned} \tag{6.14}$$

By looking at the same diagram of Fig. 2.5, we observe that the other possible limit, corresponding to an integral over $\bar{x}$, reduces the GPDs to the definition of the e.m. FFs, i.e.,

$$\int_{-1}^{1} d\bar{x} H^{q}_{\pi}(\bar{x},\xi,t) = F^{q}_{1}(t). \tag{6.15}$$

Charge conjugation

The charge conjugation mapping between different charged pions implies the following relations for the pion GPDs

$$H^q_{\pi^+}(\bar{x},\xi,t) = -H^q_{\pi^-}(-\bar{x},\xi,t). \tag{6.16}$$

The latter, combined with Eq. (6.11), leads to the isospin symmetric limit

$$H^u_{\pi}(\bar{x},\xi,t) = -H^d_{\pi}(-\bar{x},\xi,t). \tag{6.17}$$

6.3 LFWA overlap representation for pion GPDs

In this Section we present the LFWA overlap representation of the pion GPDs across the three different regions of the $\bar{x}$ range. As with the other parton distribution functions, we can separate the quark and gluon GPDs of the pion in terms of the contributions from each parton configuration.

6.3.1 Hadron-in and hadron-out frames

The matrix elements defining the GPDs in Eqs. (6.9) - (6.10) contain the kinematics of the partons in the initial and in the final states. As introduced in Section 2.1.5, the LFWAs depend only on the relative momentum coordinates of the partons with respect to the parent hadron. The identification of the arguments of the LFWAs is most easily done when hadron frames are chosen as frames of reference. We introduce the names *hadron-in* and *hadron-out* for frames where the incoming and outgoing hadron has zero transverse momentum, respectively. We want to make explicit the arguments of the LFWAs of the incoming hadron in the hadron-in frame and those of the outgoing hadron in the hadron-out frame. The formal way to connect the symmetric frame of Eqs. (6.4) - (6.5) with the hadron-in/hadron-out frame is applying the following transverse boost

$$\left[v^+, v^-, \boldsymbol{v}_\perp\right] \to \left[v^+, v^- - \frac{\boldsymbol{v}_\perp \cdot \boldsymbol{a}_\perp}{a^+} + \frac{v^+ \boldsymbol{a}_\perp^2}{2\,(a^+)^2}, \boldsymbol{v}_\perp - \frac{v^+}{a^+}\boldsymbol{a}_\perp\right]. \tag{6.18}$$

where

$$\begin{cases} a^+ = (1+\xi)\,P^+ \\ \boldsymbol{a}_\perp = -\dfrac{\boldsymbol{\Delta}_\perp}{2} \end{cases} \quad \text{for the hadron-in frame,} \tag{6.19}$$

$$\begin{cases} a^+ = (1-\xi)\,P^+ \\ \boldsymbol{a}_\perp = \dfrac{\boldsymbol{\Delta}_\perp}{2} \end{cases} \quad \text{for the hadron-out frame.} \tag{6.20}$$

For the sake of clarity we will distinguish from now the kinematic variables of the initial state (written in the hadron-in frame) from those of the final

state (expressed in the the hadron-out frame). To this end we insert the superscripts "in"/"out" in the kinematic variables relative to the partons in the initial/final states appearing as arguments of the LFWAs. In order to achieve a formulation that exhibits symmetry between incoming and outgoing quantities, we introduce, in analogy to Eq. (6.1), the following auxiliary variables for the partons

$$\bar{x}_i = \frac{\bar{k}_i^+}{P^+}, \qquad \bar{k}_i = \frac{1}{2}\left(k_i + k_i'\right). \tag{6.21}$$

These relations generate a new integration measure that will be used in the following and that replaces Eqs. (2.87) - (2.88). Such a measure is

$$[d\bar{x}]_N = \prod_{i=1}^{N} \frac{d\bar{x}_i}{\sqrt{\bar{x}_i}} \delta\left(1 - \sum_{i=1}^{N} \bar{x}_i\right), \tag{6.22}$$

$$\left[d^2\bar{\boldsymbol{k}}_\perp\right]_N = \frac{1}{\left[2(2\pi)^3\right]^{N-1}} \prod_{i=1}^{N} d^2\bar{\boldsymbol{k}}_{\perp i} \delta^{(2)}\left(\sum_{i=1}^{N} \bar{\boldsymbol{k}}_{\perp i}\right). \tag{6.23}$$

6.3.2 DGLAP1 region

The DGLAP1 region corresponds to the range $\xi \leq \bar{x} \leq 1$ and can be physically interpreted, at the parton level, as the emission of a quark or a gluon with momentum fraction $\bar{x}+\xi$ from the initial pion state and its re-absorption with momentum fraction $\bar{x}-\xi$ from the final pion. The schematic representation is shown in Fig. 6.1.

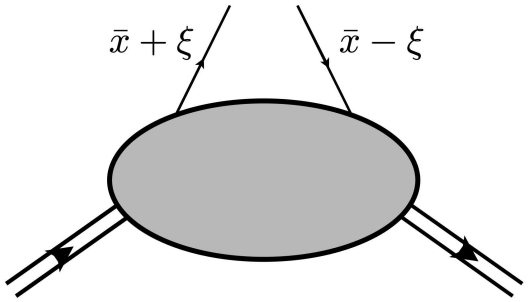


Figure 6.1: The partonic interpretation of GPDs in the DGLAP1 region $\xi \leq \bar{x} \leq 1$, corresponding to the emission of quark from the initial hadron and its re-absorption from the final hadron.

In this transition, the pion keeps the same parton content in the initial and final state. We can then write the quark GPD as a sum over contributions from each parton configuration in Eq. (2.133). The GPD of the up quark contains the contributions from all the Fock states, while for the down and the strange quark the only non-vanishing terms come from the sea Fock components. Moreover, the latter two are equivalent, due to the assumption of SU(3)

symmetry:

$$H^u_\pi(\bar{x},\xi,t) = H^{u\,(u\bar{d}\to u\bar{d})}_\pi(\bar{x},\xi,t) + H^{u\,(u\bar{d}g\to u\bar{d}g)}_\pi(\bar{x},\xi,t) + H^{u\,(u\bar{d}gg\to u\bar{d}gg)}_\pi(\bar{x},\xi,t) + \sum_{\{s\bar{s}\}} H^{u\,(u\bar{d}\{s\bar{s}\}\to u\bar{d}\{s\bar{s}\})}_\pi(\bar{x},\xi,t), \tag{6.24}$$

$$H^d_\pi(\bar{x},\xi,t) = H^{d\,(u\bar{d}d\bar{d}\to u\bar{d}d\bar{d})}_\pi(\bar{x},\xi,t), \tag{6.25}$$

$$H^s_\pi(\bar{x},\xi,t) = H^{d\,(u\bar{d}s\bar{s}\to u\bar{d}s\bar{s})}_\pi(\bar{x},\xi,t) = H^d_\pi(\bar{x},\xi,t). \tag{6.26}$$

Using the expressions for the pion states in Eqs. (2.134)-(2.137) and the symmetry properties of the functions $\psi^{(1)}_{u\bar{d}gg}$ and $\psi^{(2)}_{u\bar{d}gg}$, we find the following expressions for the individual contributions in Eqs. (6.24) - (6.26).

$$H^{u\,(u\bar{d}\to u\bar{d})}_\pi(\bar{x},\xi,t) = 2\int [d\bar{x}]_2[d^2\bar{\boldsymbol{k}}_\perp]_2\,\sqrt{\bar{x}_1\bar{x}_2}\,\delta(\bar{x}-\bar{x}_1) \times \psi^{(1)*}_{u\bar{d}}(1^{\rm out},2^{\rm out})\psi^{(1)}_{u\bar{d}}(1^{\rm in},2^{\rm in}), \tag{6.27}$$

$$H^{u\,(u\bar{d}g\to u\bar{d}g)}_\pi(\bar{x},\xi,t) = 2\frac{1}{\sqrt{1-\xi^2}}\int [d\bar{x}]_3[d^2\bar{\boldsymbol{k}}_\perp]_3\,\sqrt{\bar{x}_1\bar{x}_2\bar{x}_3}\,\delta(\bar{x}-\bar{x}_1) \times \psi^{(1)*}_{u\bar{d}g}(1^{\rm out},2^{\rm out},3^{\rm out})\psi^{(1)}_{u\bar{d}g}(1^{\rm in},2^{\rm in},3^{\rm in}), \tag{6.28}$$

$$H^{u\,(u\bar{d}gg\to u\bar{d}gg)}_\pi(\bar{x},\xi,t) = \frac{8}{1-\xi^2}\int [d\bar{x}]_4[d^2\bar{\boldsymbol{k}}_\perp]_4\sqrt{\bar{x}_1\bar{x}_2\bar{x}_3\bar{x}_4}\delta(\bar{x}-\bar{x}_1) \times \big\{\psi^{(1)*}_{u\bar{d}gg}(1^{\rm out},2^{\rm out},3^{\rm out},4^{\rm out})\psi^{(1)}_{u\bar{d}gg}(1^{\rm in},2^{\rm in},3^{\rm in},4^{\rm in}) + \psi^{(2)*}_{u\bar{d}gg}(1^{\rm out},2^{\rm out},3^{\rm out},4^{\rm out})\psi^{(2)}_{u\bar{d}gg}(1^{\rm in},2^{\rm in},3^{\rm in},4^{\rm in})\big\}, \tag{6.29}$$

$$H^{u\,(u\bar{d}\{s\bar{s}\}\to u\bar{d}\{s\bar{s}\})}_\pi(\bar{x},\xi,t) = 4\frac{1}{1-\xi^2}\int [d\bar{x}]_4[d^2\bar{\boldsymbol{k}}_\perp]_4\sqrt{\bar{x}_1\bar{x}_2\bar{x}_3\bar{x}_4} \times \Big[\delta(\bar{x}-\bar{x}_1)+\delta_{s,u}\delta_{\bar{s},\bar{u}}\delta(\bar{x}-\bar{x}_3)\Big] \times \Big[\psi^{(1)*}_{u\bar{d}\{s\bar{s}\}}(1^{\rm out},2^{\rm out},3^{\rm out},4^{\rm out})\psi^{(1)}_{u\bar{d}\{s\bar{s}\}}(1^{\rm in},2^{\rm in},3^{\rm in},4^{\rm in}) + \psi^{(2)*}_{u\bar{d}\{s\bar{s}\}}(1^{\rm out},2^{\rm out},3^{\rm out},4^{\rm out})\psi^{(2)}_{u\bar{d}\{s\bar{s}\}}(1^{\rm in},2^{\rm in},3^{\rm in},4^{\rm in}) + \frac{1}{2}\psi^{(3)*}_{u\bar{d}\{s\bar{s}\}}(1^{\rm out},2^{\rm out},3^{\rm out},4^{\rm out}) \times \psi^{(3)}_{u\bar{d}\{s\bar{s}\}}(1^{\rm in},2^{\rm in},3^{\rm in},4^{\rm in})\Big], \tag{6.30}$$

$$H^{d\,(u\bar{d}d\bar{d}\to u\bar{d}d\bar{d})}_\pi(\bar{x},\xi,t) = 4\frac{1}{1-\xi^2}\int [d\bar{x}]_4[d^2\bar{\boldsymbol{k}}_\perp]_4\sqrt{\bar{x}_1\bar{x}_2\bar{x}_3\bar{x}_4}\delta(\bar{x}-\bar{x}_3) \times \Big[\psi^{(1)*}_{u\bar{d}d\bar{d}}(1^{\rm out},2^{\rm out},3^{\rm out},4^{\rm out})\psi^{(1)}_{u\bar{d}d\bar{d}}(1^{\rm in},2^{\rm in},3^{\rm in},4^{\rm in}) + \psi^{(2)*}_{u\bar{d}d\bar{d}}(1^{\rm out},2^{\rm out},3^{\rm out},4^{\rm out})\psi^{(2)}_{u\bar{d}d\bar{d}}(1^{\rm in},2^{\rm in},3^{\rm in},4^{\rm in}) + \frac{1}{2}\psi^{(3)*}_{u\bar{d}d\bar{d}}(1^{\rm out},2^{\rm out},3^{\rm out},4^{\rm out})\psi^{(3)}_{u\bar{d}d\bar{d}}(1^{\rm in},2^{\rm in},3^{\rm in},4^{\rm in})\Big]. \tag{6.31}$$

For the gluon it holds

$$H_\pi^g(\bar{x},\xi,t) = H_\pi^{g\,(u\bar{d}g\to u\bar{d}g)}(\bar{x},\xi,t) + H_\pi^{g\,(u\bar{d}gg\to u\bar{d}gg)}(\bar{x},\xi,t), \tag{6.32}$$

where the two contributions of the r.h.s. are

$$\begin{aligned} H_\pi^{g\,(u\bar{d}g\to u\bar{d}g)}(\bar{x},\xi,t) &= 2\frac{\sqrt{\bar{x}^2-\xi^2}}{\sqrt{1-\xi^2}}\int [d\bar{x}]_3[d^2\bar{\boldsymbol{k}}_\perp]_3\sqrt{\bar{x}_1\bar{x}_2\bar{x}_3}\delta(\bar{x}-\bar{x}_3) \\ &\quad\times \psi^{(1)*}_{u\bar{d}g}(1^{\rm out},2^{\rm out},3^{\rm out})\psi^{(1)}_{u\bar{d}g}(1^{\rm in},2^{\rm in},3^{\rm in}), \end{aligned} \tag{6.33}$$

$$\begin{aligned} H_\pi^{g\,(u\bar{d}gg\to u\bar{d}gg)}(\bar{x},\xi,t) &= 16\frac{\sqrt{\bar{x}^2-\xi^2}}{\sqrt{1-\xi^2}}\int [d\bar{x}]_4[d^2\bar{\boldsymbol{k}}_\perp]_4\sqrt{\bar{x}_1\bar{x}_2\bar{x}_3\bar{x}_4}\delta(\bar{x}-\bar{x}_3) \\ &\quad\times\Big\{\psi^{(1)*}_{u\bar{d}gg}(1^{\rm out},2^{\rm out},3^{\rm out},4^{\rm out})\psi^{(1)}_{u\bar{d}gg}(1^{\rm in},2^{\rm in},3^{\rm in},4^{\rm in}) \\ &\quad+\psi^{(2)*}_{u\bar{d}gg}(1^{\rm out},2^{\rm out},3^{\rm out},4^{\rm out})\psi^{(2)}_{u\bar{d}gg}(1^{\rm in},2^{\rm in},3^{\rm in},4^{\rm in})\Big\}. \end{aligned} \tag{6.34}$$

The arguments $i = 1,2,3,4$ of the LFWAs in Eqs. (6.27)-(6.34) stand for $(x_i, \boldsymbol{k}_{\perp i})$ with the superscripts in/out denoting the momenta of the partons belonging to the incoming/outgoing hadron in the hadron-in/hadron-out frame. The parton emitted and later reabsorbed from the hadron is called the "active" parton and labelled with index j; all other partons $i \neq j$ play the role of "spectators". In the LFWA overlap representation of the GPDs the active parton is identified by the delta function that enforces its average longitudinal momentum fraction to be equal to $\bar{x}$. The LFWA arguments for the incoming hadron are related to the momenta in the average-frame by

$$x_i^{\rm in} = \frac{\bar{x}_i}{1+\xi}, \qquad \boldsymbol{k}_{\perp i}^{\rm in} = \bar{\boldsymbol{k}}_{\perp i} + \frac{\bar{x}_i}{1+\xi}\frac{\boldsymbol{\Delta}}{2}, \quad \text{for } i\neq j, \tag{6.35}$$

$$x_j^{\rm in} = \frac{\bar{x}_j+\xi}{1+\xi}, \qquad \boldsymbol{k}_{\perp j}^{\rm in} = \bar{\boldsymbol{k}}_{\perp j} - \frac{1-\bar{x}_j}{1+\xi}\frac{\boldsymbol{\Delta}}{2}. \tag{6.36}$$

Likewise, the LFWA arguments for the outgoing hadron are related to the momenta in the average-frame by

$$x_i^{\rm out} = \frac{\bar{x}_i}{1-\xi}, \qquad \boldsymbol{k}_{\perp i}^{\rm out} = \bar{\boldsymbol{k}}_{\perp i} - \frac{\bar{x}_i}{1-\xi}\frac{\boldsymbol{\Delta}}{2}, \quad \text{for } i\neq j, \tag{6.37}$$

$$x_j^{\rm out} = \frac{\bar{x}_j-\xi}{1-\xi}, \qquad \boldsymbol{k}_{\perp j}^{\rm out} = \bar{\boldsymbol{k}}_{\perp j} + \frac{1-\bar{x}_j}{1-\xi}\frac{\boldsymbol{\Delta}}{2}. \tag{6.38}$$

The fractions of momenta (6.35) - (6.38) satisfy the following relations from momentum conservation

$$x_j^{\rm in/out} = 1 - \sum_{i\neq j} x_i^{\rm in/out}, \qquad \boldsymbol{k}_{\perp j}^{\rm in/out} = -\sum_{i\neq j}\boldsymbol{k}_{\perp i}^{\rm in/out}. \tag{6.39}$$

6.3.3 DGLAP2 region

In the DGLAP2, $\bar{x}$ varies in the range $-1 \leq \bar{x} \leq -\xi$. In this region, the partonic interpretation of the quark GPD is exactly the same as in the DGLAP1 region, but with an antiquark replacing a quark (see Fig. 6.2).

