Analysis of Transverse-Momentum Distributions of Hadrons Produced in Deep-Inelastic Scattering

Jasone Garay García
Departamento de Física Teórica e Historia de la Ciencia, UPV-EHU

MASTER IN QUANTUM SCIENCE AND TECHNOLOGY

Project directed by Gunar Schnell and Charlotte Van Hulse

September 2012, Leioa

Contents

1 Inclusive and Semi-Inclusive Deep Inelastic Scattering **5**
- 1.1 Inclusive DIS . . . 5
- 1.2 Semi-Inclusive DIS . . . 8
  - 1.2.1 DIS Cross-Section & Structure Functions . . . 8

2 The HERMES Experiment **17**
- 2.1 HERA Polarised Beam . . . 17
- 2.2 The HERMES Spectrometer . . . 18
  - 2.2.1 Trigger Detectors . . . 19
  - 2.2.2 Tracking Detectors . . . 20
  - 2.2.3 Particle Detection and Identification . . . 22

3 Data Selection and Analysis **25**
- 3.1 Data Quality . . . 26
- 3.2 Event Selection . . . 27
  - 3.2.1 Geometric cuts . . . 27
  - 3.2.2 Particle Identification - PID . . . 29
  - 3.2.3 Kinematic cuts . . . 32
- 3.3 Monte Carlo Simulations . . . 34

4 Results **37**
- 4.1 Comparison of PYTHIA and Data . . . 37
- 4.2 Born data set and GMC_Trans . . . 42
  - 4.2.1 GMC_Trans . . . 42
  - 4.2.2 Born data set . . . 45

5 Conclusions **63**

Introduction

The aim of this master thesis project is to develop a better description of the momentum distributions of hadrons produced in deep inelastic processes. These momentum distributions have two components: one corresponding to the intrinsic transverse momentum of the quark inside the proton, and one corresponding to the transverse momentum acquired by the quark during the hadronisation process.

The importance of the understanding of these intrinsic transverse momentum of the quark inside the proton resides on the three-dimensional picture of the hadron it provides us. When one analyses Parton Distribution Functions (PDF) a very exact and well known set of functions give us a one-dimensional picture-like description of the nucleon. However, when we perform a study over the directions perpendicular to the collision axis, access to different information we could not perceive only with a one-dimensional study is possible.

The HERMES experiment was designed to study in more detail the spin of the proton. Since the discovery that only a 30% of the spin of the proton is provided by the quarks in it, several ideas were implemented to explain the missing 70%. Among these ideas, the possibility that quarks have orbital angular momentum was presented, and its study and analysis is mainly performed through the analysis of the transverse momentum distributions. This master thesis does not go into so deep detail in the applications of these distributions, but aims to perform a better description of them.

My personal contribution to this work with the help of my supervisors has been varied and comprises different fields. First of all the adaptation and writing of the code used to analyse the measured data had to be done, as well as the visualization of the data and comparison with PYTHIA data. Experimental data in the acceptance were compared to similar data provided by a PYTHIA Monte Carlo

simulation in order to investigate the general description of hadron production - including transverse degrees of freedom- by the PYTHIA event generator. The main focus, however, was on analysis of HERMES data corrected for acceptance and other detector effects with the goal to extract information about the transverse momentum distribution of hadrons and -in a model-dependent way- of the transverse momentum of quarks inside the proton and acquired in the hadronisation process.

This thesis is organised as follows. In Chapter 1, we introduce the theory needed to understand the physics framework used to develop the work. In it, concepts related to Inclusive and Semi-Inclusive Deep Inelastic Scattering are meant to be cleared up. In the Chapters 2 and 3, the HERMES experiment and the data analysis methodology is explained. The aim of these chapter is to give a wider view of the experimental process performed and also of the analysis work developed during the performance of this thesis. Finally, the last two chapters present a discussion of the results and conclusions obtained.

1

Inclusive and Semi-Inclusive Deep Inelastic Scattering

This section has the aim to describe how the inner structure of the nucleons can be studied through the following two interactions: Inclusive deep-inelastic scattering (DIS) and semi-inclusive deep-inelastic scattering (SIDIS).

Let's start with a historical review of the study of the structure of the proton. In the 50's the scientists at SLAC (Stanford Linear ACcelerator) were able to determine an approximate value for the size of the proton. Later in the 60's a big step was given, when it was discovered that the proton had some kind of grainy internal structure, and that therefore it was not an elementary particle. Finally, it was the joint effort of SLAC, CERN and FermiLab which gave rise to the QPM (Quark-Parton Model), a model that allows us to describe the proton in terms of its constituents quarks bound inside the proton by gluon exchange forces [1].

1.1 Inclusive DIS

Inclusive DIS is the short way of referring to Inclusive Deep Inelastic Scattering. It is an interaction in which a lepton scatters off a nucleon giving rise to the following process:

$$l(p) + N(P) \rightarrow l'(p') + X, \tag{1.1}$$

where l denotes the incident lepton (beam), N the nucleon target (the proton in our particular case), l' is the scattered lepton and X is the final unobserved hadron

system. The value between brackets represents the four-momentum of the corresponding particle. In DIS we do not look at the different hadrons created as a consequence of the interaction. In order to prove the existence of the quarks, a high resolution is required, and therefore the momentum transfer has to be of the order of $Q^2 = \frac{\hbar}{R^2}$, where R is the radius of the target nucleon. At HERA, energies of 27.5 GeV were achieved for the electron/positron beam, which gives rise to a centre of mass energy of the order of $\sqrt{57}$ GeV for HERMES.

Under these physical conditions it is possible to assume that high energy and momentum are transferred by the incident lepton to the nucleon and that this lepton interacts incoherently with the free partons inside the nucleon. Due to relativistic time dilation the interaction time between one quark and the rest of the quarks in the nucleon is much longer than the interaction time between the quark and the virtual photon. Therefore, it can be assumed that the quarks behave as free particles inside the nucleon. In addition, any momentum transverse to the direction of the nucleon can be neglected as a first approximation. This gives rise to the Quark-Parton Model frame, in which DIS can be seen as an elastic scattering of a single virtual photon from a point-like "quasi-free" spin 1/2 constituent quark in the nucleon [2].

Kinematics

In this subsection we will analyse the kinematics of this process in the concrete case of the HERMES experiment. The exact interaction we have is the following one (following the notation explained above):

$$e(p) + p(P) \rightarrow e'(p') + X. \tag{1.2}$$

We can define the invariant squared mass of this system in the rest frame of the proton in terms of the energy of the beam, E_{beam} and the mass of the proton, m_P as follows:

$$S \equiv (p+P)^2 = (E_{beam} + m_P)^2 - E_{beam}^2 = 2m_P(E_{beam} + m_P), \tag{1.3}$$

where we have neglected the masses of the electron (or positron) and the scattered lepton. This is a good approximation at HERA energies. The four-momentum

transfer is defined by

$$Q^2 = -q^2 = -(p - p')^2 \simeq 4E_{beam}E' \sin^2 \frac{\theta}{2}, \tag{1.4}$$

where θ represents the angle between the incident lepton and the scattered one and E' is the energy of the scattered lepton.

Another convenient parameter to study is the invariant mass of the hadronic system, W^2:

$$W^2 = (q + P)^2 = m_P + 2P \cdot q + q^2. \tag{1.5}$$

The kinematic ranges of W^2 and Q^2 go from zero to S, where S is the centre of mass energy of the reaction. Apart from these two important variables, it is also interesting to define the following ones to have a wider view of the outgoing process:

$$\nu = \frac{P \cdot q}{m_P} = E_{beam} - E', \tag{1.6}$$

where E' is the energy of the scattered lepton. Here ν represents the energy transferred by the virtual photon to the target proton, and it is a Lorentz invariant.

This definition allows us to rewrite W^2 (1.5) in a more convenient way:

$$W^2 = m_P + 2m_P\nu - Q^2. \tag{1.7}$$

Finally,

$$x_{Bjorken} \equiv \frac{Q^2}{2Pq} = \frac{Q^2}{2m_P\nu}. \tag{1.8}$$

The dimensionless variable $x_{Bjorken}$ is a measure of the inelasticity of the process and can only take values between 0 and 1. When $x_{Bjorken}$ tends to 1 the process tends to elasticity, and consequently, the nucleon does not break up, whereas when $x_{Bjorken}$ tends to 0 we know we are dealing with an inelastic interaction, because it means that almost all the energy has been used to create new hadrons. In the Quark-Parton model we can also interpret $x_{Bjorken}$ as the fraction of the momentum x of the proton carried by the struck quark and I will just call it x from now on.

Finally, in order to know which proportion of the whole energy available in the beginning is used to form hadrons we define the variable y (also dimensionless and bounded between 0 and 1):

$$y \equiv \frac{P \cdot q}{P \cdot p} = \frac{E_{beam} - E'}{E_{beam}}. \tag{1.9}$$

In the next sections we will explain the structure functions and we will understand how all these variables are used and where their importance resides.

1.2 Semi-Inclusive DIS

In Semi-Inclusive DIS we consider the following process:

$$e(p) + p(P) \rightarrow e'(p') + h(P_h) + X. \tag{1.10}$$

So, in this case we also take into account at least one of the hadrons created as a consequence of the interaction. At HERMES the hadrons that can be created among others are pions (π^+, π^- and π^0), kaons (k^+ and k^-) and protons and antiprotons. Therefore, apart from the variables already defined it is necessary to define the following kinematic dimensionless variable, describing the hadron fractional energy:

$$z_h = \frac{P \cdot P_h}{P \cdot q} = \frac{E_h}{\nu}, \tag{1.11}$$

where P_h is the momentum of the observed hadron.

A picture of a SIDIS process can be observed in figure 1.1.

1.2.1 DIS Cross-Section & Structure Functions

The main purpose of these electron-proton collisions and the study of the variables described above is the better understanding of the structure of the proton. This knowledge is not only important as a test of QCD, but also gives rise to new physics effects.

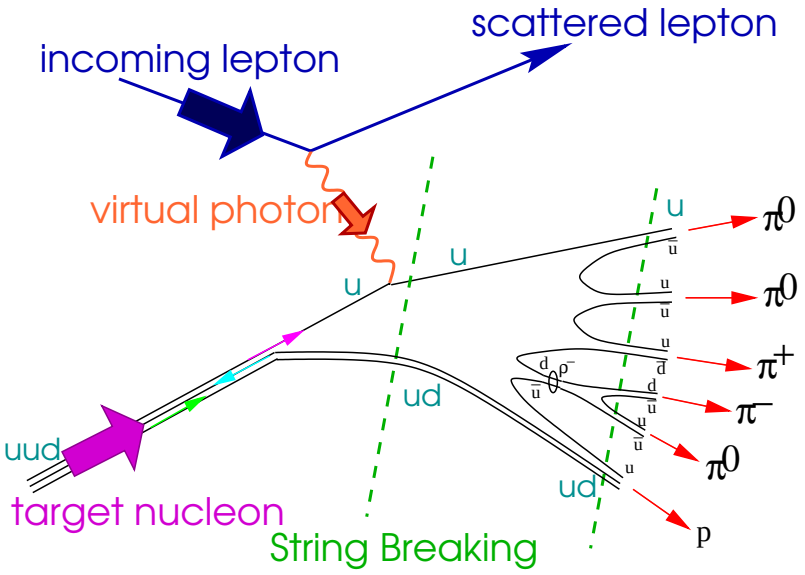

Figure 1.1: Depiction of a generic SIDIS process. The string breaking represents the recombination of the quarks after the breaking of the target nucleon as a consequence of the collision with the incoming lepton. The second green dashed line represents the hadronisation process, in which several different hadrons are created from the quarks generated [3].

What we study is the cross section of the collision between the electron and the quark inside the proton. It appears as a contraction between the leptonic and hadronic tensor as follows:

$$\frac{d^3\sigma}{dxdyd\phi_s} = \frac{\alpha^2}{2sx_BQ^2} L_{\mu\nu}(p,p',\lambda_e) 2M W^{\mu\nu}(q,P,S), \tag{1.12}$$

where $\alpha = \frac{e^2}{4\pi}$, and e is the charge of the electron.

In this cross section we can distinguish two main contributions: $L_{\mu\nu}(p,p',\lambda_e)$ corresponds to the QED part of the interaction, i.e. the photon radiation by the lepton, whereas $W^{\mu\nu}(q,P,S)$ represents the interaction of the virtual photon with the nucleon, whose structure is described by QCD. The QED interaction can be exactly calculated, but not so the latter. In this case we try to parametrize the cross-section with the help of structure functions as we will see in this chapter.

Both tensors can be divided into two parts: a symmetric (labelled with S) and an antisymmetric (labelled with A) part. The antisymmetric part only plays a role in the calculations when we have spin dependence:

$$L_{\mu\nu} = L^{(S)}_{\mu\nu} + iL^{(A)}_{\mu\nu}, \tag{1.13}$$

$$W_{\mu\nu} = W^{(S)}_{\mu\nu} + iW^{(A)}_{\mu\nu}. \tag{1.14}$$

This gives us the following expression for the cross-section:

$$\frac{d^3\sigma}{dx_B dy d\phi_s} = \frac{\alpha^2}{2sx_B Q^2}(L^{(S)}_{\mu\nu} 2MW^{\mu\nu}_{(S)} - L^{(A)}_{\mu\nu} 2MW^{\mu\nu}_{(A)}). \tag{1.15}$$

The leptonic tensor can be exactly calculated, and has the following structure:

$$L_{\mu\nu} = L^{(S)}_{\mu\nu} + iL^{(A)}_{\mu\nu} = 2[p_\mu p'_\nu + p_\nu p'_\mu - g_{\mu\nu} pp' + i\epsilon_{\mu\nu\alpha\beta} s^\alpha q^\beta] \tag{1.16}$$

where we have applied the hermiticity condition on the first step with respect to equation (1.16).

Here $g_{\mu\nu}$ is the Minkowski metric tensor and $\epsilon_{\mu\nu\alpha\beta}$ is the antisymmetric Levi-Civita tensor. As already mentioned before, here the antisymmetric (spin-dependent) part vanishes for the unpolarized case.