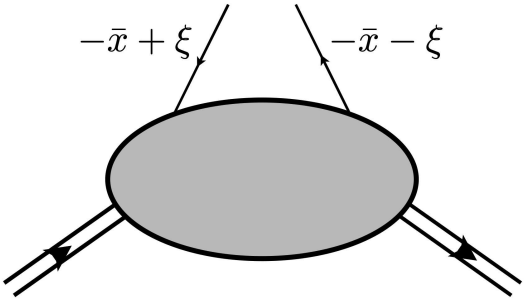


Figure 6.2: The partonic interpretation of GPDs in the DGLAP2 region $-1 \leq \bar{x} \leq -\xi$, corresponding to the emission of antiquark from the initial hadron and its re-absorption from the final hadron.

The analogous relations to Eqs. (6.24) - (6.26) that holds in this regions are as follows

$$H^u_\pi(\bar{x},\xi,t) = H^{u\,(u\bar{d}u\bar{u}\to u\bar{d}u\bar{u})}_\pi(\bar{x},\xi,t), \tag{6.40}$$

$$H^d_\pi(\bar{x},\xi,t) = H^{d\,(u\bar{d}\to u\bar{d})}_\pi(\bar{x},\xi,t) + H^{d\,(u\bar{d}g\to u\bar{d}g)}_\pi(\bar{x},\xi,t) + H^{d\,(u\bar{d}gg\to u\bar{d}gg)}_\pi(\bar{x},\xi,t) + \sum_{\{s\bar{s}\}} H^{d\,(u\bar{d}\{s\bar{s}\}\to u\bar{d}\{s\bar{s}\})}_\pi(\bar{x},\xi,t), \tag{6.41}$$

$$H^s_\pi(\bar{x},\xi,t) = H^{s\,(u\bar{d}s\bar{s}\to u\bar{d}s\bar{s})}_\pi(\bar{x},\xi,t) = H^u_\pi(\bar{x},\xi,t), \tag{6.42}$$

where the contributions from the individual Fock state are

$$\begin{aligned} H^{u\,(u\bar{d}u\bar{u}\to u\bar{d}u\bar{u})}_\pi(\bar{x},\xi,t) = -4\frac{1}{1-\xi^2}\int [d\bar{x}]_4 [d^2\bar{\boldsymbol{k}}_\perp]_4 \sqrt{\bar{x}_1\bar{x}_2\bar{x}_3\bar{x}_4}\,\delta(\bar{x}+\bar{x}_4) \\ \times \Big[\psi^{(1)*}_{u\bar{d}d\bar{d}}(1^{\rm out},2^{\rm out},3^{\rm out},4^{\rm out})\psi^{(1)}_{u\bar{d}d\bar{d}}(1^{\rm in},2^{\rm in},3^{\rm in},4^{\rm in}) \\ + \psi^{(2)*}_{u\bar{d}d\bar{d}}(1^{\rm out},2^{\rm out},3^{\rm out},4^{\rm out})\psi^{(2)}_{u\bar{d}d\bar{d}}(1^{\rm in},2^{\rm in},3^{\rm in},4^{\rm in}) \\ + \frac{1}{2}\psi^{(3)*}_{u\bar{d}d\bar{d}}(1^{\rm out},2^{\rm out},3^{\rm out},4^{\rm out})\psi^{(3)}_{u\bar{d}d\bar{d}}(1^{\rm in},2^{\rm in},3^{\rm in},4^{\rm in})\Big]. \end{aligned} \tag{6.43}$$

$$\begin{aligned} H^{d\,(u\bar{d}\to u\bar{d})}_\pi(\bar{x},\xi,t) = -2\int [d\bar{x}]_2 [d^2\bar{\boldsymbol{k}}_\perp]_2\,\sqrt{\bar{x}_1\bar{x}_2}\,\delta(\bar{x}+\bar{x}_2) \\ \times \psi^{(1)*}_{u\bar{d}}(1^{\rm out},2^{\rm out})\psi^{(1)}_{u\bar{d}}(1^{\rm in},2^{\rm in}), \end{aligned} \tag{6.44}$$

$$H^{d\,(u\bar{d}g\to u\bar{d}g)}_\pi(\bar{x},\xi,t) = -2\frac{1}{\sqrt{1-\xi^2}}\int [d\bar{x}]_3 [d^2\bar{\boldsymbol{k}}_\perp]_3\,\sqrt{\bar{x}_1\bar{x}_2\bar{x}_3}\,\delta(\bar{x}+\bar{x}_2)$$

$$\times \psi^{(1)*}_{u\bar{d}g}(1^{\rm out},2^{\rm out},3^{\rm out})\psi^{(1)}_{u\bar{d}g}(1^{\rm in},2^{\rm in},3^{\rm in}), \quad (6.45)$$

$$H^{d\,(u\bar{d}gg\to u\bar{d}gg)}_{\pi}(\bar{x},\xi,t) = -\frac{8}{1-\xi^2}\int [d\bar{x}]_4[d^2\bar{\boldsymbol{k}}_\perp]_4\sqrt{\bar{x}_1\bar{x}_2\bar{x}_3\bar{x}_4}\delta(\bar{x}+\bar{x}_2)$$
$$\times\Big\{\psi^{(1)*}_{u\bar{d}gg}(1^{\rm out},2^{\rm out},3^{\rm out},4^{\rm out})\psi^{(1)}_{u\bar{d}gg}(1^{\rm in},2^{\rm in},3^{\rm in},4^{\rm in})$$
$$+\psi^{(2)*}_{u\bar{d}gg}(1^{\rm out},2^{\rm out},3^{\rm out},4^{\rm out})\psi^{(2)}_{u\bar{d}gg}(1^{\rm in},2^{\rm in},3^{\rm in},4^{\rm in})\Big\}, \quad (6.46)$$

$$H^{d\,(u\bar{d}\{s\bar{s}\}\to u\bar{d}\{s\bar{s}\})}_{\pi}(\bar{x},\xi,t) = -4\frac{1}{1-\xi^2}\int [d\bar{x}]_4[d^2\bar{\boldsymbol{k}}_\perp]_4\sqrt{\bar{x}_1\bar{x}_2\bar{x}_3\bar{x}_4}$$
$$\times\Big[\delta(\bar{x}+\bar{x}_2)+\delta_{s,d}\delta_{\bar{s},\bar{d}}\delta(\bar{x}+\bar{x}_4)\Big]$$
$$\times\Big[\psi^{(1)*}_{u\bar{d}\{s\bar{s}\}}(1^{\rm out},2^{\rm out},3^{\rm out},4^{\rm out})\psi^{(1)}_{u\bar{d}\{s\bar{s}\}}(1^{\rm in},2^{\rm in},3^{\rm in},4^{\rm in})$$
$$+\psi^{(2)*}_{u\bar{d}\{s\bar{s}\}}(1^{\rm out},2^{\rm out},3^{\rm out},4^{\rm out})\psi^{(2)}_{u\bar{d}\{s\bar{s}\}}(1^{\rm in},2^{\rm in},3^{\rm in},4^{\rm in})$$
$$+\frac{1}{2}\psi^{(3)*}_{u\bar{d}\{s\bar{s}\}}(1^{\rm out},2^{\rm out},3^{\rm out},4^{\rm out})$$
$$\times\psi^{(3)}_{u\bar{d}\{s\bar{s}\}}(1^{\rm in},2^{\rm in},3^{\rm in},4^{\rm in})\Big]. \quad (6.47)$$

The arguments of the LFWAs in the DGLAP2 are exactly the same as those in the DLAP1 region, i.e, Eqs. (6.35) - (6.38). We observe that Eqs. (6.44) - (6.47) can be derived from Eqs. (6.27) - (6.30) by reversing the overall sign and changing $\delta(\bar{x}-\bar{x}_1)$ into $\delta(\bar{x}+\bar{x}_2)$ and $\delta(\bar{x}-\bar{x}_3)$ into $\delta(\bar{x}+\bar{x}_4)$. For the gluon, the decomposition of the GPD is exactly Eq. (6.32) and the two Fock state contributions are the same as in Eqs. (6.33) - (6.34) with the exchange of $\delta(\bar{x}-\bar{x}_j)$ with $\delta(\bar{x}+\bar{x}_j)$. This is due to the facts that the gluon is its own antiparticle and the gluon GPD is an even function in $\bar{x}$.

6.3.4 ERBL region

The ERBL region corresponds to the range $-\xi<\bar{x}<\xi$ and the partonic interpretation is the emission from the initial pion of a particle-antiparticle pair, as pictured in Fig 6.3. In the Fock-space decomposition, this implies that the parton composition of the initial hadron should contain an additional quark-antiquark, or two additional gluons, w.r.t. the Fock-state of the final hadron.

Considering the Fock states included in the model (see Eq. (2.133)), we find the following four contributions in the ERBL region

$$H^u_\pi(\bar{x},\xi,t) = H^{u\,(u\bar{d}u\bar{u}\to u\bar{d})}_\pi(\bar{x},\xi,t), \quad (6.48)$$
$$H^d_\pi(\bar{x},\xi,t) = H^{q\,(u\bar{d}d\bar{d}\to u\bar{d})}_\pi(\bar{x},\xi,t), \quad (6.49)$$
$$H^s_\pi(\bar{x},\xi,t) = H^{s\,(u\bar{d}s\bar{s}\to u\bar{d})}_\pi(\bar{x},\xi,t), \quad (6.50)$$
$$H^g_\pi(\bar{x},\xi,t) = H^{g\,(u\bar{d}gg\to u\bar{d})}_\pi(\bar{x},\xi,t). \quad (6.51)$$

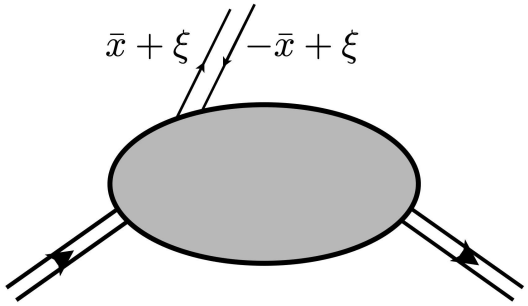


Figure 6.3: The partonic interpretation of GPDs in the ERBL region $-\xi \leq \bar{x} \leq \xi$, corresponding to the emission from the initial state of a quark-antiquark pair, or two gluons.

The corresponding LFWA overlap representation is given by

$$
\begin{aligned}
&H_\pi^{u\,(u\bar{d}u\bar{u}\to u\bar{d})}(\bar{x},\xi,t) \\
&\quad = \frac{4\sqrt{3}}{1+\xi}\Bigg\{\int d\bar{x}_1 d\bar{x}_2 d\bar{x}_3 \int \frac{d^2\bar{\boldsymbol{k}}_{\perp 1} d^2\bar{\boldsymbol{k}}_{\perp 2} d^2\bar{\boldsymbol{k}}_{\perp 3}}{(16\pi^3)^2} \\
&\qquad\qquad \times \delta(\bar{x}-\bar{x}_3)\delta(1-\xi-\bar{x}_1-\bar{x}_2)\delta^{(2)}\left(\bar{\boldsymbol{k}}_{\perp 1}+\bar{\boldsymbol{k}}_{\perp 2}-\frac{\boldsymbol{\Delta}_\perp}{2}\right) \\
&\qquad\qquad \times \psi^{(1)*}_{u\bar{d}}(1^{\rm out},2^{\rm out})\psi^{(1)}_{u\bar{d}u\bar{u}}(1^{\rm in},2^{\rm in},3^{\rm in},3'^{\rm in}) \\
&\qquad - \frac{1}{6}\int d\bar{x}_1 d\bar{x}_2 d\bar{x}_3 \int \frac{d^2\bar{\boldsymbol{k}}_{\perp 1} d^2\bar{\boldsymbol{k}}_{\perp 2} d^2\bar{\boldsymbol{k}}_{\perp 3}}{(16\pi^3)^2} \\
&\qquad\qquad \times \delta(\bar{x}-\bar{x}_1)\delta(1-\xi-\bar{x}_2-\bar{x}_3)\delta^{(2)}\left(\bar{\boldsymbol{k}}_{\perp 2}+\bar{\boldsymbol{k}}_{\perp 3}-\frac{\boldsymbol{\Delta}_\perp}{2}\right) \\
&\qquad\qquad \times \Bigg[\psi^{(1)*}_{u\bar{d}}(3^{\rm out},2^{\rm out})\psi^{(1)}_{u\bar{d}u\bar{u}}(1^{\rm in},2^{\rm in},3^{\rm in},1'^{\rm in}) \\
&\qquad\qquad\quad + \psi^{(1)*}_{u\bar{d}}(3^{\rm out},2^{\rm out})\psi^{(2)}_{u\bar{d}u\bar{u}}(1^{\rm in},2^{\rm in},3^{\rm in},1'^{\rm in}) \\
&\qquad\qquad\quad - \psi^{(1)*}_{u\bar{d}}(3^{\rm out},2^{\rm out})\psi^{(3)}_{u\bar{d}u\bar{u}}(1^{\rm in},2^{\rm in},3^{\rm in},1'^{\rm in})\Bigg]\Bigg\}, \qquad (6.52)
\end{aligned}
$$