However, for the case of the hadronic tensor we cannot exactly determine it, but only parametrize it in terms of structure functions applying the following conditions:

- Hermiticity $\rightarrow W^*_{\mu\nu}(q, P, S) = W_{\mu\nu}(q, P, S)$
- Parity $\rightarrow L^\rho_\mu L^\sigma_\nu W_{\rho\sigma}(q, P, S) = W_{\mu\nu}(\tilde{q}, \tilde{P}, \tilde{-S})$
- Time-reversal $\rightarrow L^\rho_\mu L^\sigma_\nu W_{\rho\sigma}(q, P, S) = W_{\mu\nu}(\tilde{q}, \tilde{P}, \tilde{S})$

where $\tilde{q}^\nu = L^\nu_\rho q^\rho$. In a general way we can write it as follows:

$$W_{\mu\nu} = W\mu\nu^{(S)} + W\mu\nu^{(A)} = \\ (-g_{\mu\nu} - \frac{q_\mu q_\nu}{Q^2})\frac{F_1(x,Q^2)}{M} + (P_\mu + \frac{P_q}{Q^2}q_\mu)(P_\nu + \frac{P_q}{Q^2}q_\nu)\frac{F_2(x,Q^2)}{MPq} \\ + i\epsilon_{\mu\nu\alpha\beta}q^\alpha\frac{1}{Pq}[S^\beta g_1(x,Q^2) + (S^\beta - \frac{Sq}{Pq}P^\beta)g_2(x,Q^2)]. \quad (1.17)$$

Here we have two kinds of nucleon structure function terms depending on the spin: the spin-independent $F_1(x,Q^2)$ and $F_2(x,Q^2)$ and the spin-dependent $g_1(x,Q^2)$ and $g_2(x,Q^2)$. For the purposes of this study we do not need to take care of the polarisation of the beam. The target was not polarised on the data taking period analysed here, and thus none of the terms in the contraction of $L^{\mu\nu}$ and $W^{\mu\nu}$ will involve beam polarisation. So, for our particular case, the unpolarised one, we can write the cross-section in a more explicit way as follows [4]:

$$\frac{d^2\sigma_{UU}}{dxdy} = \frac{4\pi\alpha^2}{sx^2y^2}[xy^2F_1(x,Q^2) + (1 - y - \frac{\gamma^2y^2}{4})F_2(x,Q^2)], \quad (1.18)$$

where $\gamma = \frac{2Mx}{Q}$. Here, the terms UU stand for unpolarised beam and proton.

The function $F_1(x,Q^2)$ is nothing but the Fourier transform of the electric charge distribution inside the nucleon, whereas $F_2(x,Q^2)$ is the Fourier transform of the magnetic moment, and they can be understood as the deviation from the cross-section we would obtain if a relativistic lepton would scatter by a point-like particle [2]. This is exactly a basic ingredient of the Quark Parton Model (QPM).

The inclusive cross-section shown in equation 1.18 can be easily extended to the semi-inclusive case, where new degrees of freedom arise through the presence of an additional hadron with energy fraction z, transverse momentum $P_{h\perp}$ (with reference to γ^* direction) and azimuthal angle ϕ about γ^* direction (see figure 1.2). The cross-section obtained is the following [5]:

$$d\sigma_{UU} \equiv \frac{d^5\sigma_{UU}}{dxdydzdP_{h\perp}^2d\phi} = \frac{1}{2\pi}\frac{\alpha^2}{xyQ^2}\frac{y^2}{2(1-\epsilon)}(1 + \frac{\gamma^2}{2x})\{F_{UU,T} + \epsilon F_{UU,L} \\ + \sqrt{2\epsilon(1+\epsilon)}F_{UU}^{\cos\phi}\cos\phi + \epsilon F_{UU}^{\cos 2\phi}\cos 2\phi\}. \quad (1.19)$$

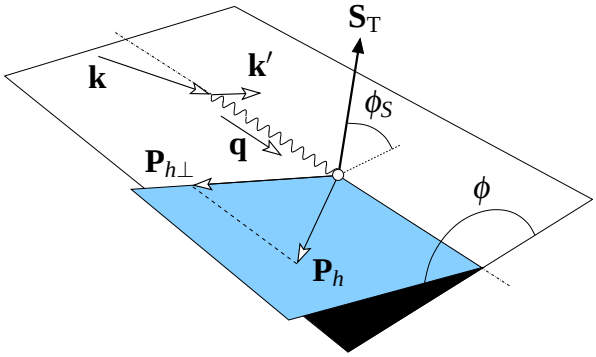

Figure 1.2: Depiction of the azimuthal angle ϕ between the scattering plane, spanned by the three-momenta $(\vec{k}, \vec{k'})$ (in this thesis the corresponding four-momenta defined as (p, p')) of incoming and outgoing leptons and the hadron plane, defined by the respective three-momenta $\vec{q}$ and $\vec{P_h}$ of the virtual photon and the produced hadron (in the theory defined as P and P_h respectively) [3].

Here, the subindexes T and L stand respectively for transversely and longitudinally polarised virtual photon. A more convenient way of presenting this equation is however performing an integration over the ϕ variable:

$$d\sigma_{UU} \equiv \frac{d^5\sigma_{UU}}{dxdydzdP_{h\perp}^2} = \frac{\alpha^2}{xyQ^2}\frac{y^2}{2(1-\epsilon)}(1+\frac{\gamma^2}{2x})\{F_{UU,T} + \epsilon F_{UU,L}\}. \quad (1.20)$$

Parton Distribution Functions

Parton distribution functions (PDFs) $f_i(x, Q^2)$ represent the probability of finding in the proton a parton (a quark or a gluon) of flavour i carrying a fraction x of the proton momentum in a scale Q of the hard interaction [6], where the spin direction of the partons is not considered. Since QCD makes no prediction on the contents of the protons, the shape of these PDFs has to be determined experimentally, through the study of the DIS process.

In the Bjorken limit ($Q^2 \to \infty$ at fixed x), DIS structure functions can be written as a sum of all the PDFs for the different quark and antiquark flavours as follows

[7]:

$$F_2 = 2xF_1 = \sum_i e_i^2 x f_i(x), \tag{1.21}$$

where we used the Callan Gross relation [8] $F_2 = 2xF_1$ and $f_i(x)$ are the PDFs for unpolarised quarks.

For the SIDIS case (also in Bjorken limit) however, the structure functions appear as a sum of the convolution between the PDFs and the fragmentation functions (FF), $\sum_{q,\bar{q}}$PDF$\otimes$ FF.

Fragmentation functions

Isolated stable quarks have never been observed in nature. When a quark is ejected from a bound ensemble of quarks and gluons, for example by absorption of a high energy photon, as it separates from the ensemble, additional quark-antiquark pairs are generated from the vacuum. The newly generated quarks and antiquarks combine with the target quarks. This recombination process continues as the separation between the partons increases until a configuration of stable singlet multiquark states is reached. This is called hadronisation (see figure 1.3), because it is the process of creating new hadrons from the quark ejected from the nucleon, and it is an essential feature of the interaction of quarks in QCD. Its understanding is basic to a complete picture of the dynamics of quark-quark and quark-gluon interactions. Experimentally, this hadronisation process is described by fragmentation functions (FFs), the number densities for the conversion of a struck quark of a given flavor q into a specific hadron type h [9].

Transverse Momentum Distributions

In order to completely describe the momentum distribution of the quarks inside the proton through a SIDIS process, we have to go beyond the collinear description of partons (quarks) in nucleon, and look at the transverse components of the momentum. Several important effects cannot be fully understood only using collinear QCD factorisation. That is the case for example of orbital angular momentum. With Transverse Momentum Distributions (TMDs) information about it can be obtained, even though indirectly (through, e.g., spin-orbit effects). Collinear QCD factorisation by itself would not be able to make predictions for

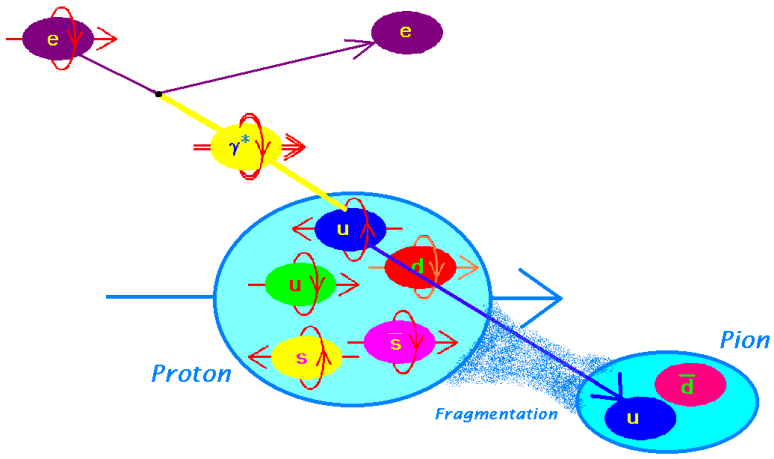

Figure 1.3: Depiction of a pion creation as a consequence of fragmentation after a SIDIS process takes part [3].

processes where transverse momentum is not integrated over, e.g., semi-inclusive DIS or Drell-Yan with the measurement of the photon transverse mometum Q_T. Another clear example is the determination of the W boson mass. It relies on some knowledge of transverse-momentum dependence of parton distributions [10].

The analysis of TMDs however is not trivial at all. In SIDIS processes transverse momentum distributions of PDFs and FFs appear as a convolution, which makes it extremely difficult to unfold. A common and very widespread Ansatz is the hypothesis of a Gaussian dependence of the momentum, for both the PDF and for the FF. The structure function $F_{UU,T}$ can be written as follows in the TMD framework [11]:

$$F_{UU,T} = \sum_q e_q^2 \mathcal{I}[f_1^q(x, \vec{p}_T^{\,2}) D_1^q(z, \vec{k}_T^{\,2})], \tag{1.22}$$

where,

$$\mathcal{I}[f^q D^q] \equiv \int d^2\vec{p}_T d^2\vec{k}_T \delta^{(2)}(\vec{p_T} - \frac{\vec{P_{h\perp}}}{z} - \vec{k_T})[f^q(x, \vec{p_T}) D^q(z, \vec{k_T})]. \tag{1.23}$$

Here, $\vec{p_T}$ is the quark intrinsic transverse momentum, $\vec{k_T}$ the fragmentation trans-

verse momentum and $\vec{P}_{h\perp}$ the hadron transverse momentum [11]. In the following we will omit quark flavours index q.

The previously introduced transverse momentum-integrated versions of the PDF f_1 and FF D_1 are obtained via:

$$f_1(x) = \int d^2\vec{p}_T f_1(x, \vec{p}_T{}^2), \tag{1.24}$$

$$D_1(z) = z^2 \int d^2\vec{k}_T D_1(z, z^2\vec{k}_T{}^2). \tag{1.25}$$

In order to de-convolute the functions in expression (1.22) the following Gaussian Ansatz are often used:

$$f_1(x, \vec{p}_T^2) = f_1(x)\frac{1}{\pi\langle p_T^2\rangle}e^{-\frac{p_T^2}{\langle p_T^2\rangle}}, \tag{1.26}$$

$$D_1(z, z^2\vec{k_T}^2) = D_1(z)\frac{1}{z^2\pi\langle k_T^2\rangle}e^{-\frac{k_T^2}{\langle k_T^2\rangle}}. \tag{1.27}$$

Introducing expressions (1.26) and (1.27) into the SIDIS unpolarised cross-section (1.22) and into the $\mathcal{I}$ integral (eq. (1.23)) we obtain these results:

$$\mathcal{I}[f_1(x, \vec{p}_T^2), D_1(z, z^2\vec{k}_T^2)] = f_1(x)\cdot D_1(z)\cdot\frac{R^2}{\pi z^2}\cdot e^{-R^2\frac{P_{h\perp}^2}{z^2}}, \tag{1.28}$$

where $\frac{1}{R^2} = \frac{\langle P_{h\perp}^2\rangle}{z^2} = \langle p_T^2\rangle + \langle k_T^2\rangle$.

Finally, we can construct our unpolarised SIDIS cross-section in a deconvoluted way [11]:

$$\frac{d^6\sigma}{dxdQ^2dzdP_{h\perp}^2d\phi_h d\phi_s} = \sum_q \frac{e_q^2}{4\pi}\frac{\alpha^2}{(MExy)^2}\frac{R^2}{z^2}e^{\frac{R^2P2_{h\perp}}{z^2}}(1 - y + \frac{y^2}{2})f_1(x)D_1(x). \tag{1.29}$$

The interesting thing is that very little is known on this field, even for the simplest

unpolarised PDFs case here developed.

This thesis aims at providing an input to the transverse momentum dependence of both f_1 and D_1 by looking at the average transverse momenta of the hadrons produced in SIDIS processes measured at the HERMES experiment at DESY and fixing specific parametrisation for $\frac{\langle P_{h\perp}^2 \rangle}{z^2} = \langle p_T^2 \rangle + \langle k_T^2 \rangle$.

2

The HERMES Experiment

The HERa MEasurement of Spin (HERMES) experiment is one of the four experiments that took place at the German national research center DESY (Deutsches Elektronen SYnchroton), Hamburg. It was proposed in the year 1990 to help resolving the so-called proton "spin puzzle". It played an important role in the understanding of the proton structure and its constituents. The data taking period was completed in 2007, but data analysis is still ongoing and particularly active with respect to the investigation of the transverse spin and transverse momentum phenomena in the nucleon.

2.1 HERA Polarised Beam

The HERMES experiment used the HERA polarised beam. The HERA infrastructure allowed for up 220 bunches, separated by 96 ns and with a length of 27 ps. This gave the beam a duty factor of 0.3×10^{-3}. At the beginning of the injection, the average beam current could reach at most 50 mA followed by and exponential decrease with time. After about 10 hours it was stopped and injected again.

The HERA storage ring ran both with electrons and positrons. Electron (positron) beams could become transversely polarised through emision of spin-flip synchroton radiation in the arcs. This effect is the so-called Sokolov-Ternov mechanism [12]. This effect involves a small asymmetry in the spin-flip amplitudes, which enhances the population of the spin state anti-parallel (parallel) to the magnetic bending field in the case of electrons (positrons) [13].

The polarisation $P(t)$ of the beam as a function of time is given by:

$$P(t) = P_{\infty}(1 - e^{-t/\tau}), \tag{2.1}$$

where P_{∞} is the asymptotic polarisation and τ is the time constant. Both P_{∞} and τ are characteristic for the storage ring and can be calculated. In the ideal HERA accelerator the maximum polarisation that could be achieved was $P_{th} = 0.924$ (asymptotic polarisation). The value of τ varies with the bending radius ρ of the storage ring and with the energy E of the beam: it increases with the third power of ρ and decreases with the fifth power of E. In the ideal situation we would have for an energy $E = 27.5$ GeV a time constant of $\tau = 37$ minutes. However, depolarisation effects affect the rise time according to:

$$\tau = (\frac{P_{\infty}}{P_{th}})\tau_{th}. \tag{2.2}$$

Under these conditions, we have typical values for the characteristics of the HERA storage ring of the order of $P_{\infty} = 0.55$ and $\tau \approx 22$ minutes. In order to achieve the highest polarisation, the precise alignment of the quadrupoles is essential. Longitudinal beam polarisation is required for the polarised electron scattering experiments at HERMES. In order to achieve this, a pair or spin rotators was used to swift the polarisation. The beam arrives transversely polarised to the first spin rotator, which rotates the polarisation towards the beam direction. After crossing the gas target zone, the second spin rotator rotates the spin back to the vertical direction.

2.2 The HERMES Spectrometer

The HERMES spectrometer has three different kinds of detectors sorted by their purpose: the trigger detectors, the tracking detectors and particle identification detectors (PIDs). The trigger detector system is based on three hodoscopes and the pre-shower, and is responsible for the event detection. The tracking detectors can be separated into three groups: the front chambers, the magnet chambers, and the back chambers. PID detectors are the transition radiation detectors (TRDs), the ring imaging Cherenkov detector (RICH), the pre-shower detector and the calorimeter.

2.2.1 Trigger Detectors

Not all the signals collected during data taking correspond to interesting physics events. Thus, a set of detectors is needed to indicate us that there is an interesting event undergoing, so that the rest of the detectors can acquire the signals corresponding to this event. In the case of HERMES, if there was a signal in the same half the hodoscopes and pre-shower, together with a deposition of energy greater than a certain value in the calorimeter, then the rest of the detectors received the order to collect this event, because there is a reasonable probability that it corresponded to a physically interesting event.

In order to achieve this, these detectors, especially the hodoscopes and the preshower, needed to have a small detection time. They were built from plastic scintillators. When crossed by a particle, these materials produce light as the result of the excitation and subsequent de-excitation of its molecules. This results in a very fast signal that activated the rest of the detectors. This signal was put in coincidence with the so-called "HERA clock", provided by the lepton beam bunch signal, indicating the time at which the lepton bunches passed through the target and thereby defined a time window for all detector triggers to arrive [14].

The pre-shower ($H2$) was similar to the other hodoscopes $H0$ and $H1$, but had in front of it a layer of lead, in which electromagnetic showers are generated by the passage of electrons. This created an electron-positron pair that converted later into a photon and this excited again the material creating another pair and so on giving rise to a signal cascade. The pre-shower can also be helpful in distinguishing hadrons and leptons, because only leptons generate electro-magnetic showers, whereas hadrons are minimum ionizing particles in lead and therefore lose only a minor part of their energy through ionization.

Finally, the coincidence of signals from the three hodoscope planes $H0$, $H1$ and $H2$ with those from a lead-glass calorimeter was required in order to be certain that the event was indeed interesting. The energy deposited in the calorimeter had to be greater than a certain value, that depended on the running conditions at HERMES. Usually a fill of gas in the target and consequently a detection period took 10 hours before the beam was dumped. During the first 9 hours after the gas target was filled, the density of gas was kept low (low luminosity), and then the threshold in the calorimeter was set to 1.4 GeV. After these 9 hours till the refill of the gas of the target, more gas was injected, such that the beam was re-

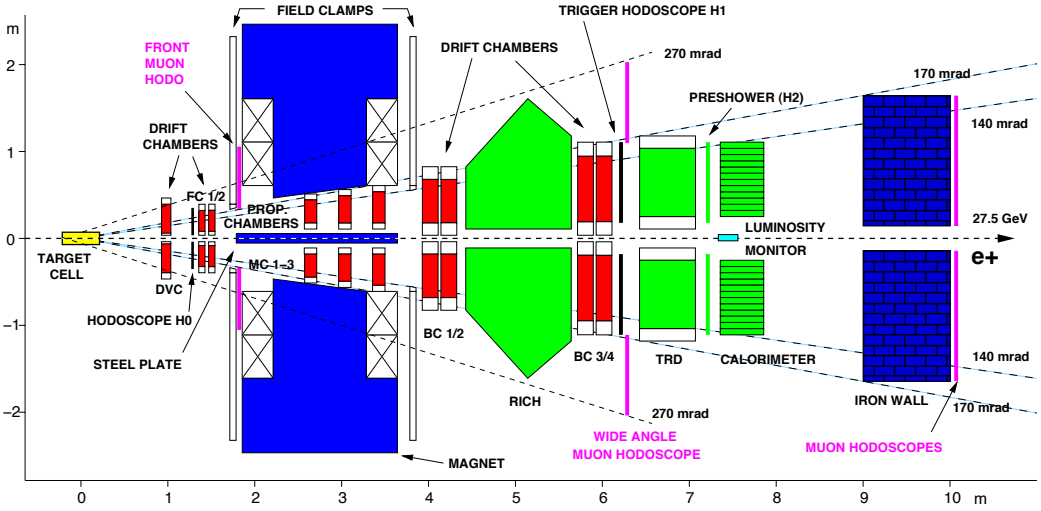

Figure 2.1: The HERMES spectrometer configuration during the years 1998-2000 [3].

ally consumed when crossing the target (high density runs). Under these running conditions, the energy threshold was set to 3.5 GeV, and the density of events was really high.

2.2.2 Tracking Detectors

The front tracking detectors, DVC and FC1/2 in figure 2.1 were located in front of the magnet, and were needed in order to determine the interaction vertex, as well as the polar and azimuthal scattering angles of the particles and the initial trajectory for the determination of the momentum. The forward angle spectrometer [15] was symmetric around the horizontal plane and was divided into two halves to allow the two HERA beams (the proton and the positron/electron beam) to pass through. The beam pipes were protected by the so-called ”septum plate”, a thick iron plate that protected them from the magnetic field (see figure 3.1).

Right after the front tracking detectors were the magnetic chambers. They were located inside of the magnet and were originally used to help resolving multiple tracks in high multiplicity events. They would improve track reconstruction

in case of missing planes in the front chamber. However, due to the low background and the good performance of the front detectors they became unnecessary for those reasons and they were used to reconstruct tracks of low momentum that were strongly deflected by the spectrometer magnet and could consequently not be detected by the back chambers.

Finally, the back chambers were located right after the magnet, and were to reconstruct the trajectories of the particles deflected by the spectrometer magnet [16]. These tracks are called "back tracks". Very low momentum particles went out of the acceptance of the detector, so they were not identified and neither included in the analysis.

At HERMES there were two different types of chambers for tracking: proportional chambers and drift chambers. Both of them are multi-wire chambers, a set of thin, equally separated anode and cathode wires placed in the middle of two cathode planes. However, their working principle is slightly different. Proportional chambers have the space between the cathode planes filled with gas, such that when a particle traverses the chamber, the gas is ionised and an applied electric field will make the freed electrons drift towards one of the anode wires. As a consequence, an avalanche of electrons will be observed near the anode wire, giving rise to an electric signal that reveals the position of the particle. In practice a general set of wires were staggered in order to obtain a good spatial resolution.

On the other side, drift chambers measured the space resolution through the time needed by the electron to reach the anode wire measured from the moment the ionizing particle traversed the detector. The drift chambers allow for a higher space resolution, whereas the multi-wire proportional chambers are more easily operated inside strong magnetic fields.

The relative momentum $\delta p/p$ and the angle $\delta\theta$ resolution that could be obtained was $\delta p/p \leq 2.6\%$ and $\delta\theta \leq 1.8$ mrad [17], always depending on the particle type and momentum. The fast tracking reconstruction was done through the so-called pattern-matching algorithm based on the use of look-up tables generated with Monte Carlo [19].

2.2.3 Particle Detection and Identification

The discrimination between electrons and hadrons, as well as the hadron identification, was based on the information from four detectors: the lead-glass calorimeter, the pre-shower detector, the TRD and the RICH detector. The separation and identification was done using a likelihood method based on empirical responses of each of the four detectors. The average efficiency of the electron/hadron discrimination is of the order or 98-99% for electron identification with only 1% hadron contamination. The identification of particles was based on the various responses of the detectors as detailed below.

When the particles passed through the RICH detector, they crossed two different radiators, one made by aerogel ($n = 1.03$), where n is the refractive index, and another one made by C_4F_{10} ($n = 1.0014$), generating two rings respectively. The radius of these rings together with the knowledge of the refractive index helped determining the mass of the detected particle. This calculation was based on the Cherenkov radiation principle. When a charged particle crosses a dielectric medium (the radiator) with a velocity that exceeds the velocity of the light in that medium, it emits electromagnetic waves (Cherenkov radiation) backwards generating a cone, analogous to shock waves observed in the Mach effect when the sound barrier is crossed, with opening angle given by:

$$\cos\theta_C = \frac{1}{n\frac{v}{c}} \tag{2.3}$$

with

$$v > \frac{c}{n}. \tag{2.4}$$

Here, v is the velocity of the particle, c is the speed of light in the vacuum, n is the refractive index of the medium and θ_C is the Cherenkov opening angle between the wave vector and the particle's direction. This angle depends on the particle mass as follows: the higher the mass of the particle, the lower the velocity of the particle. The combination of the two radiators apart from providing more accurate results, also allowed a wider range in momentum for particle identification. The HERMES RICH detector could detect and distinguish properly particles with a momentum $P \in [2, 15]GeV$, as we can see in figure 2.2. The response of the RICH is studied under the hypothesis of the particle being a pion, a kaon, or a

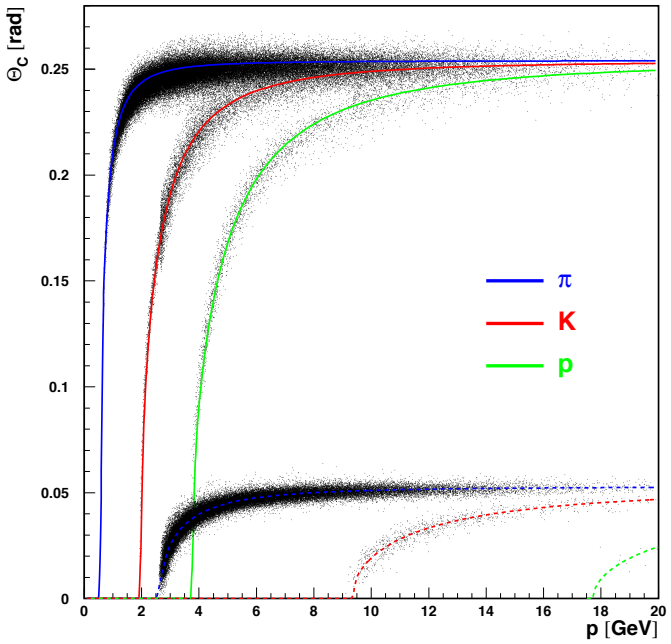

Figure 2.2: Cerenkov angles vs momentum for aerogel and C4F10 [3].

proton in order to try to determine the type of the detected particle.

The TRD is based on the emission of electromagnetic radiation (X-rays) by highly relativistic charged particles, with $\gamma > 1000$ [1] that cross the boundary between two dielectric media [18]. It consisted of 6 modules containing each a radiator and a proportional wire chamber to detect the emitted X-rays. However, at the HERMES energy regime only electrons (positrons) produced transition radiation. Since the radiation was emitted at very small angles of the order of $\theta \sim 1/\gamma$ it was not possible to separate the transition radiation from other processes by which positrons lost energy in the detector. In order to obtain a usable higher signal, several layers are needed, since when a positron crosses the TRD only a very small number of photons is radiated.

The pre-shower detector was located right behind the TRD. It used the amount of energy deposited by the crossing particle to determine if it was a hadron or an

[1] Here $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$

electron (positron). Positrons and electrons lose energy much faster when they transverse matter than hadrons. This is because hadrons transversing matter lose energy mainly through inelastic collisions, that is, excitations and ionizations of atoms, whereas high energy positrons ($E \geq 10MeV$) lose energy mostly through Brehmsstrahlung, causing electromagnetic showers.

The last step in the particle identification was the calorimeter. It consisted of a wall of lead-glass blocks in which electro-magnetic showers developed. It contributed to formation of the trigger signal based on the energy deposition in two adjacent columns. The calorimeter had three main functions: PID, trigger formation and measurement of the energy of the detected electrons and photons. The calorimeter was the only detector that detected photons. Positrons, electrons and photons were stopped inside the 50 cm (corresponding to 18 radiation lenghts) of lead glass composite of the calorimeter. Leptons and hadrons were separated looking at the E/p distribution, where E is the energy deposited in the calorimeter and p is the momentum of the particles as derived from the track curvature. The electron/positron distribution has a sharp peak around $E/p = 1$, while the hadron peak is much broader and appears at much lower energy. In combination with the pre-shower and integrated over the whole energy spectrum it had a 95 % efficiency in electron detection.

3

Data Selection and Analysis

In this section we will analyse how the data is previously selected and sorted, in order to have a good data quality. The selection and analysis of data comprises various steps. A first approach is done at "run level". One run represents a data taking period of $(3-5)$ minutes in which one file of around 560 Mb is stored. This file is later called "run" and an analysis is always done over a list of runs. This list is compiled and represents the first selection of the data, since allows us to choose the data taking period and the desired target gas type we want for our concrete study.

The second step is performed at burst level. A burst is defined as a 10 seconds period in which quality requirements about the overall good performance of HERA, the target and the spectrometer are set up. At this level cuts related to data collection, data quality or the good functioning of the detectors among others are applied.

Next step is the event level. An event is defined every time the electron interacts with the target and thus is the level at which kinematic cuts are applied over the data set in order to select the desired SIDIS events and avoid specific kinematic regions that could include for example nucleon resonances. In each event several tracks are generated, each corresponding to the different particles taking part in the interaction, namely the leading lepton (the lepton from the beam that interacted with the target), and the generated hadrons as a consequence of the hadronisation process. Since kinematic cuts were already applied, in the fourth step performed at track level we can be somewhat sure about being analysing only particles coming from the interaction. So the goal at this level is the identification of the individual particle tracks and the entrenchment of that we avoid

problematic regions at the edge of the detector acceptance.

3.1 Data Quality

In this thesis the data from the year 2000 are used [19]. At this time, HERA was run with positrons and the target size was of 21 mm high, 9 mm wide and 40 cm long, with an angular acceptance of $\theta \in (100$ mrad, 240 mrad$)$. In the present thesis the data analysed was collected on unpolarised target and no feature related to beam polarisation state is taken into account, so that effectively for both parallel and anti-parallel polarisation data are collected. In the year 2000 two different gas types were used as target: hydrogen and deuterium. In this thesis only data collected on a hydrogen target is used.

In order to make sure that the data set is suitable for this analysis, some initial cuts have to be performed from the whole data compiled during the year 2000. In this thesis it is required that the calorimeter, the RICH and the TRD were working perfectly. In addition a minimum burst length and run length is needed, and any problem derived from high voltage trips is avoided. This can be achieved choosing the following bits [20] :

- **2**: Discards data with unphysical dead times.
- **3**: Discards data with unreasonable values of the burst length.
- **4**: The lower cut on the beam current discards data that result in very small count rates. The upper cut discards data with unphysically high currents.
- **6**: Discards the first burst of each run.
- **7**: Discards problems with filling the data tables.
- **8**: Discards data with no available PID.
- **9**: Accepts only runs that are marked analysable in to the logbook.
- **16**: Rejects bad target data.
- **18**: Discards bursts in which at least one block in the Hodoscope H2 or the Luminosity Monitor was dead.