$$
\begin{aligned}
&H_\pi^{d\,(u\bar{d}d\bar{d}\to u\bar{d})}(\bar{x},\xi,t) \\
&\quad = \frac{4\sqrt{3}}{1+\xi}\Bigg\{\int d\bar{x}_1 d\bar{x}_2 d\bar{x}_3 \int \frac{d^2\bar{\boldsymbol{k}}_{\perp 1} d^2\bar{\boldsymbol{k}}_{\perp 2} d^2\bar{\boldsymbol{k}}_{\perp 3}}{(16\pi^3)^2} \\
&\qquad\qquad \times \delta(\bar{x}-\bar{x}_3)\delta(1-\xi-\bar{x}_1-\bar{x}_2)\delta^{(2)}\left(\bar{\boldsymbol{k}}_{\perp 1}+\bar{\boldsymbol{k}}_{\perp 2}-\frac{\boldsymbol{\Delta}_\perp}{2}\right) \\
&\qquad\qquad \times \psi^{(1)*}_{u\bar{d}}(1^{\rm out},2^{\rm out})\psi^{(1)}_{u\bar{d}d\bar{d}}(1^{\rm in},2^{\rm in},3^{\rm in},3'^{\rm in}) \\
&\qquad - \frac{1}{6}\int d\bar{x}_1 d\bar{x}_3 d\bar{x}_4 \int \frac{d^2\bar{\boldsymbol{k}}_{\perp 1} d^2\bar{\boldsymbol{k}}_{\perp 3} d^2\bar{\boldsymbol{k}}_{\perp 4}}{(16\pi^3)^2}
\end{aligned}
$$

$$\times \delta(\bar{x}-\bar{x}_3)\delta(1-\xi-\bar{x}_1-\bar{x}_4)\delta^{(2)}\left(\bar{\boldsymbol{k}}_{\perp 1}+\bar{\boldsymbol{k}}_{\perp 4}-\frac{\boldsymbol{\Delta}_\perp}{2}\right)$$
$$\times\Bigg\{\Bigg[\psi^{(1)*}_{u\bar{d}}(1^{\rm out},4^{\rm out})\psi^{(1)}_{u\bar{d}d\bar{d}}(1^{\rm in},3'^{\rm in},3^{\rm in},4^{\rm in})$$
$$+\psi^{(1)*}_{u\bar{d}}(1^{\rm out},4^{\rm out})\psi^{(2)}_{u\bar{d}d\bar{d}}(1^{\rm in},3'^{\rm in},3^{\rm in},4^{\rm in})$$
$$+\psi^{(1)*}_{u\bar{d}}(1^{\rm out},4^{\rm out})\psi^{(3)}_{u\bar{d}d\bar{d}}(1^{\rm in},3'^{\rm in},3^{\rm in},4^{\rm in})\Bigg]\Bigg\}, \tag{6.53}$$

$$H_\pi^{s\,(u\bar{d}s\bar{s}\to u\bar{d})}(\bar{x},\xi,t)$$
$$=\frac{4\sqrt{3}}{1+\xi}\int d\bar{x}_1 d\bar{x}_2 d\bar{x}_3\int\frac{d^2\bar{\boldsymbol{k}}_{\perp 1}d^2\bar{\boldsymbol{k}}_{\perp 2}d^2\bar{\boldsymbol{k}}_{\perp 3}}{(16\pi^3)^2}$$
$$\times\delta(\bar{x}-\bar{x}_3)\delta(1-\xi-\bar{x}_1-\bar{x}_2)\delta^{(2)}\left(\bar{\boldsymbol{k}}_{\perp 1}+\bar{\boldsymbol{k}}_{\perp 2}-\frac{\boldsymbol{\Delta}_\perp}{2}\right)$$
$$\times\psi^{(1)*}_{u\bar{d}}(1^{\rm out},2^{\rm out})\psi^{(1)}_{u\bar{d}s\bar{s}}(1^{\rm in},2^{\rm in},3^{\rm in},3'^{\rm in}), \tag{6.54}$$

$$H_\pi^{g}(\bar{x},\xi,t)$$
$$=16\sqrt{6}\frac{\sqrt{\xi^2-x^2}}{1+\xi}\int d\bar{x}_1 d\bar{x}_2 d\bar{x}_3\int\frac{d^2\bar{\boldsymbol{k}}_{\perp 1}d^2\bar{\boldsymbol{k}}_{\perp 2}d^2\bar{\boldsymbol{k}}_{\perp 3}}{(16\pi^3)^2}$$
$$\times\delta(\bar{x}-\bar{x}_3)\delta(1-\xi-\bar{x}_1-\bar{x}_2)\delta^{(2)}\left(\bar{\boldsymbol{k}}_{\perp 1}+\bar{\boldsymbol{k}}_{\perp 2}-\frac{\boldsymbol{\Delta}_\perp}{2}\right)$$
$$\times\psi^{(1)*}_{u\bar{d}}(1^{\rm out},2^{\rm out})\left[\psi^{(1)}_{u\bar{d}gg}(1^{\rm in},2^{\rm in},3^{\rm in},3'^{\rm in})+\psi^{(1)}_{u\bar{d}gg}(1^{\rm in},2^{\rm in},3'^{\rm in},3^{\rm in})\right]. \tag{6.55}$$

In Eqs. (6.52) - (6.55) the arguments of the LFWAs of the spectator partons are labelled with the same numbers in the initial and final state. The label j of the active quark is identified by the function $\delta(\bar{x}-\bar{x}_j)$, and the remaining index is automatically assigned to the active parton j'. Furthermore, the LFWA arguments of the spectator partons $i\neq j,j'$ are given as before in (6.35) and (6.37), whereas for the partons removed from the target they now read

$$x_j^{\rm in}=\frac{\bar{x}_j+\xi}{1+\xi},\qquad \boldsymbol{k}_{\perp j}^{\rm in}=\bar{\boldsymbol{k}}_{\perp j}-\frac{1-\bar{x}_j}{1+\xi}\frac{\boldsymbol{\Delta}_\perp}{2}, \tag{6.56}$$
$$x_{j'}^{\rm in}=-\frac{\bar{x}_j-\xi}{1+\xi},\qquad \boldsymbol{k}_{\perp j'}^{\rm in}=-\bar{\boldsymbol{k}}_{\perp j}-\frac{1+\bar{x}_j}{1+\xi}\frac{\boldsymbol{\Delta}_\perp}{2}, \tag{6.57}$$

and this time the auxiliary variables for the active partons, $\bar{x}_j$ and $\boldsymbol{k}_{\perp j}$, are

$$\bar{x}_j=\frac{\bar{k}_j^+}{P^+},\qquad \bar{k}_j=\frac{1}{2}\left(k_j-k_{j'}\right), \tag{6.58}$$

where $\bar{k}_j$ can be interpreted as the average between the momentum of the active particle, k_j, and the momentum of the active antiparticle, $-k_{j'}$. The

relations in Eq. (6.39) are not valid in the ERBL region and are replaced by

$$\sum_{i\neq j,j'} \bar{x}_i = 1-\xi, \qquad \sum_{i\neq j,j'} \bar{\boldsymbol{k}}_{\perp i} = \frac{\boldsymbol{\Delta}_\perp}{2}. \tag{6.59}$$

6.4 Refined parametrization

The charge-conjugation relation in Eq. (6.17) imposes certain constraints in the parametrization of the LFWAs which were not required in the case of the PDFs and TMDs. In particular, it turns out that the LFWAs with 4 partons like $q\bar{q}'s\bar{s}$ and $q\bar{q}'gg$ cannot possess complete symmetry under the exchange of the quantum numbers of the $q \leftrightarrow \bar{q}'$, $s \leftrightarrow \bar{s}$ or gg pairs. We therefore needed to revisit the parametrization of the LFWAs to ensure the correct symmetry properties while preserving the assumption in Eq. (2.159) for the asymptotic expansion of the x-dependence of the leading-twist DAs. As a result, the new parametrization reads

$$\Phi^{(1)}_{q\bar{q}'gg}(x_1,x_2,x_3,x_4) = \phi^{(1)}_{q\bar{q}'gg}(x_1,x_2,x_3,x_4), \tag{6.60}$$

$$\Phi^{(2)}_{q\bar{q}'gg}(x_1,x_2,x_3,x_4) = \frac{N^{(2)}_{q\bar{q}'gg}}{\sqrt{2}}\sqrt{x_1x_2}x_3x_4(x_1-x_2)(x_3^2-x_4^2), \tag{6.61}$$

$$\Phi^{(1)}_{q\bar{q}'s\bar{s}}(x_1,x_2,x_3,x_4) = \frac{N^{(1)}_{q\bar{q}'s\bar{s}}}{\sqrt{2}}x_1x_2\sqrt{x_3x_4}(x_3-x_4), \tag{6.62}$$

$$\Phi^{(2)}_{q\bar{q}'s\bar{s}}(x_1,x_2,x_3,x_4) = \frac{N^{(2)}_{q\bar{q}'s\bar{s}}}{\sqrt{2}}\sqrt{x_1x_2}x_3x_4(x_1-x_2), \tag{6.63}$$

$$\Phi^{(3)}_{q\bar{q}'s\bar{s}}(x_1,x_2,x_3,x_4) = \phi^{(3)}_{q\bar{q}'s\bar{s}}(x_1,x_2,x_3,x_4). \tag{6.64}$$

The new parametrization maintains the same normalization factors of the original one, hence the relative weight of the different Fock components remains unchanged. Furthermore, it reproduces the same results for the first Mellin moments of the PDF (see Tab. 3.2). While the shapes of the valence and total sea PDF slightly change, the gluon PDF remains unchanged. Due to these minor adjustments, for this preliminary exploration, we have opted to use the same model parameters as the original fit, avoiding a new fitting procedure for the PDFs, TMDs and e.m. form factor. Specifically, we will employ the parameter values corresponding to the zeroth replica of both the TMD and e.m. FF fits. We postpone to future work the update of the fit, with the inclusion of the error bands.

6.5 Model results

In this Section we present the plots relative to the up, strange and gluon GPDs computed with the new model of Section 6.4 and by using as inputs the

collinear set $\mathcal{A}^L$ from the fit of the pion PDFs with initial scale 1 GeV. For the transverse parameters, we use both the sets $\mathcal{A}^T_{FF}$ (5.29) and $\mathcal{A}^T_{TMD}$ (5.32) and we compare the predictions.

The GPDs of the quarks with the other flavours can be derived from the three plotted GPDs by applying the relations reported in Section 6.2.

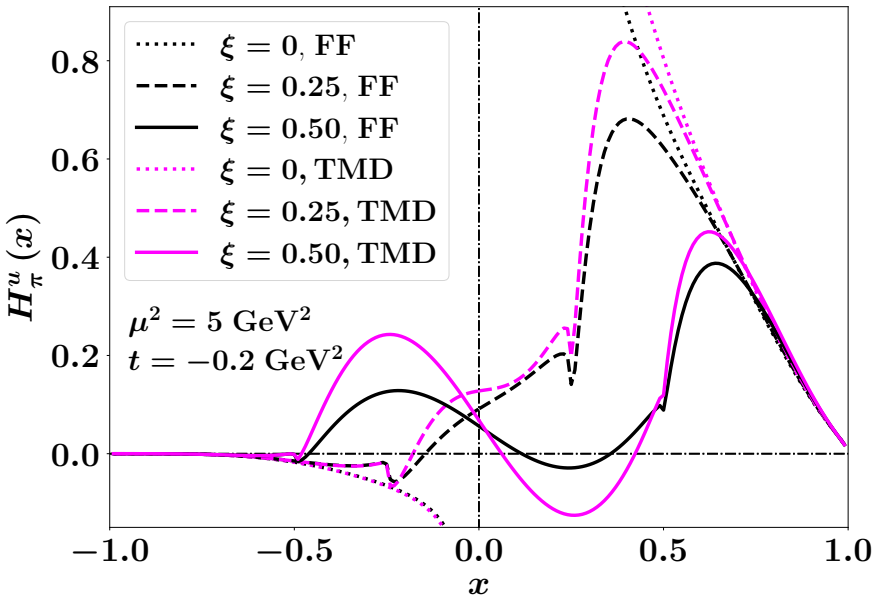
The parton distributions are plotted at $\mu^2 = 5$ GeV^2 as functions of $\bar{x}$ for three different values of ξ and for two different values of t: $t = -0.2$ GeV^2, in Fig. 6.4, and $t = -0.5$ GeV^2, in Fig. 6.5.

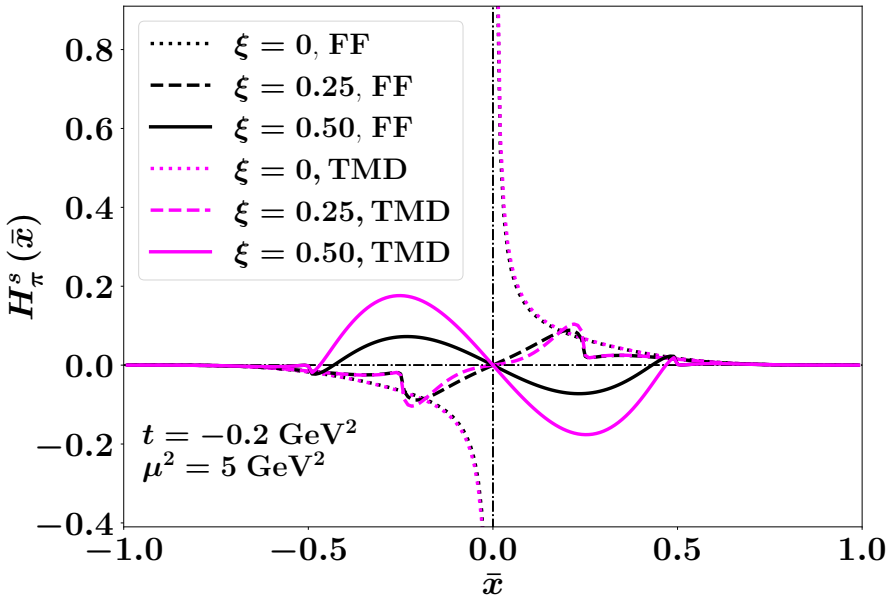
The evolution from the initial scale to the final scale is performed at leading-order[1] with the open source evolution library `APFEL++` [280–282]. The evolution equations for the GPDs are more complicated than for the PDFs, due to the presence of the three regions in $\bar{x}$. However, it must hold that in the forward limit $\xi \to 0$, the GPD evolution reduces to the PDF evolution.

Upon examining Figs. 6.4 and 6.5 we observe that the overall qualitative behaviour of the curves obtained from the two sets of transverse parameters is similar for all the plotted GPDs at different values of ξ and in the whole range of $\bar{x}$. However, there are regions where the predictions with $\mathcal{A}^T_{FF}$ significantly differ from the prediction with $\mathcal{A}^T_{TMD}$. The most notable discrepancies are associated to the ERBL regions, becoming more pronounced as ξ increases. In the DGLAP1 and DGLAP2 the agreement is remarkably good, except for the up quark in DGLAP1, where the two predictions are considerably different.

We stress again that the parameters used to generate these plots are the best fitted sets for the model of Section 4.3, which does not completely correspond to the refined model given in Eqs. (6.60) - (6.64). The comparison is therefore not conclusive, until we apply again the whole fitting procedure with the new parametrization. This work is planned in the near future. Moreover, we are also planning to refine this very qualitative comparison by including the error bands of the entire sets of the replicas. Hopefully, within the uncertainties, the results from the different fits might result compatible for all the values of $\bar{x}$, t and ξ, even if the central replicas differ significantly. This would confirm the effectiveness of the approach used throughout this work.

[1] GPD evolution at next-to-leading-order accuracy has been implemented in Ref. [279]. However, the code is not fully open-source and we were not able to obtain it.