- **19**: Discards bursts with bad TRD data quality.
- **20**: Discards bursts during which one or more planes of the FC or BC tripped.
- **21**: Discards bursts that show a strange calorimeter behaviour or a bad RICH.
- **22**: Discards bursts, during which the RICH tripped.
- **25**: Discards burst with bad RICH data quality.
- **30**: Checks that the dead time is in a physically allowed range.
- **31**: This bit is set if more than one middle block in the calorimeter halves is dead. It is used instead of bit 17 for semi-inclusive analysis of charged particles in order to increase the statistics.

Any data quality requirement related to beam polarisation is neglected, since beam polarisation is not used in this analysis. Concerning data selection one has to be strict in the needs of the analysis but without rejecting too much data, so that the statistics are not impaired. All this information about data quality can be found on the HERMES web page [20].

3.2 Event Selection

In the event selection process, one loops over all the tracks stored of each event with three main purposes. First of all, we aim to find the ”leading lepton”. This corresponds to the identification of the most energetic lepton that survived both geometric and kinematic cuts detailed below. This lepton is supposed to be the one coming from the beam that interacted with the target. Once the leading lepton is identified and the separation between hadrons and leptons is properly performed, a separation between pions, kaons and protons is implemented. All the steps will be explained in the subsequent sections.

3.2.1 Geometric cuts

The geometric cuts are applied to ensure that the track does not hit any inactive detector component. Also, border effects must also be taken into account to ensure correct kinematics. A graphical description is detailed in figure 3.1

The longitudinal vertex coordinate z at which the scattering occurs is restricted to a dimension slightly smaller than the length of the target cell: $-18\,\text{cm} \leqslant z \leqslant 18\,\text{cm}$. In addition, the horizontal and vertical cuts on the magnet and the limits of the septum plate must be taken into account as detailed in figures 3.1 and 3.2. These selection criteria are summarised it in the table 3.1.

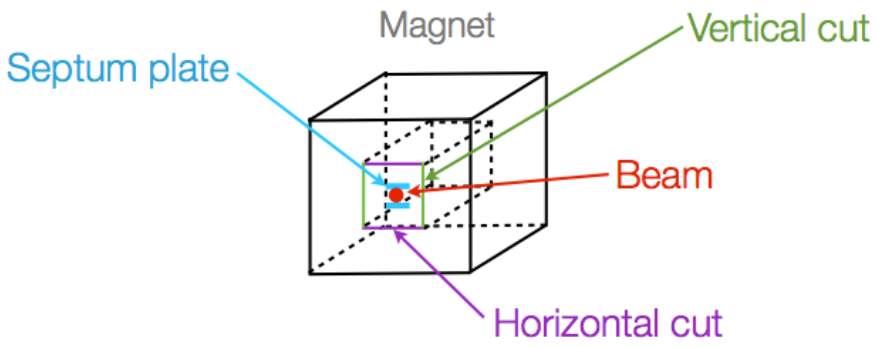

Figure 3.1: Schematic depiction of the magnet and the corresponding geometrical cuts.

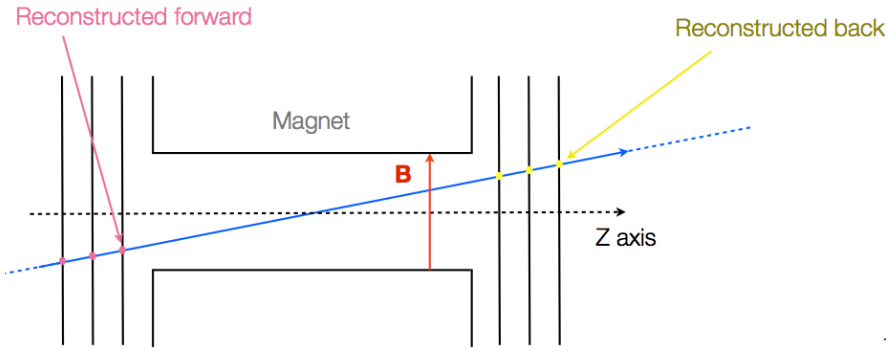

Figure 3.2: A side view of the track reconstruction before and after the particle crosses the magnet.

Limits	Cuts
vertex position	-18 cm $\leqslant z_{vertex} \leqslant 18$ cm
calorimeter position	-175 cm $\leqslant x_{Calorimeter} \leqslant 175$ cm 30 cm $\leqslant y_{Calorimeter} \leqslant 108$ cm
front field clamp position	-31 cm $< x_{ffc} < 31$ cm
septum plate position	-7 cm $> y_{sp} > 7$ cm
rear field clamp position	-54 cm $< y_{rfc} < 54$ cm 100 cm $\leqslant x_{rc} \leqslant 100$ cm
rear clamp position	-54 cm $\leqslant y_{rc} \leqslant 54$ cm

Table 3.1: Geometry cuts are applied to the detected charged particles.

3.2.2 Particle Identification - PID

The particle identification process consists in two main steps. First of all, a distinction between hadrons and leptons is implemented. Once the particles are separated into these two groups, hadron identification is accomplished.

Lepton and hadron separation

Lepton-hadron separation is carried out using calibrated signals of the different PID detectors: the RICH, the TRD, the calorimeter and the pre-shower scintillator. We have several PID detectors, because the more available, the cleaner the separation. The total probability of having a lepton or a hadron is computed as the product of the flux factor, Φ^i, and the parent distribution $\mathcal{P}_D^i$. If a particle with a specific momentum and response, $\mathcal{X}$ is observed, then the probability $\mathfrak{P}_D^i$ that this particle is of type i is given by:

$$\mathfrak{P}_D^i(\mathcal{X}) = \frac{\Phi^i \mathcal{P}_D^i(\mathcal{X})}{\sum_j \Phi^j \mathcal{P}_D^j(\mathcal{X})}. \tag{3.1}$$

The parent distribution describes the probability to observe a specific detector response for a particle of type i with a certain momentum. The flux represents the prior probability that a particle with a specific momentum is of type i.

From the parent distributions $\mathcal{P}_D^i$ the following PID values are determined:

$$PID_3 \equiv \log_{10} \frac{\mathcal{P}^l_{preshower} \mathcal{P}^l_{calorimeter} \mathcal{P}^l_{RICH}}{\mathcal{P}^h_{preshower} \mathcal{P}^h_{calorimeter} \mathcal{P}^h_{RICH}}, \tag{3.2}$$

$$PID_5 \equiv \log_{10} \frac{\prod_{n=1}^{6} \mathcal{P}^l_{TRD_n}}{\prod_{n=1}^{6} \mathcal{P}^h_{TRD_n}}. \tag{3.3}$$

Using the following easy relation of the previously defined PID values, the determination of the particle type as a lepton or a hadron is straightforward:

$$PID_3 + PID_5 - \log_{10} \frac{\Phi^h}{\Phi^l} = \log_{10} \frac{\mathfrak{P}^l_D}{\mathfrak{P}^h_D}. \tag{3.4}$$

The negative values obtained in equation 3.4 correspond to hadrons, whereas the positive values correspond to leptons. The fact that more hadrons than leptons are created in the experiment displaces the minimum of the PID distribution to positive values, as can be seen in the upper plots of the left main panel (see 3.3) and in the right upper two panels (red and blue) of figure 3.4. Correcting the PID distributions for flux factors moves the minimum to zero, as can be seen in 3.3 and in the lower plots of the right two panels (3.4).

At HERMES another PID combination can sometimes be used:

$$PID_2 \equiv \log_{10} \frac{\mathcal{P}^l_{preshower} \mathcal{P}^l_{calorimeter}}{\mathcal{P}^h_{preshower} \mathcal{P}^h_{calorimeter}}. \tag{3.5}$$

Here the RICH is not taken into account. This can for example be appropiate in the case the RICH is not working properly. A comparison of all the possible cases can be seen in figure 3.3. In the figures the negative values correspond to hadrons, whereas the positive values correspond to leptons, as already stated. The minimum then represents the separation between hadrons and leptons. Depending on the accuracy needed for the concrete distinction between leptons and hadrons slightly different cuts can be applied. For the present analysis, a clear lepton-hadron distinction is required, since SIDIS processes are analysed. That is why, examining carefully figures 3.3 and 3.4, the following cuts were chosen:

$$PID_3 + PID_5 - log_{10} \frac{\Phi^h}{\Phi^l} > 2 \longrightarrow \text{Leptons} \tag{3.6}$$

$$PID_3 + PID_5 - log_{10} \frac{\Phi^h}{\Phi^l} < 0 \longrightarrow \text{Hadrons}. \tag{3.7}$$

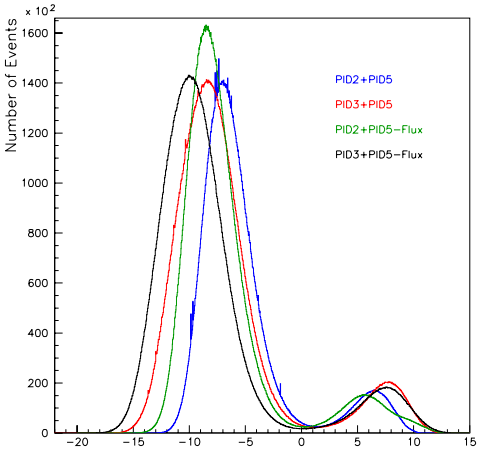

Figure 3.3: Different PID values for the same set of DIS selected data using different PID values. A clear displacement of the minimum can be appreciated.

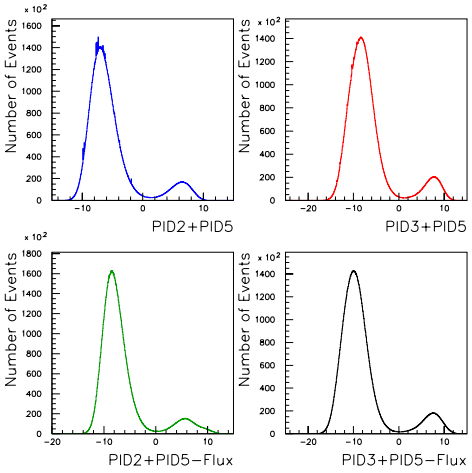

Figure 3.4: Different PID values for the same set of DIS selected data using different PID values.

Hadron identification

The identification of the different hadron types is carried out from the information provided by the RICH detector. This calculation is based on the Cherenkov radiation principle (see chapter 2).

Once a first estimation of the particle type is done, the so called quality parameter rQp has to be evaluated. This parameter helps to determine how sure one can be about the most probable hadron type being indeed the real hadron type:

$$rQp \equiv \log_{10} \frac{P(h_1)}{P(h_2)}, \tag{3.8}$$

where h_1 represents the most likely hadron type, whereas h_2 represents the second most probable hadron type. If $rQp >> 0$, then the detection worked well. However, if the algorithm [21] was not able to determine a most probable hadron type, then the parameter is directly set to zero, $rQp = 0$.

3.2.3 Kinematic cuts

Kinematic cuts are the cuts applied in order to select the desired kinematics for the analysis. In the case of this thesis, SIDIS phenomenology is sought; therefore, adequate cuts on the kinematics variables have to be applied, so that the SIDIS kinematic range is selected and it can be ensured that no other kind of process is contaminating the analysis.

1. $\mathbf{Q^2 > 1 GeV^2}$. Q^2 sets the scale of the interaction. The search for the nucleon constituents requires a sufficiently good resolution in order to resolve them in experiments. The wave length of the probe particle has to be small compared to the radius of the nucleon, $\lambda << R$ or the four-momentum transfer has to be of the order of $Q^2 >> \hbar^2/R^2$ (see picture 3.5) [22]. The exact optimal cut cannot be set, but $Q^2 > 1\text{GeV}^2$ is considered secure enough.

2. $\mathbf{W^2 > 10 GeV^2}$. W gives the centre of mass energy of the virtual photon and nucleon system. If W is large enough, then many particles can be produced, and not just the excitation of the nucleon. This cut is performed in order to suppress any contribution from target fragmentation without discarding a significant amount of useful data. Also acceptance effects are taken into account inside this cut [23].

3. $\mathbf{0.023 < x < 0.4}$. The cuts on the x are a consequence of the HERMES acceptance and of the cuts previously applied on Q^2 and W^2. They are only applied to specify a clean kinematic phase-space because outside of the given boundaries only a negligible amount of data is present.

4. $\mathbf{|E_e| > 4 GeV}$. The energy of the scattered lepton has to be higher than a certain value to select leptons with energies that are safely above the trigger threshold applied on the energy deposit in the calorimeter. This cut ensures a homogeneous data sample. Since during data collection the trigger threshold in the calorimeter is set to 1.4 GeV or 3.5 GeV, but the energy calibration is based on calibration data from the previous years, this cut removes sensitivity to possible variations over time in the calorimeter signal. Furthermore, it corresponds to an upper cut on y avoiding a region where QED radiative corrections become significant.

5. $\mathbf{0.2 < z < 0.7}$. The goal of the lower cut on z is to ensure that the hadron carries information from the struck quark [23], i.e., the hadron we detect

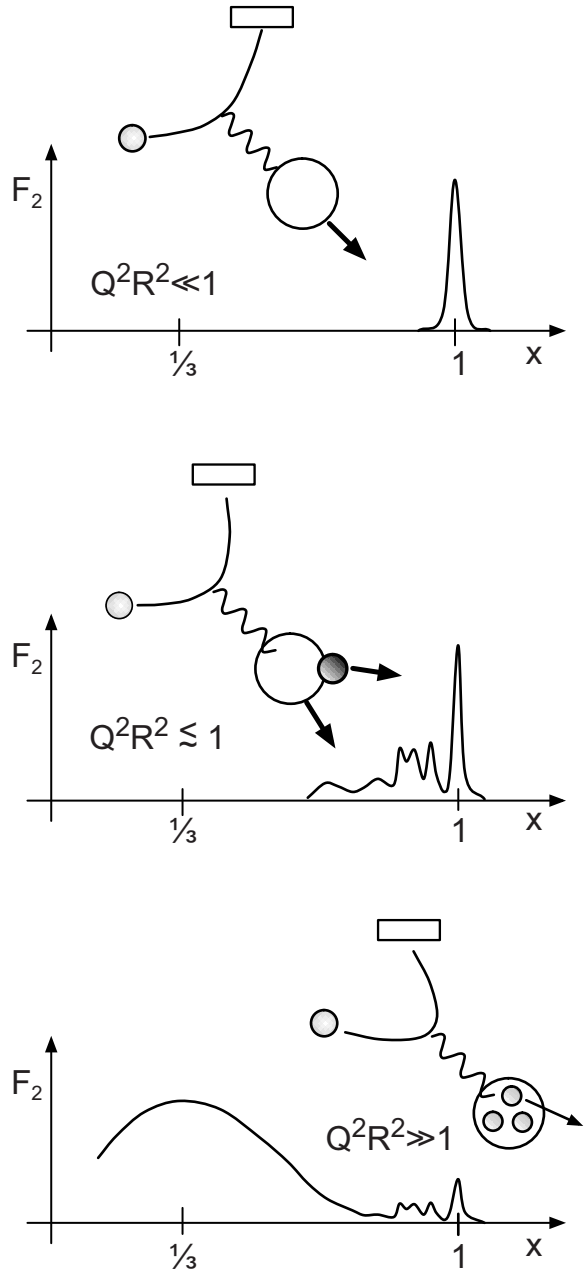

Figure 3.5: The transition from elastic over inelastic to deep inelastic scattering with increasing Q^2 [22].

was formed from a struck quark of the proton of the target the beam interacted with. We want to avoid the detection of hadrons formed from the quarks of the remnant of the target. Imposing a high z means that this hadron has to carry a high amount of energy, and therefore it is more likely that this energy comes from the energy available from the virtual photon that mediated the interaction. Sometimes also the restriction of $z < 0.7$ is established. This rejects scattering events in a region dominated by exclusively produced vector mesons, and exclusive processes are not analysed in the current analysis.