$\xi = 0$, FF
$\xi = 0.25$, FF
$\xi = 0.50$, FF
$\xi = 0$, TMD
$\xi = 0.25$, TMD
$\xi = 0.50$, TMD
$\mu^2 = 5$ GeV^2
$t = -0.2$ GeV^2
$H^u_\pi(x)$
x


$\xi = 0$, FF
$\xi = 0.25$, FF
$\xi = 0.50$, FF
$\xi = 0$, TMD
$\xi = 0.25$, TMD
$\xi = 0.50$, TMD
$t = -0.2$ GeV^2
$\mu^2 = 5$ GeV^2
$H^s_\pi(\bar{x})$
$\bar{x}$

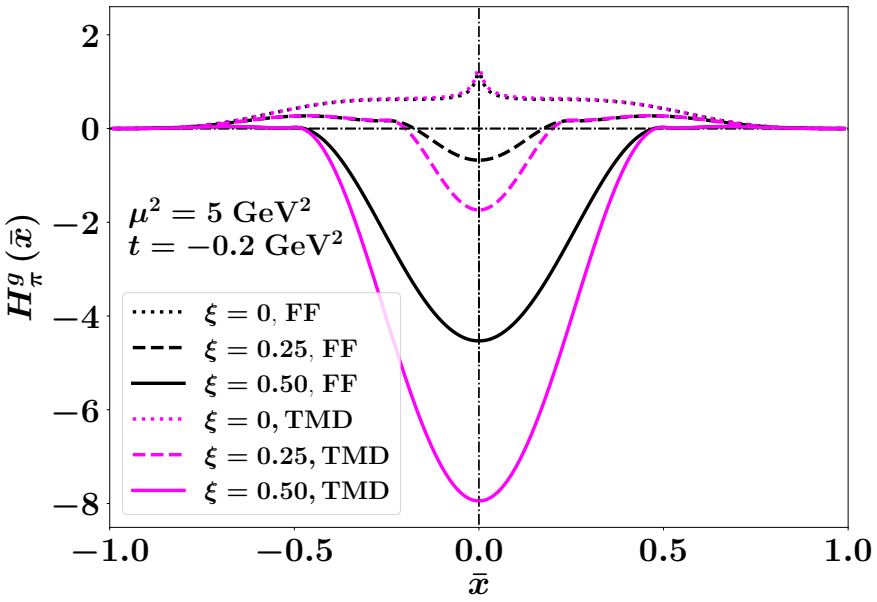
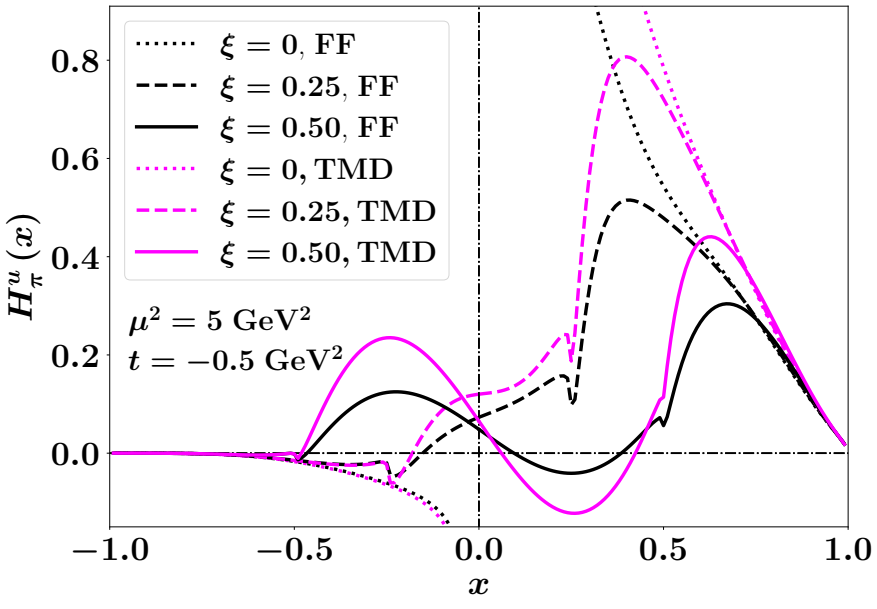


Figure 6.4: Comparison between the quark up (upper panel), strange quark (middle panel) and gluon GPDs (lower panel) with the parameters of the replica 0 from the fit of the TMDs (magenta curves) and from the fit of the FFs (black curves). The results are shown at $\mu^2 = 5$ GeV^2 for $t = -0.2$ GeV^2 and three different values of ξ: $\xi = 0$ (solid lines), $\xi = 0.25$ (dashed lines) and $\xi = 0.50$ (dotted lines).



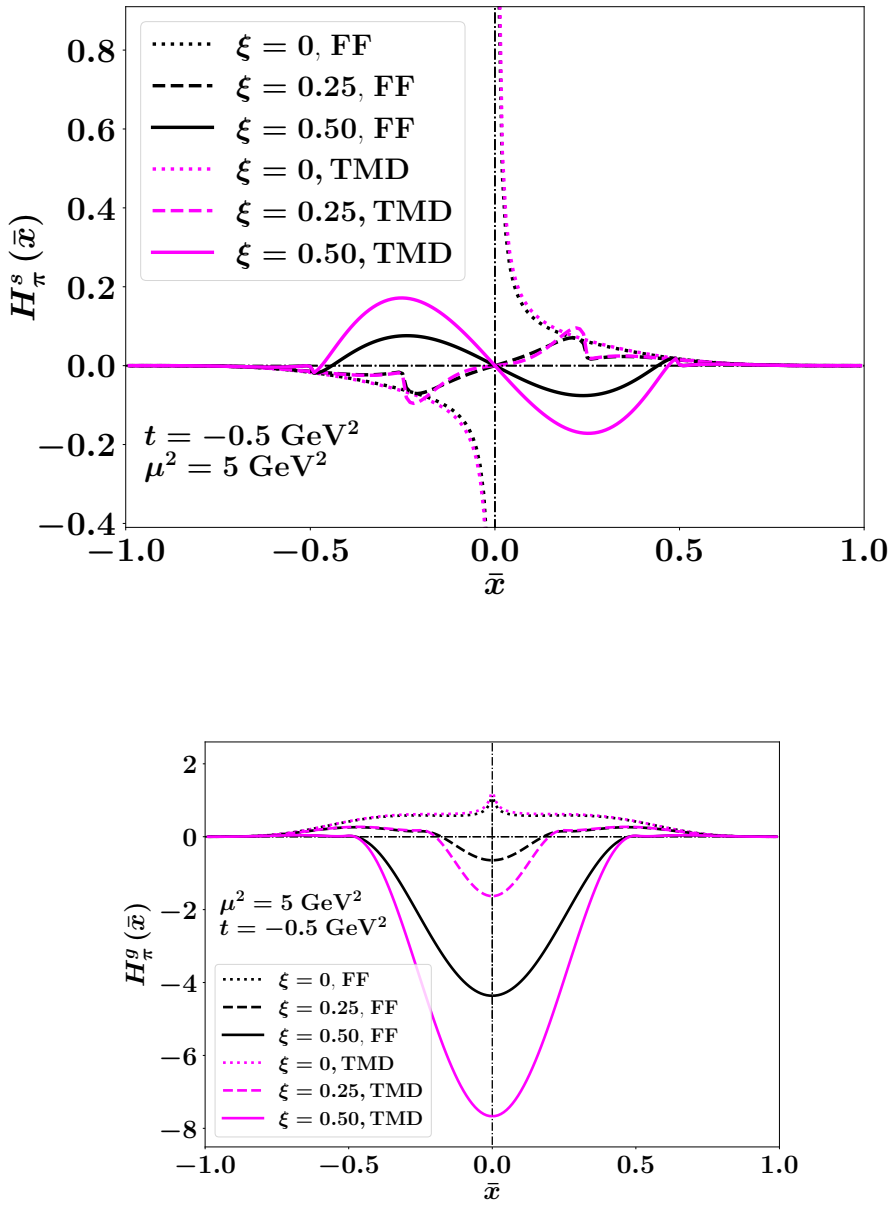


Figure 6.5: Comparison between the quark up (upper panel), strange quark (middle panel) and gluon GPDs (lower panel) with the parameters of the replica 0 from the fit of the TMDs (magenta curves) and from the fit of the FFs (black curves). The results are shown at 5 GeV^2 for $t = -0.5$ GeV^2 and three different values of ξ: $\xi = 0$ (solid lines), $\xi = 0.25$ (dashed lines) and $\xi = 0.50$ (dotted lines).

Chapter 7

Conclusions

The primary focus of this thesis was to investigate the internal structure of the pion using parton distribution functions as a tool. The intention was to approach such a task step by step, presenting the topics as an authentic *journey in phenomenology*.

In Ch. 2 all the theoretical ingredients have been systematically presented. The journey started from the very basic concepts underlying quantum field theory, examining the geometrical properties of the light-cone and introducing the light-front coordinates and light cone operators. Then, the focus shifted on a specific quantum field theory, QCD, delineating the properties and the quantum fields conceptualized as fundamental quantities. We showed the efficacy of the light front formalism in describing hadrons in terms of its constituent partons. This involves decomposing colourless states into a Fock-state expansion, with the emergence of the LFWFs and LFWAs. The hadron states enter into the definition of the most general correlator, from which all the partonic distributions originate. This is the starting point to parametrize the parton distribution functions in terms of LFWF-based models. Thus, we proposed our theoretical parametrization of the LFWFs based on two sets of parameters to be independently fixed through two consecutive fitting procedures for different distribution functions. The two sets are referred to as *collinear* and *transverse* parameters. This is due to the fact that the parametrization is built in such a way that the dependence on the transverse parameters is integrated out in the collinear PDFs and therefore the PDFs involve exclusively the longitudinal parameters. On the other hand, all the other parton distributions depend on both collinear and transverse parameters. Hence, the approach follows this logic: an initial fitting of the collinear distribution functions determines the collinear set, which is then utilized as input for the parton distribution functions that are dependent on the transverse momentum. A subsequent fitting of the transverse set is carried out accordingly. The decision regarding the number of Fock components to incorporate in the decomposition of the pion state was driven by the aim to effectively parametrize the valence, gluon, and sea at the initial scale. To the best of our knowledge, our framework is the first

approach among the existing light-front model calculations which incorporates both sea and gluon components. Another very interesting aspect of the theoretical parametrization is its strong correlation with the pion DAs, explicitly formulated at the leading-twist level.

We implemented our strategy to analyze various types of parton distributions by gradually increasing the level of complexity.

The initial focus was on examining the pion collinear parton distribution functions (PDFs), presented in Chapter 3. This involved conducting fits of the available Drell-Yan (DY) and prompt photon production data leading to the determination of the optimal values of the collinear set of parameters. The results were then compared with other existing extractions and the behaviour of the valence PDF in the limit $x \to 1$ was studied.

Ch. 4 was devoted to fitting the existing measurements of the pion e.m. form factor (FF), using the second set of the model parameters.

After the analysis of the 1-dimensional PDFs we delved into the more complicated 3-dimensional pion transverse-momentum dependent parton distributions (TMDs) in Ch. 5. This exploration included a brief introduction to the DY formalism and to the framework of Nanga Parbat, an open source QCD software developed in Pavia for fitting hadron TMDs. The available DY data were fitted within two different models: a pure phenomenological model and the LFWF-based model developed in this work. The fit of the TMDs provided, like the FFs, the best fit parameters of the transverse set. The results obtained from the phenomenological fits and the LFWA-based approach have been thoroughly compared and rigorously analyzed. The two models for pion TMDs have also been used to obtain predictions for future experimental measurements.

The last part of the thesis was devoted to provide predictions for the pion generalized parton distributions (GPDs), which represent very challenging distributions both to extract from experimental data and to calculate theoretically.

From a purely theoretical perspective, the main goal of the thesis was to develop a model based on the LFWF approach enabling to access to complementary information regarding the internal structure of the pion. The model proposed in Ch. 2 made significant strides in achieving this goal. It successfully contributed to fitting pion PDFs in Ch. 3, yielding remarkable result. Additionally, the model effectively explained the experimental data for pion FFs in Ch. 4 and appropriately described measurements of pion TMDs in Ch. 5. However, it did not incorporate in the LFWFs the constraints imposed from the symmetry properties of the GPDs. We then proposed a refined version of the original model in Chapter 6 to address the GPD symmetry property issue. The new parametrization kept the relative weight of the different Fock components of the pion state unaltered. Furthermore, it reproduced the results for the first Mellin moments of the PDF at the initial scale. While the shapes of the valence and total sea PDF slightly changed w.r.t. the original parametrization,

the gluon PDF remained unchanged. Considering these minor adjustments, for our initial exploration, we choose to maintain the same model parameters as the original fit, bypassing the need for a new fitting procedure for the PDFs, TMDs, and e.m. FF. The update of the fit has been deferred to future work.

From a phenomenological perspective, it firmly emerged the necessity of new future experimental data for observables sensitive to all the parton distribution functions of the pion studied in this work. For the PDFs, new cross sections are crucial to better constrain the gluon parameters, to discriminate between different independent components of the Fock states with 4 partons and to shed light on the behaviour of the valence PDF at high x. Regarding the TMDs, the current DY cross sections are characterized by large systematic uncertainties which contribute to the high values of the penalty chi-squared in the fits. Therefore, collecting new experimental data is essential to increase the quality of the fits. In the case of the GPDs, all the existing information is derived from theoretical model calculations since no experimental data is currently available. It is necessary to validate these theoretical predictions through direct comparison with experimental measurements and then to aim to a global fit with PDFs and TMDs. For the e.m. FF, new measurements are needed at higher values of Q^2, for the reconstruction of the pion transverse charge radius as discussed in Refs. [283–288].

In conclusion, it can be asserted that the efforts in this work represent a profound exploration towards the core of pions. We have delved into the internal structure of pions, mapping multi-dimensional landscapes that have provided substantial insights and guided us throughout our journey. Our compass on this journey has been the fundamental essence of scientific advancements, namely the rigorous comparison between theoretical models and experimental data. It is important to highlight that we have not yet achieved a clear and definitive depiction to fully *revealing the internal structure of the pion*; consequently, the *journey in phenomenology* must continue.

7.1 Outlook

What are the next steps in this journey in phenomenology?

One future exploration involves to repeating the fitting procedure within the refined parametrization introduced in Chapter 6. This refinement necessitates determining two new sets of parameters, which can be obtained by conducting fits for either both PDFs and FFs, or alternatively, fitting PDFs and TMDs. A very interesting compatibility cross check can be performed by comparing the uncertainty bands from both these methods in the predictions of pion GPDs. Beyond comparing the two parametrizations, it is valuable to consider a comparison of the GPDs with their uncertainty bands with various other theoretical models to study their compatibility. Moreover, predicting the values of potential observables that can be measured in exclusive processes connected to pion GPDs presents an intriguing aspect to explore.

Another important avenue for future improvement involves the potential for conducting simultaneous fits of PDFs, TMDs, and FFs. We have observed that the optimal parameters for the fit of FFs do not adequately explain the TMD data, and vice versa. This issue arises from the distinct kinematics inherent in the involved different physical processes. We expect that this challenge can be addressed by implementing a global fitting approach. In such an approach, the best-fit parameters would be determined by minimizing a common chi-squared function that incorporates all experimental data associated with the various partonic distributions.

As repeatedly emphasized within this work, another crucial future development concerns the experimental aspect, where new measurements related to the internal structure of pions are essential. These measurements are necessary across all kinematic regimes, for both inclusive and exclusive processes. The new data can play a multifaceted role: better constraining existing theoretical fits, enhancing sensitivity to specific parameters, shedding light on still debated theoretical issues, and ultimately, facilitating the realization of new future fits. This is a genuine appeal to the hadronic physics community, encouraging active engagement in carrying out measurements for the pion in the future.

Bibliography

[1] A. Aprahamian et al., "Reaching for the horizon: The 2015 long range plan for nuclear science", .

[2] E. Rutherford, "Collision of α particles with light atoms. IV. An anomalous effect in nitrogen", *Phil. Mag. Ser. 6* **37** (1919) 581.

[3] J. Chadwick, "Possible Existence of a Neutron", *Nature* **129** (1932) 312.

[4] European Muon collaboration, J. Ashman et al., "A Measurement of the Spin Asymmetry and Determination of the Structure Function g(1) in Deep Inelastic Muon-Proton Scattering", *Phys. Lett. B* **206** (1988) 364.

[5] C. M. G. Lattes, H. Muirhead, G. P. S. Occhialini and C. F. Powell, "PROCESSES INVOLVING CHARGED MESONS", *Nature* **159** (1947) 694.

[6] J. Steinberger, W. K. H. Panofsky and J. Steller, "EVIDENCE FOR THE PRODUCTION OF NEUTRAL MESONS BY PHOTONS", *Phys. Rev.* **78** (1950) 802.

[7] P. A. M. Dirac, "Forms of relativistic dynamics", *Rev. Mod. Phys.* **21** (1949) 392.

[8] J. B. Kogut and D. E. Soper, "Quantum Electrodynamics in the Infinite Momentum Frame", *Phys. Rev. D* **1** (1970) 2901.

[9] G. Leibbrandt, "The Light Cone Gauge in Yang-Mills Theory", *Phys. Rev. D* **29** (1984) 1699.

[10] D. Hwang, "Light cone representation of the spin and orbital angular momentum of relativistic composite systems", *Nucl. Phys. B Proc. Suppl.* **105** (2002) 160.

[11] D. Soper, "Infinite-momentum helicity states", *Phys. Rev. D* **5** (1972) 1956.

[12] B. D. Keister and W. N. Polyzou, "Relativistic Hamiltonian dynamics in nuclear and particle physics", *Adv. Nucl. Phys.* **20** (1991) 225.

[13] M. Burkardt, "Light front quantization", *Adv. Nucl. Phys.* **23** (1996) 1 [hep-ph/9505259].

[14] J.-H. Gao, "Singularities, boundary conditions and gauge link in the light cone gauge", *Phys. Part. Nucl.* **45** (2014) 704 [1309.4970].

[15] R. L. Jaffe, "Spin, twist and hadron structure in deep inelastic processes", in *Ettore Majorana International School of Nucleon Structure: 1st Course: The Spin Structure of the Nucleon*, pp. 42–129, 1, 1996, hep-ph/9602236.

[16] J. Collins, *Foundations of Perturbative QCD*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2011, 10.1017/CBO9780511975592.

[17] S. Meißner, K. Goeke, A. Metz and M. Schlegel, "Generalized parton correlation functions for a spin-0 hadron", *Journal of High Energy Physics* **2008** (2008) 038–038.

[18] S. Meißner, A. Metz and M. Schlegel, "Generalized parton correlation functions for a spin-1/2 hadron", *Journal of High Energy Physics* **2009** (2009) 056–056.

[19] M. Diehl, "Introduction to gpds and tmds", *The European Physical Journal A* **52** (2016) .

[20] C. Lorcé, B. Pasquini and M. Vanderhaeghen, "Unified framework for generalized and transverse-momentum dependent parton distributions within a 3q light-cone picture of the nucleon", *Journal of High Energy Physics* **2011** (2011) .

[21] R. L. Jaffe, H. Meyer and G. Piller, "Spin, twist and hadron structure in deep inelastic processes", *Lect. Notes Phys.* **496** (1997) 178.

[22] A. V. Belitsky, X. Ji and F. Yuan, "Final state interactions and gauge invariant parton distributions", *Nucl. Phys. B* **656** (2003) 165 [hep-ph/0208038].

[23] S. Meissner, A. Metz, M. Schlegel and K. Goeke, "Generalized parton correlation functions for a spin-0 hadron", *JHEP* **08** (2008) 038 [0805.3165].

[24] C. Lorcé and B. Pasquini, "Structure analysis of the generalized correlator of quark and gluon for a spin-1/2 target", *JHEP* **09** (2013) 138 [1307.4497].

[25] K. Goeke, A. Metz, P. V. Pobylitsa and M. V. Polyakov, "Lorentz invariance relations among parton distributions revisited", *Phys. Lett. B* **567** (2003) 27 [hep-ph/0302028].

[26] X.-d. Ji and F. Yuan, "Parton distributions in light cone gauge: Where are the final state interactions?", *Phys. Lett. B* **543** (2002) 66 [hep-ph/0206057].

[27] S. J. Brodsky, B. Pasquini, B.-W. Xiao and F. Yuan, "Phases of Augmented Hadronic Light-Front Wave Functions", *Phys. Lett. B* **687** (2010) 327 [1001.1163].

[28] M. Burkardt, "Impact parameter space interpretation for generalized parton distributions", *Int. J. Mod. Phys. A* **18** (2003) 173 [hep-ph/0207047].

[29] M. Diehl, "Generalized parton distributions in impact parameter space", *Eur. Phys. J. C* **25** (2002) 223 [hep-ph/0205208].

[30] M. Diehl, "Introduction to GPDs and TMDs", *Eur. Phys. J. A* **52** (2016) 149 [1512.01328].

[31] A. V. Belitsky, X.-d. Ji and F. Yuan, "Quark imaging in the proton via quantum phase space distributions", *Phys. Rev. D* **69** (2004) 074014 [hep-ph/0307383].

[32] X.-d. Ji, "Viewing the proton through 'color' filters", *Phys. Rev. Lett.* **91** (2003) 062001 [hep-ph/0304037].

[33] C. Lorce and B. Pasquini, "Quark Wigner Distributions and Orbital Angular Momentum", *Phys. Rev. D* **84** (2011) 014015 [1106.0139].

[34] C. Lorce, B. Pasquini, X. Xiong and F. Yuan, "The quark orbital angular momentum from Wigner distributions and light-cone wave functions", *Phys. Rev. D* **85** (2012) 114006 [1111.4827].

[35] N. Y. Lee, P. V. Pobylitsa, M. V. Polyakov and K. Goeke, "Meson twist four parton distributions in terms of twist two distribution amplitudes at large N(c)", *J. Phys. G* **27** (2001) L127 [hep-ph/0011166].

[36] D. Mueller, "The Evolution of the pion distribution amplitude in next-to-leading-order", *Phys. Rev. D* **51** (1995) 3855 [hep-ph/9411338].

[37] P. Ball, "Theoretical update of pseudoscalar meson distribution amplitudes of higher twist: The Nonsinglet case", *JHEP* **01** (1999) 010 [hep-ph/9812375].

[38] Lattice Parton collaboration, J. Hua et al., "Pion and Kaon Distribution Amplitudes from Lattice QCD", *Phys. Rev. Lett.* **129** (2022) 132001 [2201.09173].

[39] Lattice Parton collaboration, J. Hua, M.-H. Chu, P. Sun, W. Wang, J. Xu, Y.-B. Yang et al., "Distribution Amplitudes of K* and ϕ at the Physical Pion Mass from Lattice QCD", *Phys. Rev. Lett.* **127** (2021) 062002 [2011.09788].

[40] P. Ball, V. M. Braun and A. Lenz, "Twist-4 distribution amplitudes of the K* and phi mesons in QCD", *JHEP* **08** (2007) 090 [0707.1201].

[41] H.-M. Choi and C.-R. Ji, "Distribution amplitudes and decay constants for (pi, K, rho, K*) mesons in light-front quark model", *Phys. Rev. D* **75** (2007) 034019 [hep-ph/0701177].

[42] H.-M. Choi and C.-R. Ji, "Two-particle twist-3 distribution amplitudes of the pion and kaon in the light-front quark model", *Phys. Rev. D* **95** (2017) 056002 [1701.02402].

[43] V. M. Braun and I. E. Filyanov, "Conformal Invariance and Pion Wave Functions of Nonleading Twist", *Z. Phys. C* **48** (1990) 239.

[44] X.-d. Ji, J.-P. Ma and F. Yuan, "Classification and asymptotic scaling of hadrons' light cone wave function amplitudes", *Eur. Phys. J. C* **33** (2004) 75 [hep-ph/0304107].

[45] X.-d. Ji, J.-P. Ma and F. Yuan, "Generalized counting rule for hard exclusive processes", *Phys. Rev. Lett.* **90** (2003) 241601 [hep-ph/0301141].

[46] S. J. Brodsky, T. Huang and P. Lepage, *in Particle and Fields*. edited by Capri, A.Z. and Kamal, A.N. (Plenum, New York), 1983.

[47] L. Chang, I. C. Cloët, J. J. Cobos-Martinez, C. D. Roberts, S. M. Schmidt and P. C. Tandy, "Imaging dynamical chiral-symmetry breaking: Pion wave function on the light front", *Phys. Rev. Lett.* **110** (2013) 132001.

[48] M. Diehl, T. Feldmann, R. Jakob and P. Kroll, "The overlap representation of skewed quark and gluon distributions", *Nucl. Phys. B* **596** (2001) 33 [hep-ph/0009255].

[49] D. S. Hwang and D. Mueller, "Implication of the overlap representation for modelling generalized parton distributions", *Phys. Lett. B* **660** (2008) 350 [0710.1567].

[50] N. Chouika, C. Mezrag, H. Moutarde and J. Rodríguez-Quintero, "Covariant Extension of the GPD overlap representation at low Fock states", *Eur. Phys. J. C* **77** (2017) 906 [1711.05108].

[51] N. Chouika, C. Mezrag, H. Moutarde and J. Rodríguez-Quintero, "A Nakanishi-based model illustrating the covariant extension of the pion GPD overlap representation and its ambiguities", *Phys. Lett. B* **780** (2018) 287 [1711.11548].

[52] S. Alekhin et al., "HERAFitter", *Eur. Phys. J. C* **75** (2015) 304 [1410.4412].

[53] I. Novikov et al., "Parton Distribution Functions of the Charged Pion Within The xFitter Framework", *Phys. Rev. D* **102** (2020) 014040 [2002.02902].

[54] NA10 collaboration, B. Betev et al., "Differential Cross-section of High Mass Muon Pairs Produced by a 194-GeV/$c\pi^-$ Beam on a Tungsten Target", *Z. Phys. C* **28** (1985) 9.

[55] J. S. Conway et al., "Experimental Study of Muon Pairs Produced by 252-GeV Pions on Tungsten", *Phys. Rev. D* **39** (1989) 92.

[56] WA70 collaboration, M. Bonesini et al., "High Transverse Momentum Prompt Photon Production by π^- and π^+ on Protons at 280-GeV/c", *Z. Phys. C* **37** (1988) 535.

[57] E598 collaboration, J. J. Aubert et al., "Experimental Observation of a Heavy Particle J", *Phys. Rev. Lett.* **33** (1974) 1404.

[58] G. S. Abrams et al., "The Discovery of a Second Narrow Resonance in e+ e- Annihilation", *Phys. Rev. Lett.* **33** (1974) 1453.

[59] E288 collaboration, S. W. Herb et al., "Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions", *Phys. Rev. Lett.* **39** (1977) 252.

[60] R. D. Ball, S. Carrazza, L. D. Debbio, S. Forte, J. Gao, N. Hartland et al., "Parton distribution benchmarking with LHC data", *Journal of High Energy Physics* **2013** (2013) .

[61] F. James and M. Roos, "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations", *Comput. Phys. Commun.* **10** (1975) 343.

[62] J. Arrington et al., "Physics with CEBAF at 12 GeV and future opportunities", *Prog. Part. Nucl. Phys.* **127** (2022) 103985 [2112.00060].

[63] R. Abdul Khalek et al., "Science Requirements and Detector Concepts for the Electron-Ion Collider: EIC Yellow Report", *Nucl. Phys. A* **1026** (2022) 122447 [2103.05419].

[64] D. P. Anderle et al., "Electron-ion collider in China", *Front. Phys. (Beijing)* **16** (2021) 64701 [2102.09222].

[65] J. D. Sullivan, "One pion exchange and deep inelastic electron - nucleon scattering", *Phys. Rev. D* **5** (1972) 1732.

[66] A. C. Aguilar et al., "Pion and Kaon Structure at the Electron-Ion Collider", *Eur. Phys. J. A* **55** (2019) 190 [1907.08218].

[67] J. Arrington et al., "Revealing the structure of light pseudoscalar mesons at the electron–ion collider", *J. Phys. G* **48** (2021) 075106 [2102.11788].

[68] B. Adams et al., "Letter of Intent: A New QCD facility at the M2 beam line of the CERN SPS (COMPASS++/AMBER)", 8, 2018.

[69] M. Gluck, E. Reya and A. Vogt, "Pionic parton distributions", *Z. Phys. C* **53** (1992) 651.

[70] Jefferson Lab Angular Momentum (JAM) collaboration, P. C. Barry, C.-R. Ji, N. Sato and W. Melnitchouk, "Global QCD Analysis of Pion Parton Distributions with Threshold Resummation", *Phys. Rev. Lett.* **127** (2021) 232001 [2108.05822].

[71] C. Bourrely, W.-C. Chang and J.-C. Peng, "Pion Partonic Distributions in the Statistical Model from Pion-induced Drell-Yan and J/Ψ Production Data", *Phys. Rev. D* **105** (2022) 076018 [2202.12547].

[72] C. Bourrely, F. Buccella and J.-C. Peng, "A new extraction of pion parton distributions in the statistical model", *Phys. Lett. B* **813** (2021) 136021 [2008.05703].

[73] G. Sterman, "Summation of large corrections to short-distance hadronic cross sections", *Nuclear Physics B* **281** (1987) 310.

[74] S. Catani and L. Trentadue, "Resummation of the qcd perturbative series for hard processes", *Nuclear Physics B* **327** (1989) 323.

[75] C. W. Bauer, S. Fleming, D. Pirjol, I. Z. Rothstein and I. W. Stewart, "Hard scattering factorization from effective field theory", *Phys. Rev. D* **66** (2002) 014017.

[76] A. V. Manohar, "Deep inelastic scattering as $\vec{x}$ 1 using soft-collinear effective theory", *Phys. Rev. D* **68** (2003) 114019.

[77] J. Chay and C. Kim, "Deep inelastic scattering near the endpoint in soft-collinear effective theory", *Phys. Rev. D* **75** (2007) 016003.

[78] M. Bonvini, S. Forte, M. Ghezzi and G. Ridolfi, "Threshold resummation in scet vs. perturbative qcd: An analytic comparison", *Nuclear Physics B* **861** (2012) 337.

[79] T. Becher, M. Neubert and B. D. Pecjak, "Factorization and momentum-space resummation in deep-inelastic scattering", *Journal of High Energy Physics* **2007** (2007) 076.

[80] A. Idilbi, X. Ji and F. Yuan, "Resummation of threshold logarithms in effective field theory for dis, drell–yan and higgs production", *Nuclear Physics B* **753** (2006) 42.

[81] B. D. Pecjak, "Non-factorizable contributions to deep inelastic scattering at large x", *Journal of High Energy Physics* **2005** (2005) 040.

[82] M. Aicher, A. Schafer and W. Vogelsang, "Soft-gluon resummation and the valence parton distribution function of the pion", *Phys. Rev. Lett.* **105** (2010) 252003 [1009.2481].