6. $\mathbf{2}\text{GeV} < \mathbf{P_h} < \mathbf{15}\text{GeV}$. This cut on the momentum is only a consequence of the RICH acceptance. The data collected from the rings generated on the aerogel and the gas only include this momentum range (See figure 2.2).

3.3 Monte Carlo Simulations

A Monte Carlo generator is a suite of physics generators that simulate different physics processes. These generators are the tool used to try to describe the observed kinematic distributions as well as possible. The event generation at HERMES is developed in two different steps. First of all a set of programs - Generator Monte Carlo (GMC) - is used in order to generate events. Once you have your generated events, the second step consists on simulating the HERMES experiment - HERMES Monte Carlo (HMC) - which tries to reconstruct the initial conditions.

GMC - Generator Monte Carlo

As explained on the previous paragraph, these kind of Monte Carlo generators are used to generate final events of different physical processes. That is the case for example of the gmc_dis, gmc_DVCS, PYTHIA or gmc_trans generators. At HERMES very different generators are used depending on the necessities of the analysis one wants to develop. They generate DIS events, DVCS, SIDIS events, and SIDIS events on transversely polarized beams respectively. In this thesis both PYTHIA and gmc_trans have been used.

Basically what these programs do is to generate random vertex events following different physical processes, giving as a final result a kinematical description of the vertex for each generated event, as well as information about the particles present in the final state. This means that we can say they also compute the

hadronization of the events, which is crucial to know what happens with the struck quarks.

There are several ways of generating the event normalization. The events must be generated according to the cross section (probability) each event has. The bigger the cross section, the higher the probability. So, somehow this has to be reflected on the events generated. However, there is not just a single way of achieving this. These probabilities can be obtained whether repeating the process several times or simply giving each process a proper weight. These concepts can be better understood through the following example. Let's suppose we have a die with six faces but only two available numbers. One of the faces has a six written on it, whereas the other five faces have a one. So, the situation can be evaluated as follows; you throw the dice several times, and each time you generate events according to the cross section. Somehow this faces are just telling you what reality is, without any previous intervention. This is the case of gmc_trans generator. Now let's imagine that we have a coin and we want to obtain the same binomial distribution we obtained in the previous example. Each of the two faces of the coin, the proper face and the tail, have in principle a probability of one half of appearing. Now we want to weight each of these outputs with a different probability factor, i.e. we multiply the face by one sixth and the tail by five sixths. Done this, we arrive to the same binomial distribution we had in the previous case.

HMC - HERMES Monte Carlo

Until now we have just generated the events, and we have the final result of the whole process. However, now that the particle is created we have to go through its path till the beginning and generate how it crossed the detector. Using different geometry files available at HERMES, the probability of every single step has to be modelled. The HMC generator simulates a perfect detector. Hence, several imperfections such as attenuation or structural and geometrical accidents have to be included during the digitalization process.

GMC_Trans Monte Carlo

The GMC_Trans Monte Carlo is used in this thesis in order to check the methodology used to analyse the data. GMC_Trans Monte Carlo was implemented to provide a better understanding of the transverse momentum distributions, both for unpolarised and polarised cases. In our analysis only unpolarised case is going to

be relevant. This simulator was conceived under the hypothesis that the transverse momentum dependence in Structure Functions is Gaussian. A better explanation of this is collected in the section 1.2.1[11].

4

Results

In this chapter the different procedures and results obtained will be presented. The idea of studying the intrinsic transverse momentum dependence of the quarks inside the proton has been developed in several different ways along this thesis. The methodology followed for this purposes is the same in the very basis, even if the way of extracting the values is different for each procedure. The ultimate idea is the extraction of the relevant parameters from the fit of real data using some different models.

First of all, a comparison between SIDIS real data (in acceptance) and PYTHIA Monte Carlo data in the same conditions is studied in order to evaluate how well both data sets agree. Afterwards, data corrected for acceptance is analysed using function (4.2), and the results obtained from these calculations are inserted into two different algorithms that will allow us to obtain $\langle \vec{P}_{h\perp}^{2} \rangle$. If the gaussian behaviour hypothesised for it is correct, both sets of results would have to agree. In order to check the validity of the method the full analysis chain is first applied on gmc_trans data, which are generated according to the Gaussian Ansatz.

4.1 Comparison of PYTHIA and Data

A comparison between the $\langle \vec{P}_{h\perp}^{2} \rangle$ distribution extracted from the PYTHIA Monte Carlo under acceptance and that of real HERMES data (under acceptance) sorted with the procedure explained in chapter 3.

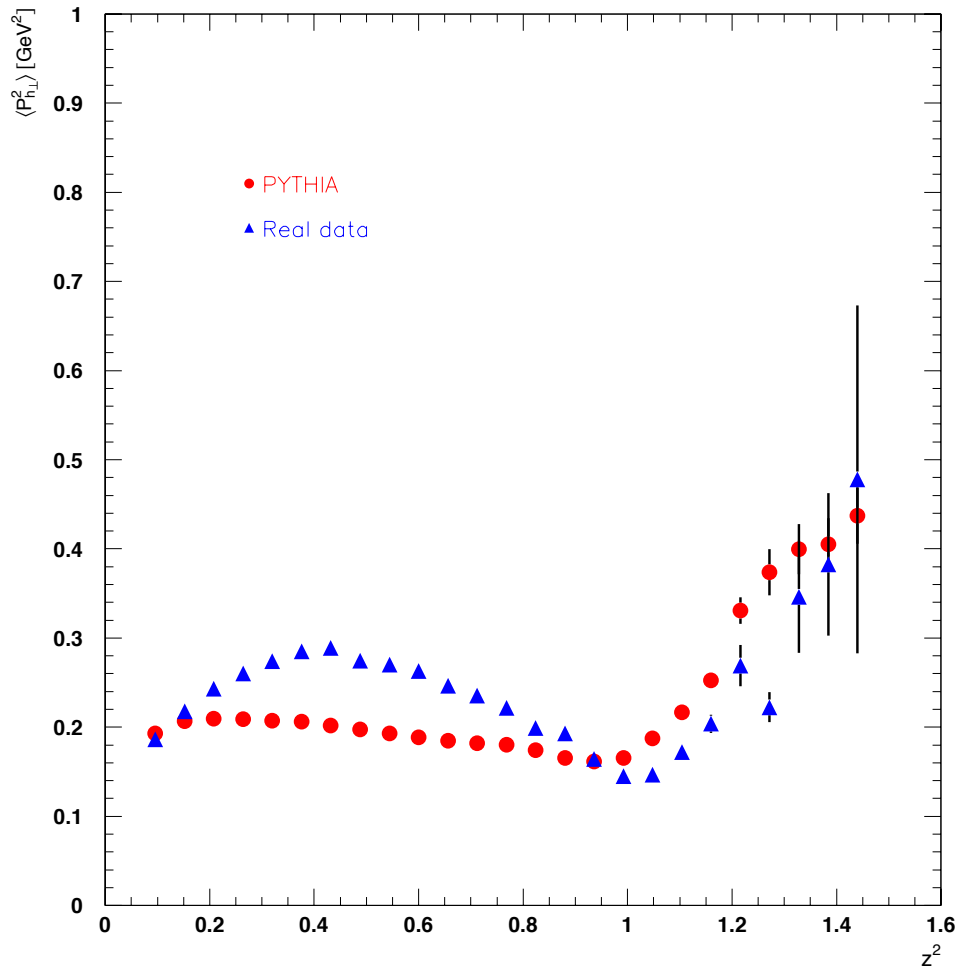

Figure 4.1: HERMES data and PYTHIA Monte Carlo comparison for positive pions, π^+. The red points represent the Monte Carlo output and the blue points the real data.

Pions

An important disagreement between the two sets of data can be easily perceived for both π^+ and π^-. See figures 4.1 and 4.2.

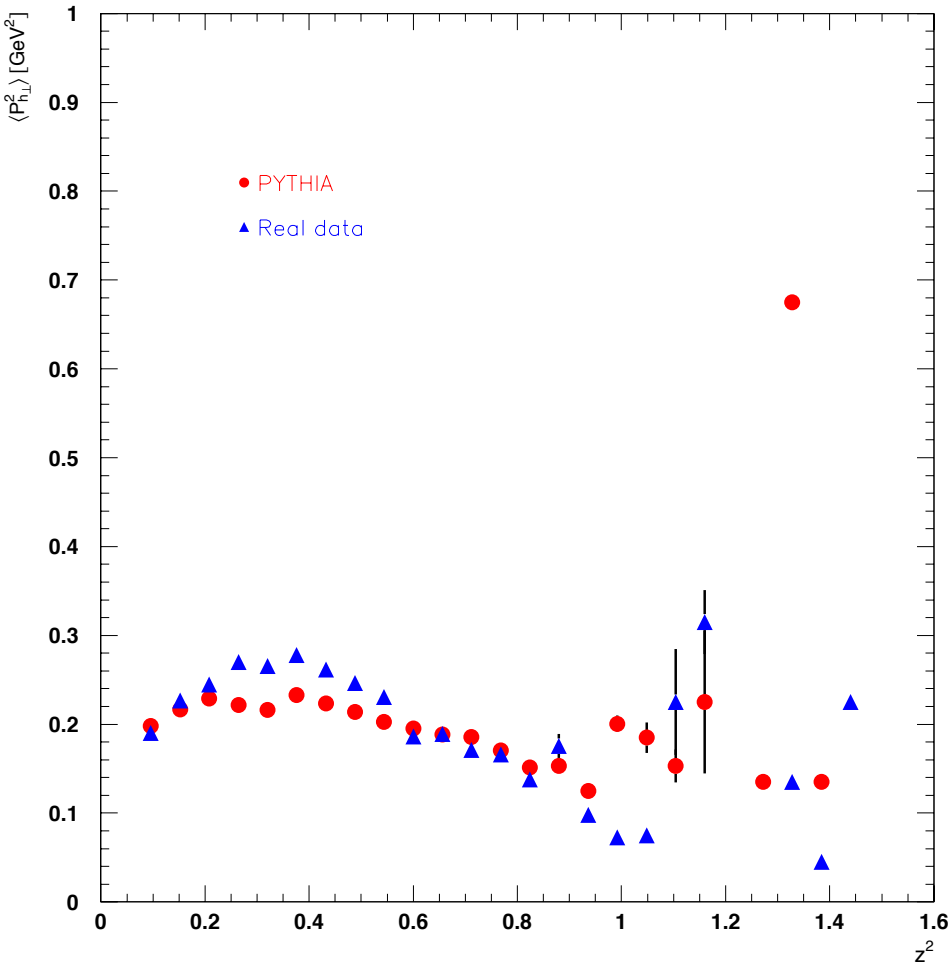

Figure 4.2: HERMES data and PYTHIA Monte Carlo comparison for negative pions, π^-. The red points represent the Monte Carlo output and the blue points the real data.

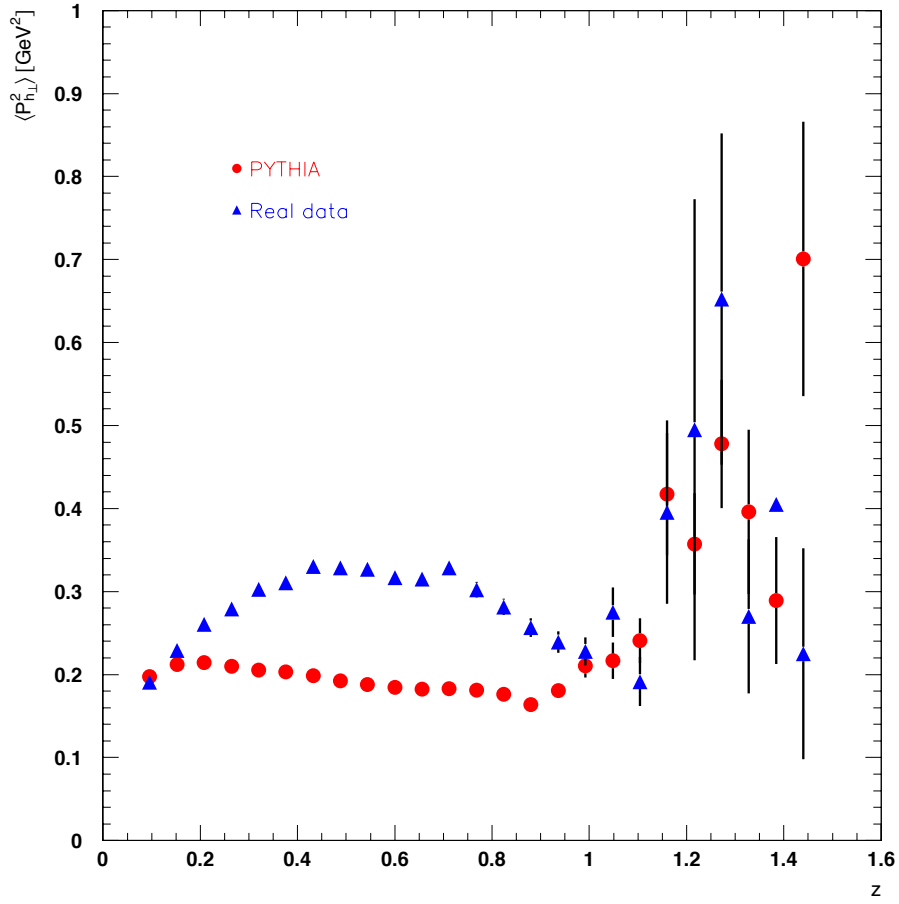

Figure 4.3: HERMES data and PYTHIA Monte Carlo comparison for positive kaons, K^+. The red points represent the Monte Carlo output and the blue points the real data.