[83] B. Pasquini and P. Schweitzer, "Pion transverse momentum dependent parton distributions in a light-front constituent approach, and the Boer-Mulders effect in the pion-induced Drell-Yan process", *Phys. Rev. D* **90** (2014) 014050 [1406.2056].

[84] BLFQ collaboration, J. Lan, K. Fu, C. Mondal, X. Zhao and j. P. Vary, "Light mesons with one dynamical gluon on the light front", *Phys. Lett. B* **825** (2022) 136890 [2106.04954].

[85] M. Ding, K. Raya, D. Binosi, L. Chang, C. D. Roberts and S. M. Schmidt, "Symmetry, symmetry breaking, and pion parton distributions", *Phys. Rev. D* **101** (2020) 054014 [1905.05208].

[86] Z.-F. Cui, M. Ding, F. Gao, K. Raya, D. Binosi, L. Chang et al., "Kaon and pion parton distributions", *Eur. Phys. J. C* **80** (2020) 1064.

[87] W. de Paula, E. Ydrefors, J. H. A. Nogueira, T. Frederico and G. Salmè, "Parton distribution function in a pion with minkowskian dynamics", *Phys. Rev. D* **105** (2022) L071505.

[88] H. B. Meyer and J. W. Negele, "Gluon contributions to the pion mass and light cone momentum fraction", *Phys. Rev. D* **77** (2008) 037501 [0707.3225].

[89] P. E. Shanahan and W. Detmold, "Gluon gravitational form factors of the nucleon and the pion from lattice QCD", *Phys. Rev. D* **99** (2019) 014511 [1810.04626].

[90] Extended Twisted Mass collaboration, C. Alexandrou et al., "Quark and Gluon Momentum Fractions in the Pion from Nf=2+1+1 Lattice QCD", *Phys. Rev. Lett.* **127** (2021) 252001 [2109.10692].

[91] Z. Fan and H.-W. Lin, "Gluon parton distribution of the pion from lattice QCD", *Phys. Lett. B* **823** (2021) 136778 [2104.06372].

[92] ETM Collaboration collaboration, C. Alexandrou, S. Bacchio, I. Cloët, M. Constantinou, K. Hadjiyiannakou, G. Koutsou et al., "Pion and kaon $<x^3>$ from lattice qcd and pdf reconstruction from mellin moments", *Phys. Rev. D* **104** (2021) 054504.

[93] P. C. Barry, N. Sato, W. Melnitchouk and C.-R. Ji, "First Monte Carlo Global QCD Analysis of Pion Parton Distributions", *Phys. Rev. Lett.* **121** (2018) 152001 [1804.01965].

[94] M. Oehm, C. Alexandrou, M. Constantinou, K. Jansen, G. Koutsou, B. Kostrzewa et al., "$\langle x \rangle$ and $\langle x^2 \rangle$ of the pion PDF from lattice QCD with $N_f = 2 + 1 + 1$ dynamical quark flavors", *Phys. Rev. D* **99** (2019) 014508 [1810.09743].

[95] B. Joó, J. Karpie, K. Orginos, A. V. Radyushkin, D. G. Richards, R. S. Sufian et al., "Pion valence structure from Ioffe-time parton pseudodistribution functions", *Phys. Rev. D* **100** (2019) 114512 [1909.08517].

[96] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language", *Nucl. Phys. B* **126** (1977) 298.

[97] W. Melnitchouk, "Quark hadron duality in electron pion scattering", *Eur. Phys. J. A* **17** (2003) 223 [hep-ph/0208258].

[98] G. R. Farrar and D. R. Jackson, "The Pion Form-Factor", *Phys. Rev. Lett.* **43** (1979) 246.

[99] E. L. Berger and S. J. Brodsky, "Quark Structure Functions of Mesons and the Drell-Yan Process", *Phys. Rev. Lett.* **42** (1979) 940.

[100] M. B. Hecht, C. D. Roberts and S. M. Schmidt, "Valence quark distributions in the pion", *Phys. Rev. C* **63** (2001) 025213 [nucl-th/0008049].

[101] Z. F. Ezawa, "Wide-Angle Scattering in Softened Field Theory", *Nuovo Cim. A* **23** (1974) 271.

[102] P. V. Landshoff and J. C. Polkinghorne, "Threshold properties of electroproduction and annihilation", *Nucl. Phys. B* **53** (1973) 473.

[103] J. F. Gunion, S. J. Brodsky and R. Blankenbecler, "Large Angle Scattering and the Interchange Force", *Phys. Rev. D* **8** (1973) 287.

[104] T. Shigetani, K. Suzuki and H. Toki, "Pion structure function in the Nambu and Jona-Lasinio model", *Phys. Lett. B* **308** (1993) 383 [hep-ph/9402286].

[105] A. Szczepaniak, C.-R. Ji and S. R. Cotanch, "Generalized relativistic meson wave function", *Phys. Rev. D* **49** (1994) 3466 [hep-ph/9309284].

[106] R. M. Davidson and E. Ruiz Arriola, "Structure functions of pseudoscalar mesons in the SU(3) NJL model", *Phys. Lett. B* **348** (1995) 163.

[107] D. G. Dumm, S. Noguera, N. N. Scoccola and S. Scopetta, "Pion distribution amplitude and the pion-photon transition form factor in a nonlocal chiral quark model", *Phys. Rev. D* **89** (2014) 054031 [1311.3595].

[108] P. T. P. Hutauruk, I. C. Cloet and A. W. Thomas, "Flavor dependence of the pion and kaon form factors and parton distribution functions", *Phys. Rev. C* **94** (2016) 035201 [1604.02853].

[109] T. J. Hobbs, "Quantifying finite-momentum effects in the quark quasidistribution functions of mesons", *Phys. Rev. D* **97** (2018) 054028 [1708.05463].

[110] K. D. Bednar, I. C. Cloët and P. C. Tandy, "Distinguishing Quarks and Gluons in Pion and Kaon Parton Distribution Functions", *Phys. Rev. Lett.* **124** (2020) 042002 [1811.12310].

[111] HLFHS collaboration, G. F. de Teramond, T. Liu, R. S. Sufian, H. G. Dosch, S. J. Brodsky and A. Deur, "Universality of Generalized Parton Distributions in Light-Front Holographic QCD", *Phys. Rev. Lett.* **120** (2018) 182001 [1801.09154].

[112] J. Lan, C. Mondal, S. Jia, X. Zhao and J. P. Vary, "Parton Distribution Functions from a Light Front Hamiltonian and QCD Evolution for Light Mesons", *Phys. Rev. Lett.* **122** (2019) 172001 [1901.11430].

[113] L. Chang, K. Raya and X. Wang, "Pion Parton Distribution Function in Light-Front Holographic QCD", *Chin. Phys. C* **44** (2020) 114105 [2001.07352].

[114] A. Kock, Y. Liu and I. Zahed, "Pion and kaon parton distributions in the QCD instanton vacuum", *Phys. Rev. D* **102** (2020) 014039 [2004.01595].

[115] L. Chang and C. D. Roberts, "Regarding the Distribution of Glue in the Pion", *Chin. Phys. Lett.* **38** (2021) 081101 [2106.08451].

[116] M. Gluck, E. Reya and I. Schienbein, "Pionic parton distributions revisited", *Eur. Phys. J. C* **10** (1999) 313 [hep-ph/9903288].

[117] P. J. Sutton, A. D. Martin, R. G. Roberts and W. J. Stirling, "Parton distributions for the pion extracted from Drell-Yan and prompt photon experiments", *Phys. Rev. D* **45** (1992) 2349.

[118] C. D. Roberts, D. G. Richards, T. Horn and L. Chang, "Insights into the emergence of mass from studies of pion and kaon structure", *Prog. Part. Nucl. Phys.* **120** (2021) 103883 [2102.01765].

[119] M. Bonvini, S. Forte and G. Ridolfi, "Soft gluon resummation of Drell-Yan rapidity distributions: Theory and phenomenology", *Nucl. Phys. B* **847** (2011) 93 [1009.5691].

[120] D. Westmark and J. F. Owens, "Enhanced threshold resummation formalism for lepton pair production and its effects in the determination of parton distribution functions", *Phys. Rev. D* **95** (2017) 056024 [1701.06716].

[121] G. P. Lepage and S. J. Brodsky, "Exclusive Processes in Quantum Chromodynamics: Evolution Equations for Hadronic Wave Functions and the Form-Factors of Mesons", *Phys. Lett. B* **87** (1979) 359.

[122] A. V. Efremov and A. V. Radyushkin, "Factorization and Asymptotical Behavior of Pion Form-Factor in QCD", *Phys. Lett. B* **94** (1980) 245.

[123] E. Ydrefors, W. de Paula, J. H. A. Nogueira, T. Frederico and G. Salmé, "Pion electromagnetic form factor with Minkowskian dynamics", *Phys. Lett. B* **820** (2021) 136494 [2106.10018].

[124] T. Gutsche, V. E. Lyubovitskij, I. Schmidt and A. Vega, "Pion light-front wave function, parton distribution and the electromagnetic form factor", *J. Phys. G* **42** (2015) 095005 [1410.6424].

[125] T. Frederico, E. Pace, B. Pasquini and G. Salmè, "Pion Generalized Parton Distributions with covariant and Light-front constituent quark models", *Phys. Rev. D* **80** (2009) 054021 [0907.5566].

[126] J. P. B. C. de Melo, T. Frederico, E. Pace and G. Salmè, "Space-like and time-like pion electromagnetic form-factor and Fock state components within the light-front dynamics", *Phys. Rev. D* **73** (2006) 074013 [hep-ph/0508001].

[127] J. P. B. C. de Melo, T. Frederico, E. Pace and G. Salmè, "Electromagnetic form-factor of the pion in the space and time - like regions within the front form dynamics", *Phys. Lett. B* **581** (2004) 75 [hep-ph/0311369].

[128] J. P. B. C. de Melo, T. Frederico, E. Pace and G. Salmè, "Pair term in the electromagnetic current within the front form dynamics: Spin-0 case", *Nucl. Phys. A* **707** (2002) 399 [nucl-th/0205010].

[129] C.-W. Hwang, "A Consistent treatment for pion form-factors in space - like and time - like regions", *Phys. Rev. D* **64** (2001) 034011 [hep-ph/0105016].

[130] B. L. G. Bakker, H.-M. Choi and C.-R. Ji, "Regularizing the fermion loop divergencies in the light front meson currents", *Phys. Rev. D* **63** (2001) 074014 [hep-ph/0008147].

[131] F. Cardarelli, I. L. Grach, I. M. Narodetsky, E. Pace, G. Salmè and S. Simula, "Charge form-factor of pi and K mesons", *Phys. Rev. D* **53** (1996) 6682 [nucl-th/9507038].

[132] F. Cardarelli, E. Pace, G. Salmè and S. Simula, "Nucleon and pion electromagnetic form-factors in a light front constituent quark model", *Phys. Lett. B* **357** (1995) 267 [nucl-th/9507037].

[133] T. Frederico and G. A. Miller, "Null plane phenomenology for the pion decay constant and radius", *Phys. Rev. D* **45** (1992) 4207.

[134] P. L. Chung, F. Coester and W. N. Polyzou, "Charge Form-Factors of Quark Model Pions", *Phys. Lett. B* **205** (1988) 545.

[135] X. Gao, N. Karthik, S. Mukherjee, P. Petreczky, S. Syritsyn and Y. Zhao, "Pion form factor and charge radius from lattice QCD at the physical point", *Phys. Rev. D* **104** (2021) 114515 [2102.06047].

[136] ETM collaboration, C. Alexandrou, S. Bacchio, I. Cloet, M. Constantinou, J. Delmar, K. Hadjiyiannakou et al., "Scalar, vector, and tensor form factors for the pion and kaon from lattice QCD", *Phys. Rev. D* **105** (2022) 054502 [2111.08135].

[137] χQCD collaboration, G. Wang, J. Liang, T. Draper, K.-F. Liu and Y.-B. Yang, "Lattice Calculation of Pion Form Factor with Overlap Fermions", *Phys. Rev. D* **104** (2021) 074502 [2006.05431].

[138] ETM collaboration, C. Alexandrou et al., "Pion vector form factor from lattice QCD at the physical point", *Phys. Rev. D* **97** (2018) 014508 [1710.10401].

[139] JLQCD collaboration, S. Aoki, G. Cossu, X. Feng, S. Hashimoto, T. Kaneko, J. Noaki et al., "Light meson electromagnetic form factors from three-flavor lattice QCD with exact chiral symmetry", *Phys. Rev. D* **93** (2016) 034504 [1510.06470].

[140] J. Koponen, F. Bursa, C. T. H. Davies, R. J. Dowdall and G. P. Lepage, "Size of the pion from full lattice QCD with physical u , d , s and c quarks", *Phys. Rev. D* **93** (2016) 054503 [1511.07382].

[141] H. Fukaya, S. Aoki, S. Hashimoto, T. Kaneko, H. Matsufuru and J. Noaki, "Computation of the electromagnetic pion form factor from lattice QCD in the ϵ regime", *Phys. Rev. D* **90** (2014) 034506 [1405.4077].

[142] B. B. Brandt, A. Jüttner and H. Wittig, "The pion vector form factor from lattice QCD and NNLO chiral perturbation theory", *JHEP* **11** (2013) 034 [1306.2916].

[143] O. H. Nguyen, K.-I. Ishikawa, A. Ukawa and N. Ukita, "Electromagnetic form factor of pion from $N_f = 2 + 1$ dynamical flavor QCD", *JHEP* **04** (2011) 122 [1102.3652].

[144] JLQCD, TWQCD collaboration, S. Aoki et al., "Pion form factors from two-flavor lattice QCD with exact chiral symmetry", *Phys. Rev. D* **80** (2009) 034508 [0905.2465].

[145] ETM collaboration, R. Frezzotti, V. Lubicz and S. Simula, "Electromagnetic form factor of the pion from twisted-mass lattice QCD at N(f) = 2", *Phys. Rev. D* **79** (2009) 074506 [0812.4042].

[146] P. A. Boyle, J. M. Flynn, A. Juttner, C. Kelly, H. P. de Lima, C. M. Maynard et al., "The Pion's electromagnetic form-factor at small momentum transfer in full lattice QCD", *JHEP* **07** (2008) 112 [0804.3971].

[147] QCDSF/UKQCD collaboration, D. Brömmel et al., "The Pion form-factor from lattice QCD with two dynamical flavours", *Eur. Phys. J. C* **51** (2007) 335 [hep-lat/0608021].

[148] Lattice Hadron Physics collaboration, F. D. R. Bonnet, R. G. Edwards, G. T. Fleming, R. Lewis and D. G. Richards, "Lattice computations of the pion form-factor", *Phys. Rev. D* **72** (2005) 054506 [hep-lat/0411028].

[149] X. Feng, Y. Fu and L.-C. Jin, "Lattice QCD calculation of the pion charge radius using a model-independent method", *Phys. Rev. D* **101** (2020) 051502 [1911.04064].

[150] E. B. Dally et al., "Measurement of the π^- Form-factor", *Phys. Rev. D* **24** (1981) 1718.

[151] E. B. Dally et al., "Elastic Scattering Measurement of the Negative Pion Radius", *Phys. Rev. Lett.* **48** (1982) 375.

[152] S. R. Amendolia et al., "A Measurement of the Pion Charge Radius", *Phys. Lett. B* **146** (1984) 116.

[153] NA7 collaboration, S. R. Amendolia et al., "A Measurement of the Space - Like Pion Electromagnetic Form-Factor", *Nucl. Phys. B* **277** (1986) 168.

[154] C. J. Bebek, C. N. Brown, M. Herzlinger, S. D. Holmes, C. A. Lichtenstein, F. M. Pipkin et al., "Measurement of the pion form-factor up to $q^2 = 4$ GeV^2", *Phys. Rev. D* **13** (1976) 25.

[155] C. J. Bebek et al., "Scalar transverse separation for single π^+ electroproduction", *Phys. Rev. Lett.* **37** (1976) 1326.

[156] C. J. Bebek et al., "Electroproduction of single pions at low epsilon and a measurement of the pion form-factor up to $q^2 = 10$ GeV^2", *Phys. Rev. D* **17** (1978) 1693.

[157] P. Brauel, T. Canzler, D. Cords, R. Felst, G. Grindhammer, M. Helm et al., "Electroproduction of $\pi^+ n$, $\pi^- p$ and $K^+\Lambda$, $K^+\Sigma^0$ Final States Above the Resonance Region", *Z. Phys. C* **3** (1979) 101.

[158] H. Ackermann, T. Azemoon, W. Gabriel, H. D. Mertiens, H. D. Reich, G. Specht et al., "Determination of the Longitudinal and the Transverse Part in pi+ Electroproduction", *Nucl. Phys. B* **137** (1978) 294.

[159] Jefferson Lab F(pi) collaboration, J. Volmer et al., "Measurement of the Charged Pion Electromagnetic Form-Factor", *Phys. Rev. Lett.* **86** (2001) 1713 [nucl-ex/0010009].

[160] Jefferson Lab F(pi) collaboration, V. Tadevosyan et al., "Determination of the pion charge form-factor for $Q^2 = 0.60$ GeV^2 - 1.60 GeV^2", *Phys. Rev. C* **75** (2007) 055205 [nucl-ex/0607007].

[161] Jefferson Lab F(pi)-2 collaboration, T. Horn et al., "Determination of the Charged Pion Form Factor at $Q^2 = 1.60$ and 2.45-(GeV/c)^2", *Phys. Rev. Lett.* **97** (2006) 192001 [nucl-ex/0607005].

[162] Jefferson Lab collaboration, H. P. Blok et al., "Charged pion form factor between Q^2=0.60 and 2.45 GeV^2. I. Measurements of the cross section for the ^{1}H$(e, e'\pi^+)n$ reaction", *Phys. Rev. C* **78** (2008) 045202 [0809.3161].

[163] Jefferson Lab collaboration, G. M. Huber et al., "Charged pion form-factor between $Q^2 = 0.60$ GeV^2 and 2.45 GeV^2. II. Determination of, and results for, the pion form-factor", *Phys. Rev. C* **78** (2008) 045203 [0809.3052].

[164] P. Pedroni and S. Sconfietti, "A new Monte Carlo-based fitting method", *J. Phys. G* **47** (2020) 054001 [1909.03885].

[165] L. Chang, I. C. Cloet, J. J. Cobos-Martinez, C. D. Roberts, S. M. Schmidt and P. C. Tandy, "Imaging dynamical chiral symmetry breaking: pion wave function on the light front", *Phys. Rev. Lett.* **110** (2013) 132001 [1301.0324].

[166] Z.-F. Cui, M. Ding, F. Gao, K. Raya, D. Binosi, L. Chang et al., "Higgs modulation of emergent mass as revealed in kaon and pion parton distributions", *Eur. Phys. J. A* **57** (2021) 5 [2006.14075].

[167] Z. F. Cui, M. Ding, J. M. Morgado, K. Raya, D. Binosi, L. Chang et al., "Concerning pion parton distributions", *Eur. Phys. J. A* **58** (2022) 10 [2112.09210].

[168] Z. F. Cui, M. Ding, J. M. Morgado, K. Raya, D. Binosi, L. Chang et al., "Emergence of pion parton distributions", *Phys. Rev. D* **105** (2022) L091502 [2201.00884].

[169] M. Ding, K. Raya, D. Binosi, L. Chang, C. D. Roberts and S. M. Schmidt, "Drawing insights from pion parton distributions", *Chin. Phys. C* **44** (2020) 031002 [1912.07529].

[170] F. A. Ceccopieri, A. Courtoy, S. Noguera and S. Scopetta, "Pion nucleus Drell–Yan process and parton transverse momentum in the pion", *Eur. Phys. J. C* **78** (2018) 644 [1801.07682].

[171] M. Ahmady, C. Mondal and R. Sandapen, "Predicting the light-front holographic TMDs of the pion", *Phys. Rev. D* **100** (2019) 054005 [1907.06561].

[172] N. Kaur and H. Dahiya, "Transverse momentum dependent parton distributions of pion in the light-front holographic model", *Int. J. Mod. Phys. A* **36** (2021) 2150052 [1908.08657].

[173] C. Shi, K. Bednar, I. C. Cloët and A. Freese, "Spatial and Momentum Imaging of the Pion and Kaon", *Phys. Rev. D* **101** (2020) 074014 [2003.03037].

[174] C. Lorcé, B. Pasquini and P. Schweitzer, "Transverse pion structure beyond leading twist in constituent models", *Eur. Phys. J. C* **76** (2016) 415 [1605.00815].

[175] S. Noguera and S. Scopetta, "Pion transverse momentum dependent parton distributions in the Nambu and Jona-Lasinio model", *JHEP* **11** (2015) 102 [1508.01061].

[176] A. Bacchetta, S. Cotogno and B. Pasquini, "The transverse structure of the pion in momentum space inspired by the AdS/QCD correspondence", *Phys. Lett. B* **771** (2017) 546 [1703.07669].

[177] I. Scimemi and A. Vladimirov, "Analysis of vector boson production within TMD factorization", *Eur. Phys. J. C* **78** (2018) 89 [1706.01473].

[178] I. Scimemi and A. Vladimirov, "Non-perturbative structure of semi-inclusive deep-inelastic and Drell-Yan scattering at small transverse momentum", *JHEP* **06** (2020) 137 [1912.06532].

[179] A. Bacchetta, F. Delcarro, C. Pisano, M. Radici and A. Signori, "Extraction of partonic transverse momentum distributions from semi-inclusive deep-inelastic scattering, Drell-Yan and Z-boson production", *JHEP* **06** (2017) 081 [1703.10157].

[180] A. Bacchetta, V. Bertone, C. Bissolotti, G. Bozzi, F. Delcarro, F. Piacenza et al., "Transverse-momentum-dependent parton distributions up to N^3LL from Drell-Yan data", *JHEP* **07** (2020) 117 [1912.07550].

[181] V. Bertone, I. Scimemi and A. Vladimirov, "Extraction of unpolarized quark transverse momentum dependent parton distributions from Drell-Yan/Z-boson production", *JHEP* **06** (2019) 028 [1902.08474].

[182] A. Bacchetta, V. Bertone, C. Bissolotti, G. Bozzi, M. Cerutti, F. Piacenza et al., "Unpolarized Transverse Momentum Distributions from a global fit of Drell-Yan and Semi-Inclusive Deep-Inelastic Scattering data", 2206.07598.

[183] M. Bury, F. Hautmann, S. Leal-Gomez, I. Scimemi, A. Vladimirov and P. Zurita, "PDF bias and flavor dependence in TMD distributions", 2201.07114.

[184] X. Wang, Z. Lu and I. Schmidt, "Transverse momentum spectrum of dilepton pair in the unpolarized $\pi^- N$ Drell-Yan process within TMD factorization", *JHEP* **08** (2017) 137 [1707.05207].

[185] A. Vladimirov, "Pion-induced Drell-Yan processes within TMD factorization", *JHEP* **10** (2019) 090 [1907.10356].

[186] R. Boussarie et al., "TMD Handbook", 2304.03302.

[187] S. Arnold, A. Metz and M. Schlegel, "Dilepton production from polarized hadron hadron collisions", *Phys. Rev. D* **79** (2009) 034005 [0809.2262].

[188] D. Boer and W. Vogelsang, "Drell-Yan lepton angular distribution at small transverse momentum", *Phys. Rev. D* **74** (2006) 014004 [hep-ph/0604177].

[189] J. Collins and T. C. Rogers, "Connecting Different TMD Factorization Formalisms in QCD", *Phys. Rev. D* **96** (2017) 054011 [1705.07167].

[190] M. G. Echevarria, A. Idilbi and I. Scimemi, "Factorization Theorem For Drell-Yan At Low q_T And Transverse Momentum Distributions On-The-Light-Cone", *JHEP* **07** (2012) 002 [1111.4996].

[191] M. Grewal, Z.-B. Kang, J.-W. Qiu and A. Signori, "Predictive power of transverse-momentum-dependent distributions", *Phys. Rev. D* **101** (2020) 114023 [2003.07453].

[192] J. Collins, *Foundations of perturbative QCD*, vol. 32. Cambridge University Press, 11, 2013, 10.1017/9781009401845.

[193] G. Bozzi, S. Catani, D. de Florian and M. Grazzini, "Transverse-momentum resummation and the spectrum of the Higgs boson at the LHC", *Nucl. Phys. B* **737** (2006) 73 [hep-ph/0508068].

[194] W. Bizoń, X. Chen, A. Gehrmann-De Ridder, T. Gehrmann, N. Glover, A. Huss et al., "Fiducial distributions in Higgs and Drell-Yan production at N^{3}LL+NNLO", *JHEP* **12** (2018) 132 [1805.05916].

[195] W. J. Stirling and M. R. Whalley, "A Compilation of Drell-Yan cross-sections", *J. Phys. G* **19** (1993) D1.

[196] E. Anassontzis et al., "High mass dimuon production in $\bar{p}n$ and $\pi^- n$ interactions at 125-GeV/c", *Phys. Rev. D* **38** (1988) 1377.

[197] L. A. Harland-Lang, A. D. Martin, P. Motylinski and R. S. Thorne, "Parton distributions in the LHC era: MMHT 2014 PDFs", *Eur. Phys. J. C* **75** (2015) 204 [1412.3989].

[198] J. Pumplin, D. Stump, R. Brock, D. Casey, J. Huston, J. Kalk et al., "Uncertainties of predictions from parton distribution functions. 2. The Hessian method", *Phys. Rev. D* **65** (2001) 014013 [hep-ph/0101032].

[199] D. de Florian, R. Sassot, M. Stratmann and W. Vogelsang, "Extraction of Spin-Dependent Parton Densities and Their Uncertainties", *Phys. Rev. D* **80** (2009) 034030 [0904.3821].

[200] D. de Florian, R. Sassot, M. Epele, R. J. Hernández-Pinto and M. Stratmann, "Parton-to-Pion Fragmentation Reloaded", *Phys. Rev. D* **91** (2015) 014035 [1410.6027].

[201] Y.-H. Lien and A. Chumakov, "Measurement of Target Spin (in)dependent Asymmetries in Dimuon Production in Pion-Nucleon Collisions at COMPASS", *SciPost Phys. Proc.* (2022) 028.

[202] D. Müller, D. Robaschik, B. Geyer, F. M. Dittes and J. Hořejši, "Wave functions, evolution equations and evolution kernels from light ray operators of QCD", *Fortsch. Phys.* **42** (1994) 101 [hep-ph/9812448].

[203] X.-D. Ji, "Gauge-Invariant Decomposition of Nucleon Spin", *Phys. Rev. Lett.* **78** (1997) 610 [hep-ph/9603249].

[204] A. V. Radyushkin, "Scaling limit of deeply virtual Compton scattering", *Phys. Lett. B* **380** (1996) 417 [hep-ph/9604317].

[205] A. V. Radyushkin, "Asymmetric gluon distributions and hard diffractive electroproduction", *Phys. Lett. B* **385** (1996) 333 [hep-ph/9605431].

[206] X.-D. Ji, "Deeply virtual Compton scattering", *Phys. Rev. D* **55** (1997) 7114 [hep-ph/9609381].

[207] J. C. Collins, L. Frankfurt and M. Strikman, "Factorization for hard exclusive electroproduction of mesons in QCD", *Phys. Rev. D* **56** (1997) 2982 [hep-ph/9611433].

[208] A. V. Radyushkin, "Nonforward parton distributions", *Phys. Rev. D* **56** (1997) 5524 [hep-ph/9704207].

[209] M. Diehl, T. Feldmann, R. Jakob and P. Kroll, "Skewed parton distributions in real and virtual Compton scattering", *Phys. Lett. B* **460** (1999) 204 [hep-ph/9903268].

[210] K. Goeke, M. V. Polyakov and M. Vanderhaeghen, "Hard exclusive reactions and the structure of hadrons", *Prog. Part. Nucl. Phys.* **47** (2001) 401 [hep-ph/0106012].

[211] M. Diehl, "Generalized parton distributions", *Phys. Rept.* **388** (2003) 41 [hep-ph/0307382].

[212] A. V. Belitsky and A. V. Radyushkin, "Unraveling hadron structure with generalized parton distributions", *Phys. Rept.* **418** (2005) 1 [hep-ph/0504030].

[213] S. Boffi and B. Pasquini, "Generalized parton distributions and the structure of the nucleon", *Riv. Nuovo Cim.* **30** (2007) 387 [0711.2625].

[214] M. Vanderhaeghen, P. A. M. Guichon and M. Guidal, "Hard electroproduction of photons and mesons on the nucleon", *Phys. Rev. Lett.* **80** (1998) 5064.

[215] M. Vanderhaeghen, P. A. M. Guichon and M. Guidal, "Deeply virtual electroproduction of photons and mesons on the nucleon: Leading order amplitudes and power corrections", *Phys. Rev. D* **60** (1999) 094017 [hep-ph/9905372].

[216] S. V. Goloskokov and P. Kroll, "Vector meson electroproduction at small Bjorken-x and generalized parton distributions", *Eur. Phys. J. C* **42** (2005) 281 [hep-ph/0501242].

[217] L. Favart, M. Guidal, T. Horn and P. Kroll, "Deeply Virtual Meson Production on the nucleon", *Eur. Phys. J. A* **52** (2016) 158 [1511.04535].

[218] A. V. Belitsky, D. Mueller and A. Kirchner, "Theory of deeply virtual Compton scattering on the nucleon", *Nucl. Phys. B* **629** (2002) 323 [hep-ph/0112108].

[219] D. Müller, T. Lautenschlager, K. Passek-Kumericki and A. Schaefer, "Towards a fitting procedure to deeply virtual meson production - the next-to-leading order case", *Nucl. Phys. B* **884** (2014) 438 [1310.5394].

[220] K. Kumerički and D. Müller, "Description and interpretation of DVCS measurements", *EPJ Web Conf.* **112** (2016) 01012 [1512.09014].

[221] K. Kumericki, S. Liuti and H. Moutarde, "GPD phenomenology and DVCS fitting: Entering the high-precision era", *Eur. Phys. J. A* **52** (2016) 157 [1602.02763].

[222] X.-D. Ji, "Off forward parton distributions", *J. Phys. G* **24** (1998) 1181 [hep-ph/9807358].

[223] H. Pagels, "Energy-Momentum Structure Form Factors of Particles", *Phys. Rev.* **144** (1966) 1250.

[224] M. V. Polyakov, "Generalized parton distributions and strong forces inside nucleons and nuclei", *Phys. Lett. B* **555** (2003) 57 [hep-ph/0210165].

[225] M. V. Polyakov and P. Schweitzer, "Forces inside hadrons: pressure, surface tension, mechanical radius, and all that", *Int. J. Mod. Phys. A* **33** (2018) 1830025 [1805.06596].

[226] V. D. Burkert, L. Elouadrhiri, F. X. Girod, C. Lorcé, P. Schweitzer and P. E. Shanahan, "Colloquium: Gravitational Form Factors of the Proton", 2303.08347.