Kaons

Protons and kaons have no valence quarks in common. This explains why the modelling of the hadronisation process is more difficult as two quarks have to be produced on the fragmentation process. The results of the real data and PYTHIA Monte Carlo comparison can be observed in figures 4.3 and 4.4.

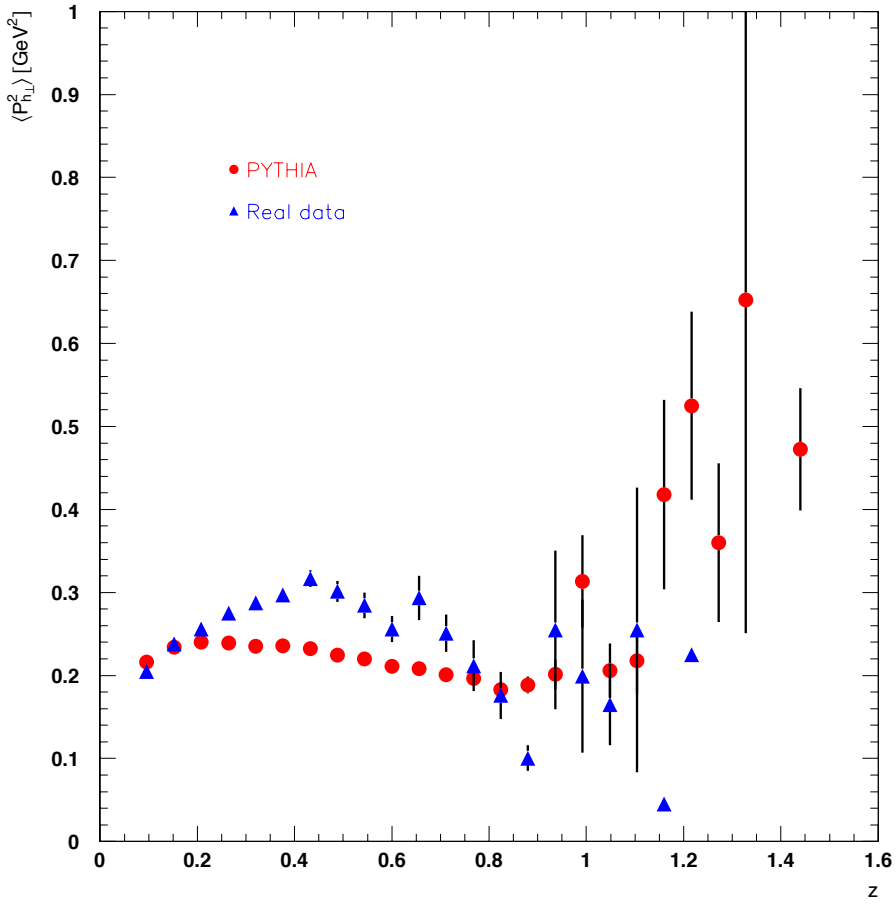

Figure 4.4: HERMES data and PYTHIA Monte Carlo comparison for negative kaons, K^-. The red points represent the Monte Carlo output and the blue points the real data.

4.2 Born data set and GMC_Trans

In section 4.1 both data sets are affected by acceptance effects. In order to eliminate the artifacts arising from acceptance effect, the study developed in the current analysis is based on a data set unfolded for radiative and instrumental smearing effects. In figure 4.5 the difference between data in acceptance and unfolded data can be easily appreciated. The unfolded data presents a more flat shape compared to the pronounced curvature of data in acceptance as a consequence of the large acceptance effects that characterises the HERMES spectrometer.

As a first step, a set of data extracted from the GMC_Trans Monte Carlo simulator will be analysed using the same two procedures that will later be used for the unfolded real data. With this step we aim to test that the analysis method used to check if the transverse momentum dependence is Gaussian works fine. The expectation is that if it this is the case, both procedures should give the same results.

4.2.1 GMC_Trans

"Mean" method

The "mean method" consists on the direct extraction of the mean value of the $P_{h\perp}^2$ distribution from the gmc_trans data. This mean value will be identified as $\langle P_{h\perp}^2 \rangle$. From this results we plan to extract the average squared quark transverse momentum $\langle p_T^2 \rangle$ through a fitting procedure assuming the following relation between the different transverse momentum compontents:

$$\langle P_{h\perp}^2(z) \rangle = \langle p_T^2 \rangle z^2 + \langle K_T^2 \rangle. \tag{4.1}$$

In this expression, the term $\langle p_T^2 \rangle$ corresponds to the intrinsic transverse momentum of the quark inside the proton and $\langle K_T^2 \rangle = \langle k_T^2 \rangle z^2$ to the transverse momentum obtained by the quark during the fragmentation process. z is the fraction of the momentum taken by the hadron from the quark during hadronisation process.

In order to extract information about $\langle p_T^2 \rangle$ and $\langle K_T^2 \rangle$, the mean values of $P_{h\perp}^2(z)$ for different z bins are were fitted using the following fit function, also known as the "Hashi parametrisation":

$$f \equiv \gamma z^2 + \delta[(1-z)^\beta z^\alpha] \tag{4.2}$$

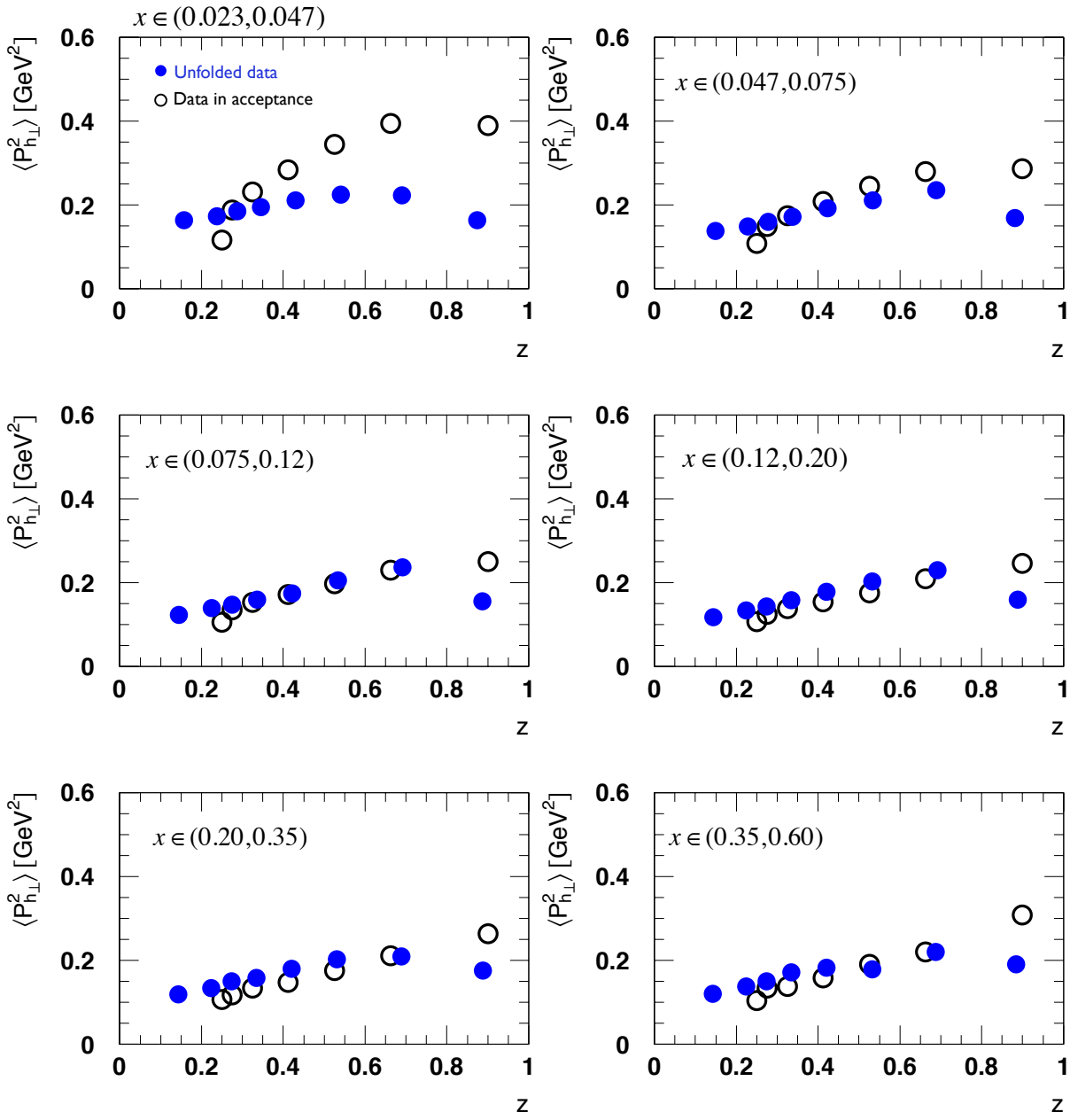

Figure 4.5: Mean value of $P_{h\perp}^2$ as a function of z for π^+ for real data (in acceptance) and unfolded data.

Here α, β, γ and δ are the parameters that will be extracted in the fit. In particular, γ corresponds to the value of $\langle p_T^2 \rangle$.

The z dependence of average $P_{h\perp}^2$ can be seen in figure 4.6. The resulting parameters from the fit are presented in table 4.1.

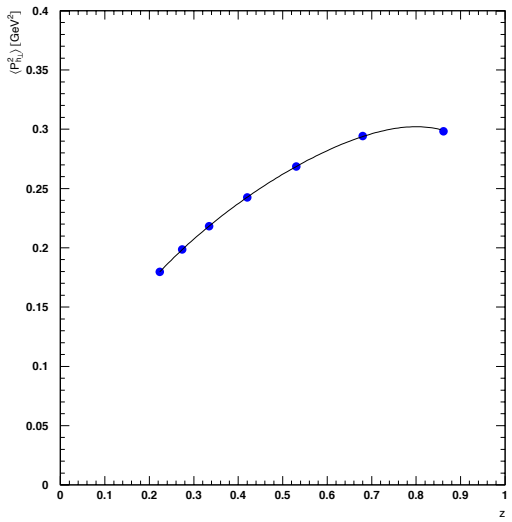

Figure 4.6: Fit of $\langle P_{h\perp}^2 \rangle$ from GMC_Trans using the "Mean" method for π^+ is presented.

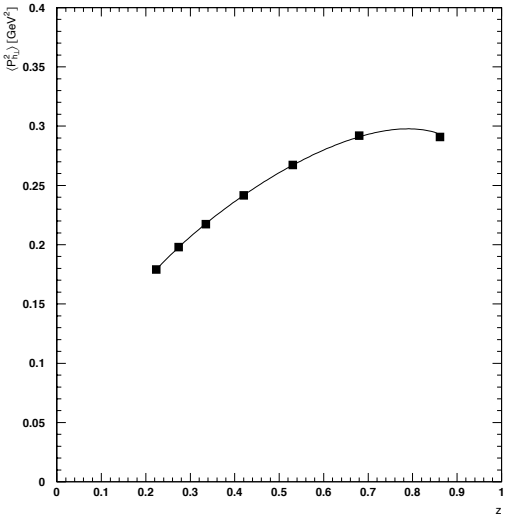

Figure 4.7: Fit of $\langle P_{h\perp}^2 \rangle$ from GMC_Trans using the "Gauss" method for π^+ is presented.

"Gauss" method

In this case the procedure is the following: the transverse momentum distribution is fitted using the hypothesis of having a Gaussian shape (see eq. (4.3)). One of the parameters of this fits corresponds to $\langle P_{h\perp}^2 \rangle$. These values are extracted and then plotted against z and fitted using the same fit presented on the previous subsection as "Hashi parametrisation".

So, first of all, the mean value of the squared transverse momentum is fitted via:

$$g \equiv \frac{N}{\langle P_{h\perp}^2 \rangle} e^{-\frac{P_{h\perp}^2}{\langle P_{h\perp}^2 \rangle}}. \tag{4.3}$$

	α	β	γ	δ
"Mean" method	0.528±0.045	0.365± 0.440	0.150±0.192	0.417±0.051
"Gauss" method	0.533±0.014	0.339±0.137	0.129±0.064	0.418±0.015
Input parameters	0.536	0.366	0.144	0.422

Table 4.1: Fit results for the parameters from the "Mean" method and the "Gauss" method applied over GMC_Trans data, where the input parameters represent the "Hashi set".

The Gaussian fit of the transverse momentum distribution is plotted in figures 4.8 and 4.9 for the 7 different z bins analysed. From each of these bins, the mean transverse momentum is extracted (see fig. 4.7) and these mean values are used to extract the parameters α, β, γ and δ using (4.2). The results are shown in table 4.1.

Finally, a plot comparing both fits is shown in figure 4.10. Generally, a good agreement between the two methods can be observed, though there is a slight shift at large z.

4.2.2 Born data set

Once we have checked the methodology using the GMC_Trans presented data, we repeat the procedure for real unfolded data and execute the same procedure. First of all, the so-called "mean method" is applied, and afterwards, its results are compared to the ones obtained using the "Gauss method".

"Mean" method

Here again the fit (4.2) is applied to distribution of the mean values of $P_{h\perp}^2$ as a function of z. In the plots 4.11 and 4.12 the result of the fitting function are presented and in plots 4.13 and 4.14 the values of the parameters and their x-dependence can be appreciated.

The results are stored in tables 4.2 and 4.3.