[227] M. Burkardt, "Impact parameter dependent parton distributions and off forward parton distributions for zeta —> 0", *Phys. Rev. D* **62** (2000) 071503 [hep-ph/0005108].

[228] QCDSF collaboration, M. Gockeler, R. Horsley, D. Pleiter, P. E. L. Rakow, A. Schafer, G. Schierholz et al., "Generalized parton distributions from lattice QCD", *Phys. Rev. Lett.* **92** (2004) 042002 [hep-ph/0304249].

[229] LHPC, SESAM collaboration, P. Hagler, J. W. Negele, D. B. Renner, W. Schroers, T. Lippert and K. Schilling, "Moments of nucleon generalized parton distributions in lattice QCD", *Phys. Rev. D* **68** (2003) 034505 [hep-lat/0304018].

[230] LHPC, SESAM collaboration, LHPC, P. Hagler, J. W. Negele, D. B. Renner, W. Schroers, T. Lippert et al., "Transverse structure of nucleon parton distributions from lattice QCD", *Phys. Rev. Lett.* **93** (2004) 112001 [hep-lat/0312014].

[231] LHPC, SESAM collaboration, W. Schroers et al., "Moments of nucleon spin dependent generalized parton distributions", *Nucl. Phys. B Proc. Suppl.* **129** (2004) 907 [hep-lat/0309065].

[232] QCDSF, UKQCD collaboration, M. Gockeler, P. Hagler, R. Horsley, D. Pleiter, P. E. L. Rakow, A. Schafer et al., "Quark helicity flip generalized parton distributions from two-flavor lattice QCD", *Phys. Lett. B* **627** (2005) 113 [hep-lat/0507001].

[233] QCDSF, UKQCD collaboration, M. Göckeler, P. Hägler, R. Horsley, Y. Nakamura, D. Pleiter, P. E. L. Rakow et al., "Transverse spin structure of the nucleon from lattice QCD simulations", *Phys. Rev. Lett.* **98** (2007) 222001 [hep-lat/0612032].

[234] P. Hagler, "Hadron structure from lattice quantum chromodynamics", *Phys. Rept.* **490** (2010) 49 [0912.5483].

[235] Y.-S. Liu, W. Wang, J. Xu, Q.-A. Zhang, J.-H. Zhang, S. Zhao et al., "Matching generalized parton quasidistributions in the RI/MOM scheme", *Phys. Rev. D* **100** (2019) 034006 [1902.00307].

[236] X. Ji, A. Schäfer, X. Xiong and J.-H. Zhang, "One-Loop Matching for Generalized Parton Distributions", *Phys. Rev. D* **92** (2015) 014039 [1506.00248].

[237] S. Bhattacharya, C. Cocuzza and A. Metz, "Generalized quasi parton distributions in a diquark spectator model", *Phys. Lett. B* **788** (2019) 453 [1808.01437].

[238] S. Bhattacharya, C. Cocuzza and A. Metz, "Exploring twist-2 GPDs through quasidistributions in a diquark spectator model", *Phys. Rev. D* **102** (2020) 054021 [1903.05721].

[239] M. Constantinou et al., "Parton distributions and lattice-QCD calculations: Toward 3D structure", *Prog. Part. Nucl. Phys.* **121** (2021) 103908 [2006.08636].

[240] A. V. Radyushkin, "Generalized parton distributions and pseudodistributions", *Phys. Rev. D* **100** (2019) 116011 [1909.08474].

[241] K. Cichy, L. Del Debbio and T. Giani, "Parton distributions from lattice data: the nonsinglet case", *JHEP* **10** (2019) 137 [1907.06037].

[242] C. Alexandrou, K. Cichy, M. Constantinou, K. Hadjiyiannakou, K. Jansen, A. Scapellato et al., "Unpolarized and helicity generalized parton distributions of the proton within lattice QCD", *Phys. Rev. Lett.* **125** (2020) 262001 [2008.10573].

[243] S. Bhattacharya, K. Cichy, M. Constantinou, J. Dodson, X. Gao, A. Metz et al., "Generalized parton distributions from lattice QCD with asymmetric momentum transfer: Unpolarized quarks", *Phys. Rev. D* **106** (2022) 114512 [2209.05373].

[244] A. V. Radyushkin, "Symmetries and structure of skewed and double distributions", *Phys. Lett. B* **449** (1999) 81 [hep-ph/9810466].

[245] V. Guzey and M. V. Polyakov, "Dual parameterization of generalized parton distributions and description of DVCS data", *Eur. Phys. J. C* **46** (2006) 151 [hep-ph/0507183].

[246] M. Guidal, M. V. Polyakov, A. V. Radyushkin and M. Vanderhaeghen, "Nucleon form-factors from generalized parton distributions", *Phys. Rev. D* **72** (2005) 054013 [hep-ph/0410251].

[247] M. Diehl, T. Feldmann, R. Jakob and P. Kroll, "Generalized parton distributions from nucleon form-factor data", *Eur. Phys. J. C* **39** (2005) 1 [hep-ph/0408173].

[248] S. Ahmad, H. Honkanen, S. Liuti and S. K. Taneja, "Generalized Parton Distributions from Hadronic Observables: Zero Skewness", *Phys. Rev. D* **75** (2007) 094003 [hep-ph/0611046].

[249] X.-D. Ji, W. Melnitchouk and X. Song, "A Study of off forward parton distributions", *Phys. Rev. D* **56** (1997) 5511 [hep-ph/9702379].

[250] V. Y. Petrov, P. V. Pobylitsa, M. V. Polyakov, I. Bornig, K. Goeke and C. Weiss, "Off - forward quark distributions of the nucleon in the large N(c) limit", *Phys. Rev. D* **57** (1998) 4325 [hep-ph/9710270].

[251] M. Penttinen, M. V. Polyakov and K. Goeke, "Helicity skewed quark distributions of the nucleon and chiral symmetry", *Phys. Rev. D* **62** (2000) 014024 [hep-ph/9909489].

[252] J. Ossmann, M. V. Polyakov, P. Schweitzer, D. Urbano and K. Goeke, "The Generalized parton distribution function (E**u + E**d)(x,xi,t) of the nucleon in the chiral quark soliton model", *Phys. Rev. D* **71** (2005) 034011 [hep-ph/0411172].

[253] M. Wakamatsu and H. Tsujimoto, "The Generalized parton distribution functions and the nucleon spin sum rules in the chiral quark soliton model", *Phys. Rev. D* **71** (2005) 074001 [hep-ph/0502030].

[254] H. Mineo, S. N. Yang, C.-Y. Cheung and W. Bentz, "The Generalized parton distributions of the nucleon in the NJL model based on the Faddeev approach", *Nucl. Phys. B Proc. Suppl.* **141** (2005) 281 [hep-ph/0502017].

[255] S. Scopetta and V. Vento, "Generalized parton distributions and composite constituent quarks", *Phys. Rev. D* **69** (2004) 094004 [hep-ph/0307150].

[256] A. Mukherjee and M. Vanderhaeghen, "Off forward matrix elements in light front Hamiltonian QCD", *Phys. Lett. B* **542** (2002) 245 [hep-ph/0206159].

[257] S. Boffi, B. Pasquini and M. Traini, "Linking generalized parton distributions to constituent quark models", *Nucl. Phys. B* **649** (2003) 243 [hep-ph/0207340].

[258] S. Boffi, B. Pasquini and M. Traini, "Helicity dependent generalized parton distributions in constituent quark models", *Nucl. Phys. B* **680** (2004) 147 [hep-ph/0311016].

[259] B. Pasquini, M. Traini and S. Boffi, "Nonperturbative versus perturbative effects in generalized parton distributions", *Phys. Rev. D* **71** (2005) 034022 [hep-ph/0407228].

[260] B. Pasquini, M. Pincetti and S. Boffi, "Chiral-odd generalized parton distributions in constituent quark models", *Phys. Rev. D* **72** (2005) 094029 [hep-ph/0510376].

[261] B. Pasquini and S. Boffi, "Virtual meson cloud of the nucleon and generalized parton distributions", *Phys. Rev. D* **73** (2006) 094001 [hep-ph/0601177].

[262] T. Frederico, E. Pace, B. Pasquini and G. Salme, "Generalized parton distributions of the pion in a covariant Bethe-Salpeter model and

light-front models", *Nucl. Phys. B Proc. Suppl.* **199** (2010) 264 [0911.1736].

[263] C. Mezrag, L. Chang, H. Moutarde, C. D. Roberts, J. Rodríguez-Quintero, F. Sabatié et al., "Sketching the pion's valence-quark generalised parton distribution", *Phys. Lett. B* **741** (2015) 190 [1411.6634].

[264] K. Raya, Z.-F. Cui, L. Chang, J.-M. Morgado, C. D. Roberts and J. Rodriguez-Quintero, "Revealing pion and kaon structure via generalised parton distributions *", *Chin. Phys. C* **46** (2022) 013105 [2109.11686].

[265] J.-L. Zhang, K. Raya, L. Chang, Z.-F. Cui, J. M. Morgado, C. D. Roberts et al., "Measures of pion and kaon structure from generalised parton distributions", *Phys. Lett. B* **815** (2021) 136158 [2101.12286].

[266] J. M. M. Chávez, V. Bertone, M. Defurne, F. de Soto, C. Mezrag, H. Moutarde et al., "Accessing pion GPDs through the Sullivan process: is it feasible?", *Rev. Mex. Fis. Suppl.* **3** (2022) 0308099 [2203.16947].

[267] C. Mezrag, "Generalised Parton Distributions in Continuum Schwinger Methods: Progresses, Opportunities and Challenges", *Particles* **6** (2023) 262.

[268] C. D. Roberts, "Hadron Structure Using Continuum Schwinger Function Methods", *Few Body Syst.* **64** (2023) 51 [2304.00154].

[269] M. V. Polyakov and C. Weiss, "Skewed and double distributions in pion and nucleon", *Phys. Rev. D* **60** (1999) 114017 [hep-ph/9902451].

[270] B. C. Tiburzi and G. A. Miller, "Exploring skewed parton distributions with two-body models on the light front 2: Covariant Bethe-Salpeter approach", *Phys. Rev. D* **65** (2002) 074009 [hep-ph/0109174].

[271] A. Mukherjee, I. V. Musatov, H. C. Pauli and A. V. Radyushkin, "Power law wave functions and generalized parton distributions for pion", *Phys. Rev. D* **67** (2003) 073014 [hep-ph/0205315].

[272] L. Theussl, S. Noguera and V. Vento, "Generalized parton distributions of the pion in a Bethe-Salpeter approach", *Eur. Phys. J. A* **20** (2004) 483 [nucl-th/0211036].

[273] M. Diehl, A. Manashov and A. Schafer, "Generalized parton distributions for the pion in chiral perturbation theory", *Phys. Lett. B* **622** (2005) 69 [hep-ph/0505269].

[274] C.-R. Ji, Y. Mishchenko and A. Radyushkin, "Higher Fock state contributions to the generalized parton distribution of pion", *Phys. Rev. D* **73** (2006) 114013 [hep-ph/0603198].

[275] W. Broniowski, E. Ruiz Arriola and K. Golec-Biernat, "Generalized parton distributions of the pion in chiral quark models and their QCD evolution", *Phys. Rev. D* **77** (2008) 034023 [0712.1012].

[276] W. Broniowski and E. Ruiz Arriola, "Gravitational and higher-order form factors of the pion in chiral quark models", *Phys. Rev. D* **78** (2008) 094011 [0809.1744].

[277] A. Van Dyck, T. Van Cauteren and J. Ryckebusch, "Support of generalized parton distributions in Bethe-Salpeter models of hadrons", *Phys. Lett. B* **662** (2008) 413 [0710.2271].

[278] V. Shastry, W. Broniowski and E. Ruiz Arriola, "Generalized quasi-, Ioffe-time-, and pseudodistributions of the pion in the Nambu–Jona-Lasinio model", *Phys. Rev. D* **106** (2022) 114035 [2209.02619].

[279] A. Freund and M. F. McDermott, "Next-to-leading order evolution of generalized parton distributions for DESY HERA and HERMES", *Phys. Rev. D* **65** (2002) 056012 [hep-ph/0106115].

[280] APFEL collaboration, V. Bertone, S. Carrazza and J. Rojo, "APFEL: A PDF Evolution Library with QED corrections", *Comput. Phys. Commun.* **185** (2014) 1647 [1310.1394].

[281] V. Bertone, "APFEL++: A new PDF evolution library in C++", *PoS* **DIS2017** (2018) 201 [1708.00911].

[282] V. Bertone, H. Dutrieux, C. Mezrag, J. M. Morgado and H. Moutarde, "Revisiting evolution equations for generalised parton distributions", *Eur. Phys. J. C* **82** (2022) 888 [2206.01412].

[283] G. A. Miller, "Singular Charge Density at the Center of the Pion?", *Phys. Rev. C* **79** (2009) 055204 [0901.1117].

[284] J. A. Rinehimer and G. A. Miller, "Neutron Charge Density from Simple Pion Cloud Models", *Phys. Rev. C* **80** (2009) 025206 [0906.5020].

[285] G. A. Miller, M. Strikman and C. Weiss, "Pion transverse charge density from timelike form factor data", *Phys. Rev. D* **83** (2011) 013006 [1011.1472].

[286] Y. Dong, "Transverse charge density of the pion in a realistic effective Lagrangian approach", *Phys. Rev. C* **81** (2010) 018201.

[287] A. N. Vall, I. A. Perevalova, M. V. Polyakov and A. K. Sokolnikova, "On the problem of singular charge density at the center of pion", *Phys. Part. Nucl. Lett.* **10** (2013) 607.

[288] M. Carmignotto, T. Horn and G. A. Miller, "Pion transverse charge density and the edge of hadrons", *Phys. Rev. C* **90** (2014) 025211 [1404.1539].

List of publications during the PhD

Articles included in the thesis

✯ Barbara Pasquini, Simone Rodini, and Simone Venturini, "Valence quark, sea, and gluon content of the pion from the parton distribution functions and the electromagnetic form factor", *Phys. Rev. D* **107** (2023) 114023 [2303.01789].

✯ Matteo Cerutti, Lorenzo Rossi, Simone Venturini, Alessandro Bacchetta, Valerio Bertone, Chiara Bissolotti, and Marco Radici, "Extraction of pion transverse momentum distributions from Drell-Yan data", *Phys. Rev. D* **107** (2023) 014024 [2210.01733].

Conference Proceedings

❄ Simone Venturini, Barbara Pasquini, Simone Rodini, "A model for pion collinear parton distribution function and form factor", *proceedings of the 30th International Workshop on Deep-Inelastic Scattering and Related Subjects*, **7** (2023) [2307.10735].

Articles in preparation

→ A. Bacchetta, M. Cerutti, B. Pasquini, M. Radici, S. Rodini, L. Rossi, S. Venturini, "Fitting the pion TMD within a light-front model", *in preparation.*

→ V. Bertone, B. Pasquini, S. Rodini, S. Venturini, "Insight in the pion generalized parton distributions within a light-front model", *in preparation.*

Articles not included in the thesis

❄ Simone Venturini, Chiara Aimè, Arianna Armanetti, Daniele Aurelio, Ettore Budassi, Davide Gianatti, Diego Maragnano, Paolo Montagna, Michele Pirola, Davide Santostasi, Simone Restelli, Luca Zatti, "PER (Particle Escape Room) me si va ne la fisica recente." Proceedings of 41st International Conference on High Energy Physics, (2022), PoS ICHEP2022 1164.