"Gauss" method

In order to prove our hypothesis of a Gaussian dependence of the transverse momentum, the extraction of the value of $\langle p_T^2 \rangle$ can be achieved from the transverse

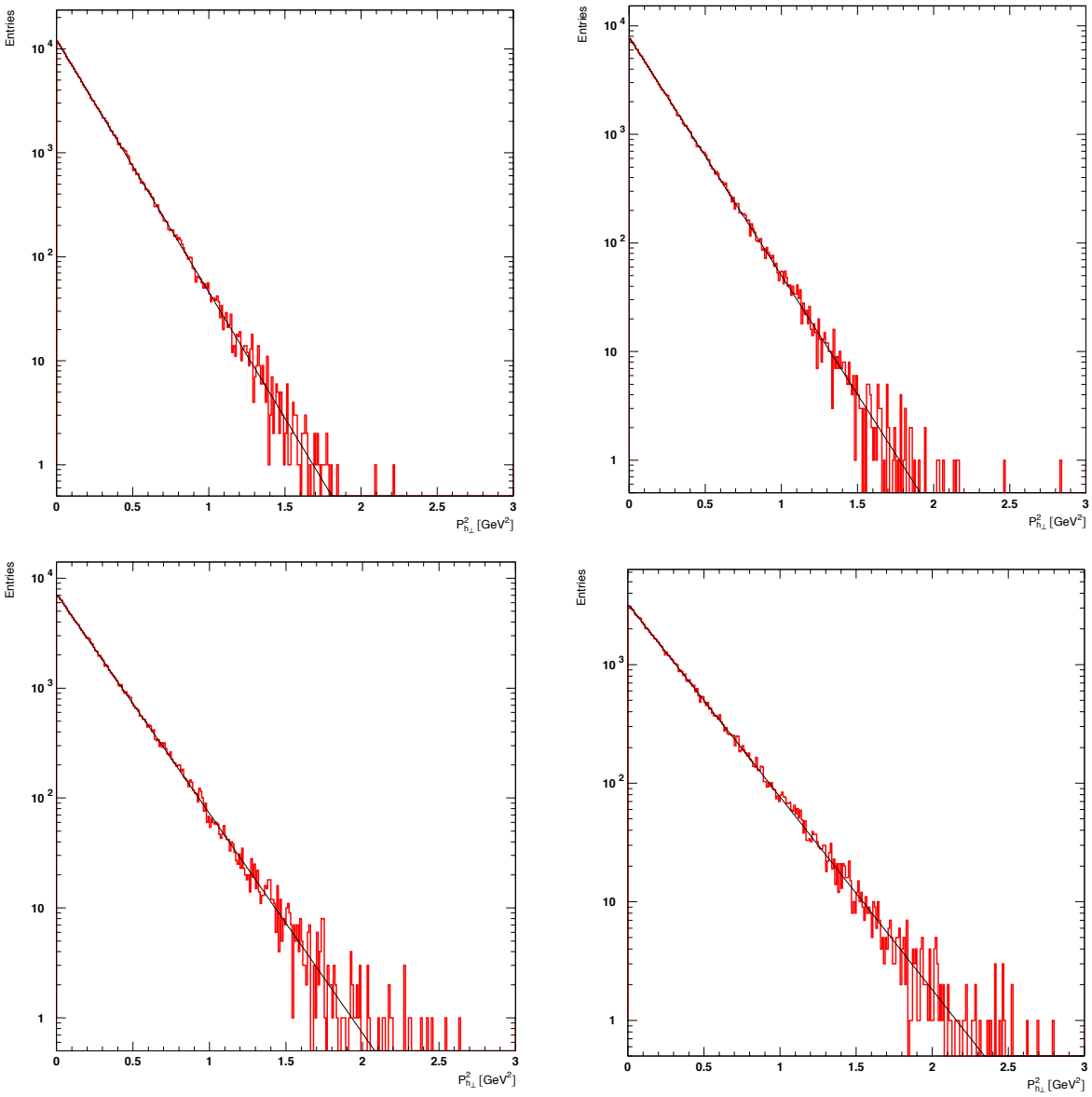

Figure 4.8: Gaussian fit of $P^2_{h\perp}$ distribution for π^+ over Monte Carlo data for the first four z bins.

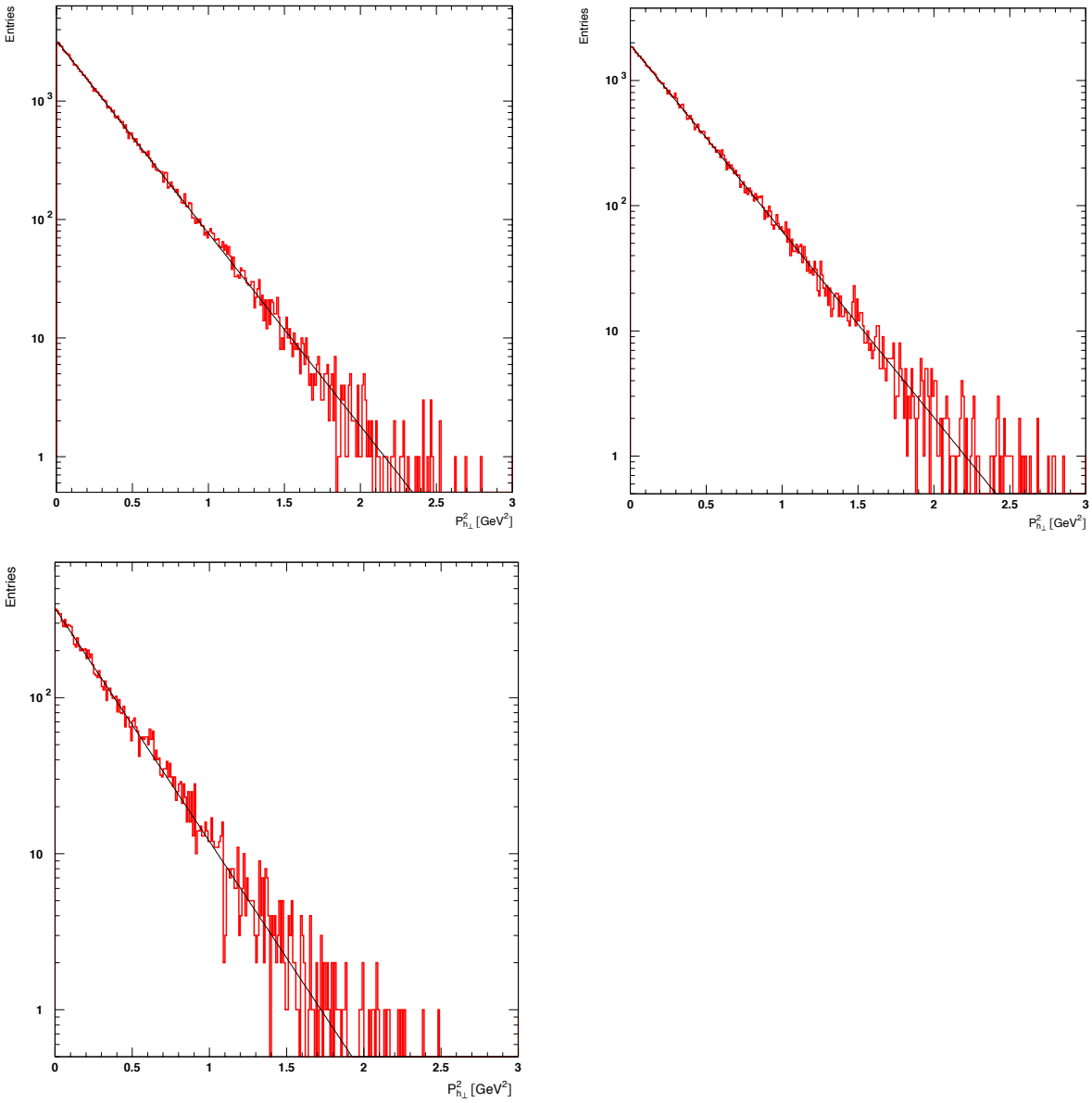

Figure 4.9: Gaussian fit of $P^2_{h\perp}$ distribution for π^+ over Monte Carlo data for the three last z bins.

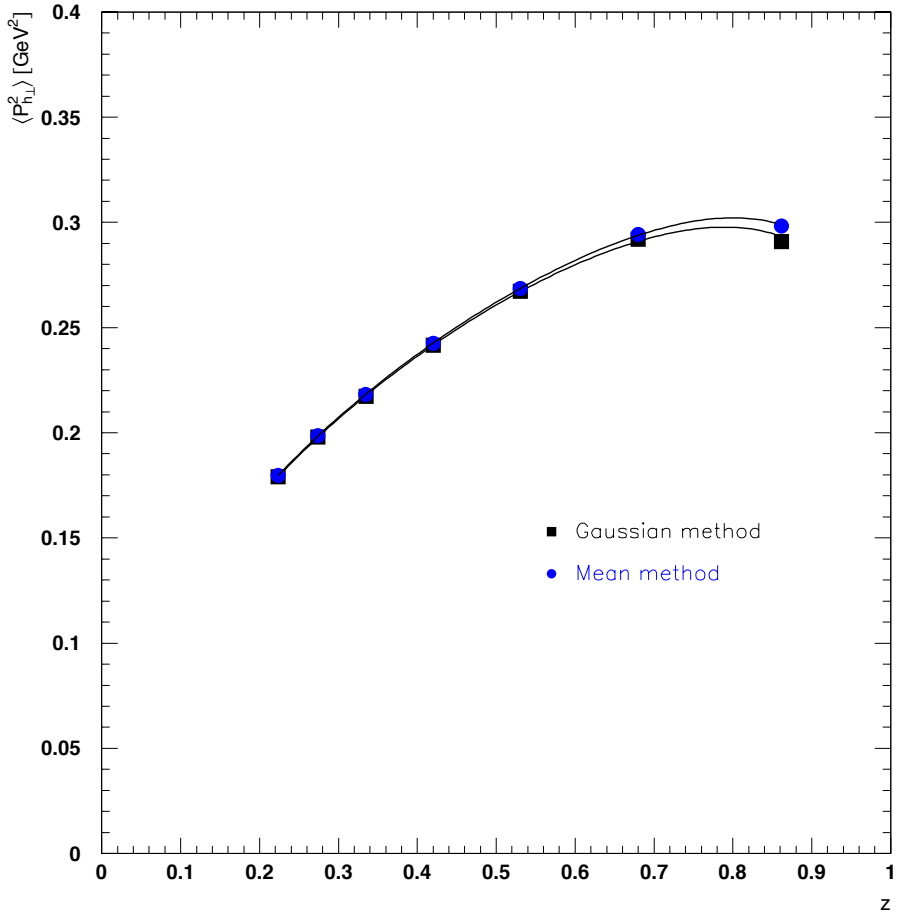

Figure 4.10: Fits of$\langle P_{h\perp}^2 \rangle$ for π^+ over Monte Carlo data using both methods.

	α	β	γ	δ
x_0	0.268±0.012	0.233 ± 0.073	0.092±0.192	0.270±0.007
x_1	0.240±0.010	0.215±0.065	0.183±0.192	0.019±0.007
x_2	0.217±0.017	0.284±0.137	0.238±0.032	0.189±0.009
x_3	0.294±0.012	0.238±0.091	0.199±0.025	0.207±0.006
x_4	0.367±0.016	0.266±0.107	0.132±0.033	0.243±0.009
x_5	0.487±0.128	0.928±0.667	0.240±0.113	0.340±0.110

Table 4.2: Fit results for the parameters from Mean method applied over unfolded real data for the case of π^+, where x_i are the different x-bins used.

	α	β	γ	δ
x_0	0.360±0.012	0.272 ±0.069	-0.007±0.028	0.325± 0.008
x_1	0.298± 0.011	0.240±0.064	0.089±0.021	0.251±0.006
x_2	0.296±0.010	0.245±0.074	0.102±0.023	0.234± 0.006
x_3	0.348±0.013	0.261± 0.093	0.086±0.030	0.251± 0.008
x_4	0.387±0.019	0.251±0.116	0.080±0.042	0.272±0.011
x_5	0.603±0.151	1.403±0.718	0.336±0.100	0.483±0.183

Table 4.3: Fit results for the parameters from Mean method applied over unfolded real data for the case of π^-, where x_i are the different x-bins used.

momentum distribution using the fit presented in equation (4.4)

$$g \equiv \frac{N}{\langle P_{h\perp}^2 \rangle} P_{h\perp} e^{-\frac{P_{h\perp}^2}{\langle P_{h\perp}^2 \rangle}} \tag{4.4}$$

where N and $\langle P_{h\perp}^2 \rangle$ are the two parameters to be extracted, and N represents the total amount of entries for the considered bin in z. $P_{h\perp}^2$ is the independent variable of the function.

Here the fit function (4.4) is slightly different than the one presented in the GMC_Trans section. This is due to the way data is extracted and stored in each case. For the GMC_Trans case, the cross-section is defined as $\frac{d^2\sigma}{dP_{h\perp}^2}$, whereas in the case of real unfolded data it is set as $\frac{d\sigma}{dP_{h\perp}}$. This simply implies that in the second case, a $P_{h\perp}$ has to be included in the functional dependence, and consequently, gives rise to the extra $P_{h\perp}$ factor in equation (4.4). The output of this fit can be seen in figures 4.15 and 4.16.

	α	β	γ	δ
x_0	0.677±0.022	0.347 ± 0.152	0.153±0.093	0.532±0.024
x_1	0.191±0.023	0.308±0.191	0.307±0.085	0.371±0.020
x_2	0.360±0.033	0.284±0.285	0.411±0.091	0.258±0.023
x_3	0.342±0.038	0.198±0.373	0.413±0.128	0.243±0.025
x_4	0.419±0.049	0.281±0.270	0.345±0.103	0.296±0.031
x_5	0.495±0.240	-0.344±0.519	-0.181±0.817	0.343±0.165

Table 4.4: Fit results for the parameters from Gauss method applied over unfolded real data for the case of π^+, where x_i are the different x-bins used.

	α	β	γ	δ
x_0	0.832± 0.020	0.377 ±0.133	-0.110±0.110	0.734±0.028
x_1	0.639±0.021	0.337± 0.138	0.110± 0.080	0.495±0.021
x_2	0.519±0.025	0.306±0.170	0.164±0.081	0.391±0.021
x_3	0.518± 0.033	0.312±0.225	0.177± 0.106	0.388±0.028
x_4	0.523±0.055	0.293±0.326	0.205±0.166	0.405±0.046
x_5	0.570±0.157	0.339±1.054	0.166±0.575	0.469± 0.157

Table 4.5: Fit results for the parameters from Gauss method applied over unfolded real data for the case of π^-, where x_i are the different x-bins used.

Once the mean transverse momentum is extracted, the same fit as in the previous case based on equation (4.2) (see figures 4.17 and 4.18). Different α, β, γ and δ parameters are extracted, and then compared with the previous ones. Plots corresponding to these results can be seen in figures 4.19 and 4.20. The numerical results are collected in tables 4.4 and 4.5.

Finally, a comparison of the fit results obtained with both methods can be seen in the plots 4.21 and 4.22.

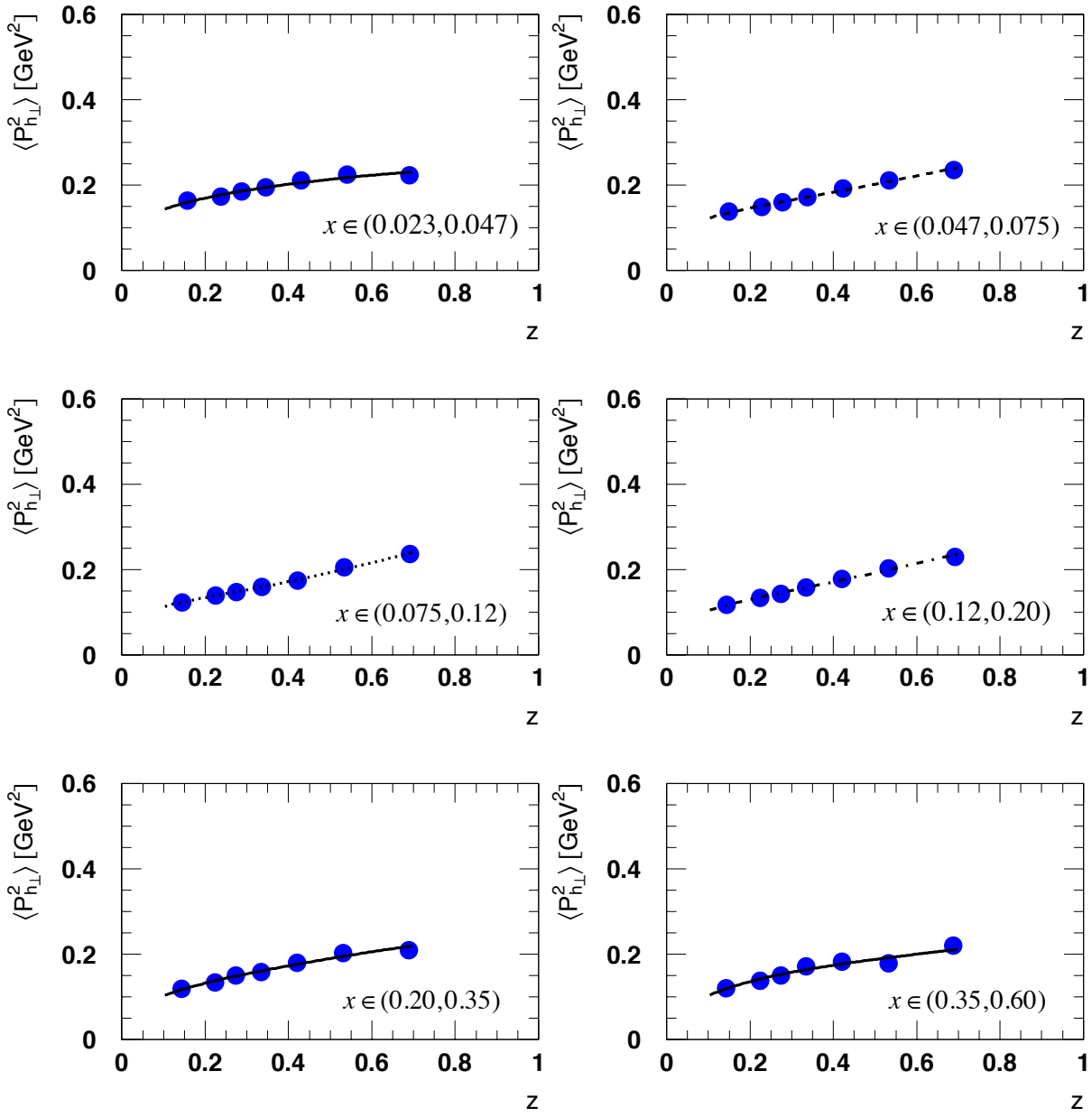

Figure 4.11: Fit of $\langle P_{h\perp}^2 \rangle$ using the "Mean" method over unfolded real data for π^+ is presented.

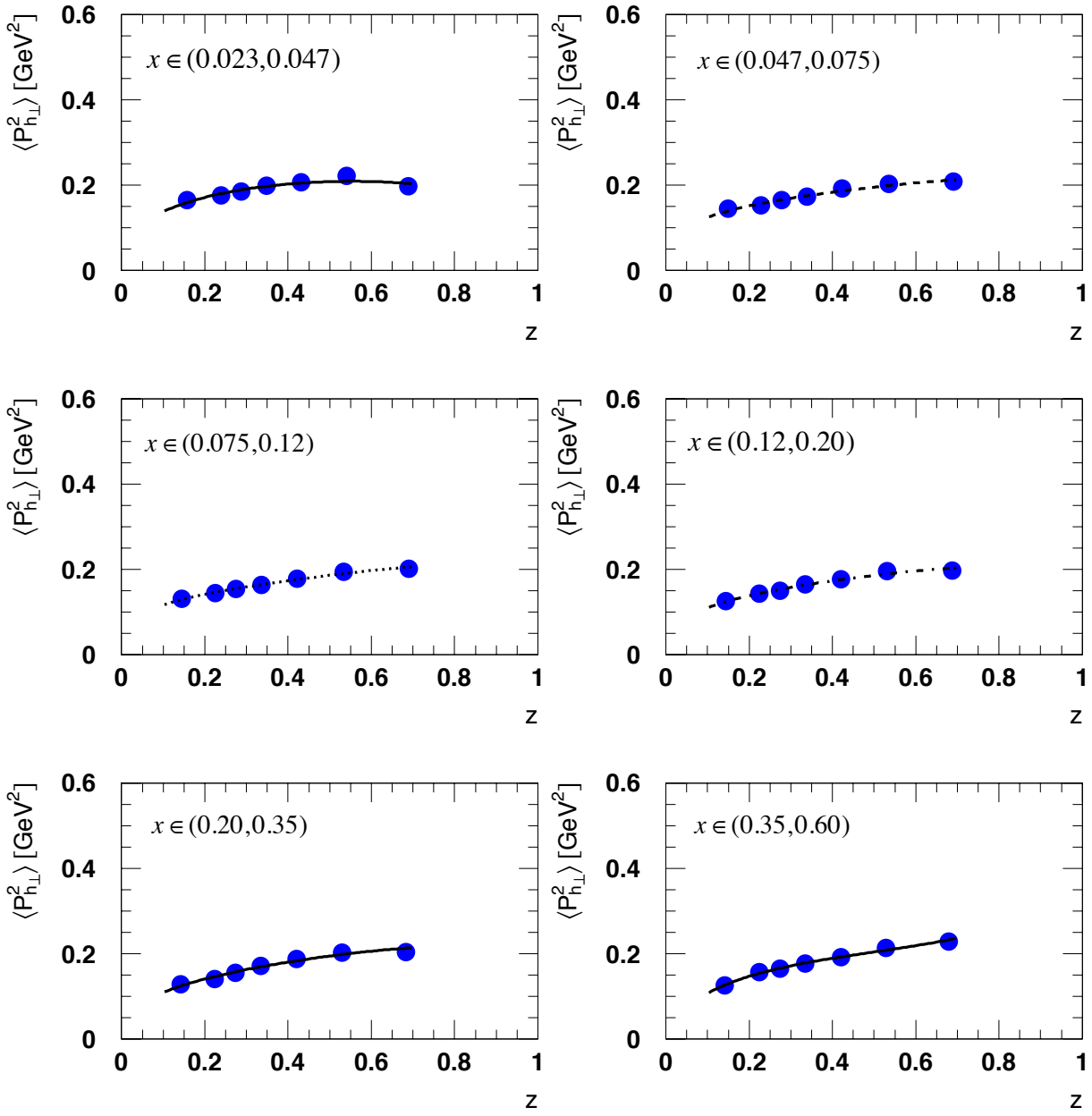

Figure 4.12: Fit of $\langle P_{h\perp}^2 \rangle$ using the "Mean" method over unfolded real data for π^- is presented.

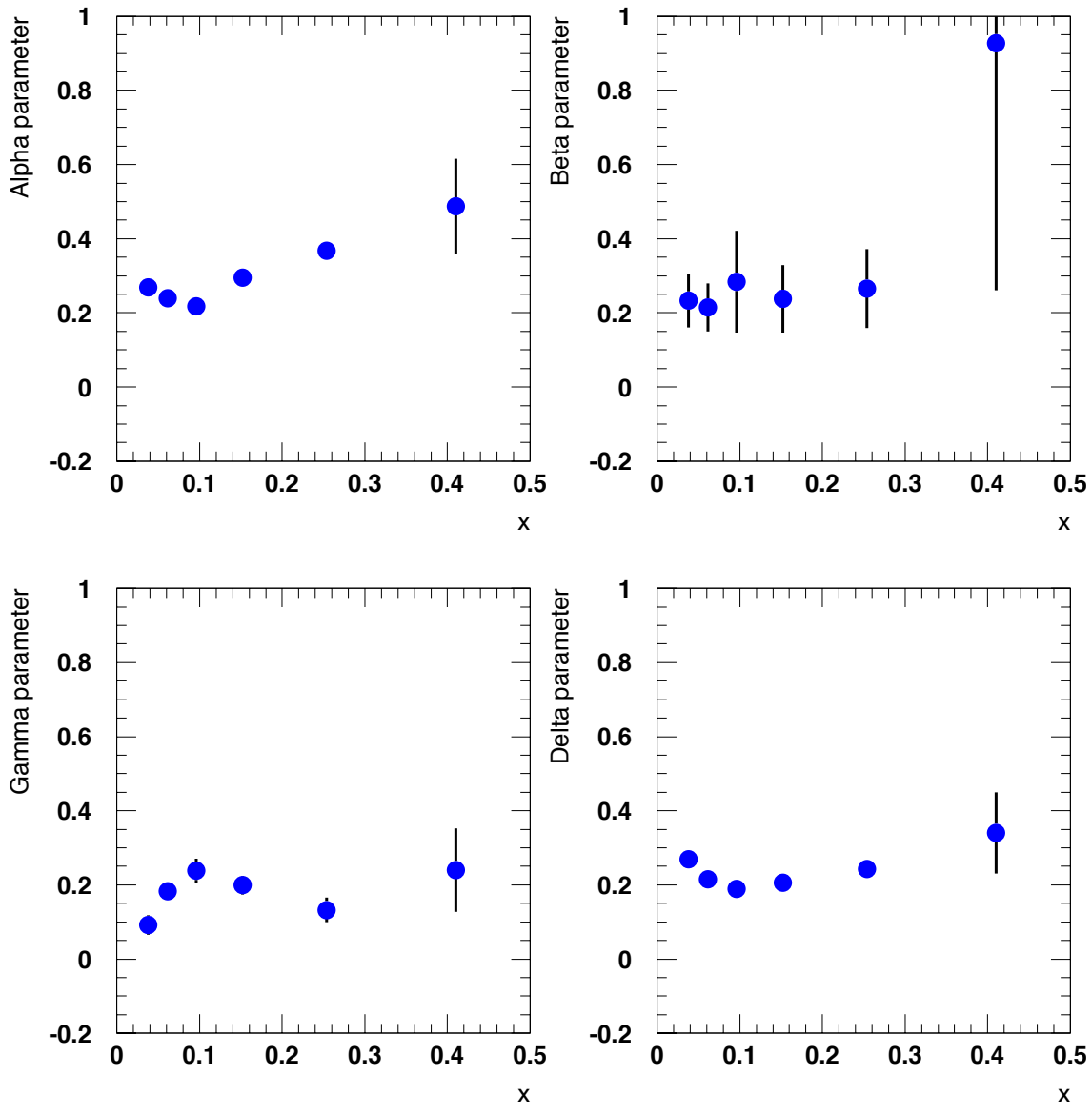

Figure 4.13: The dependence of the four fitted parameters with respect to x for the case of π^+.

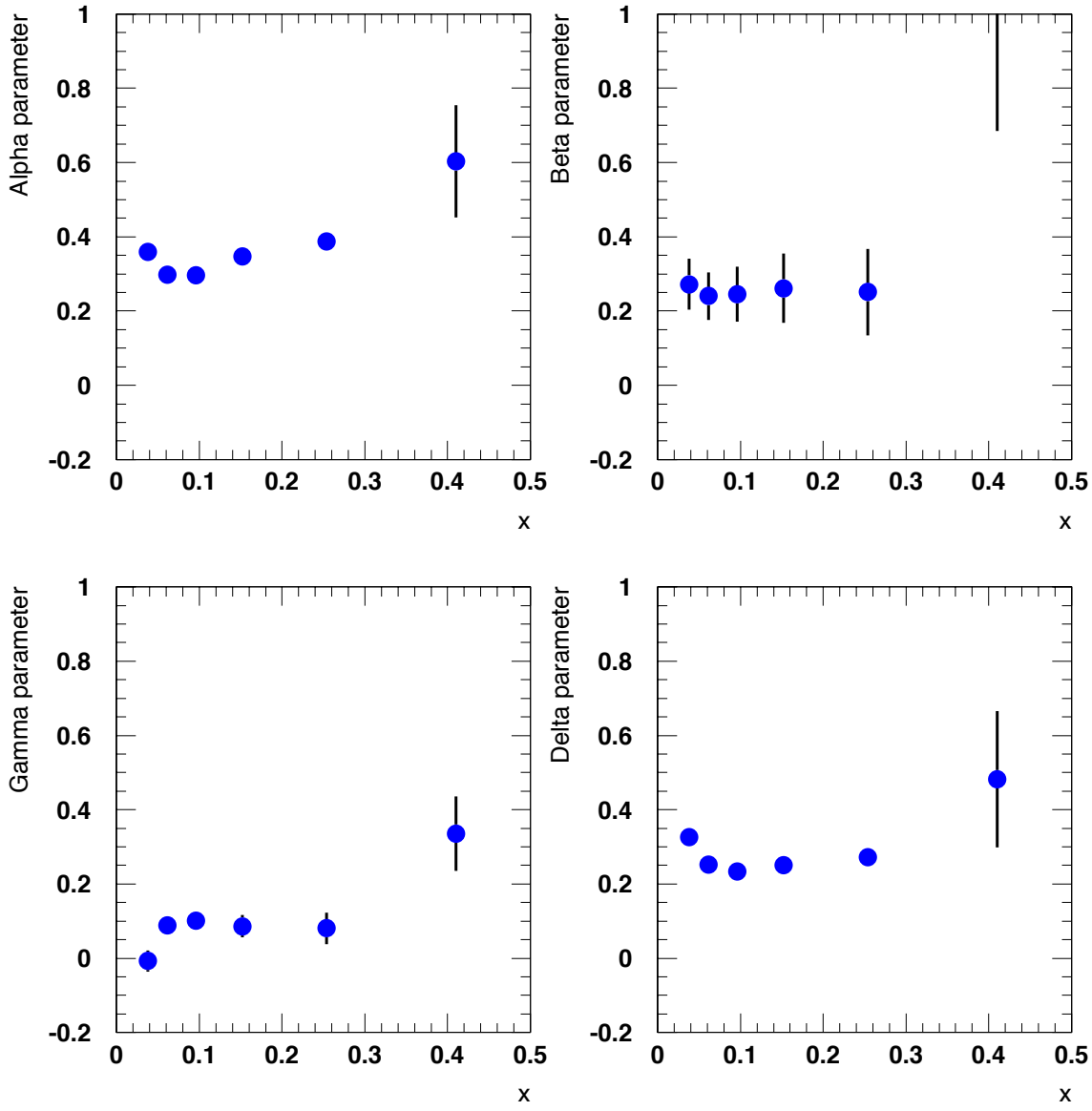

Figure 4.14: The dependence of the four fitted parameters with respect to x for the case of π^-.

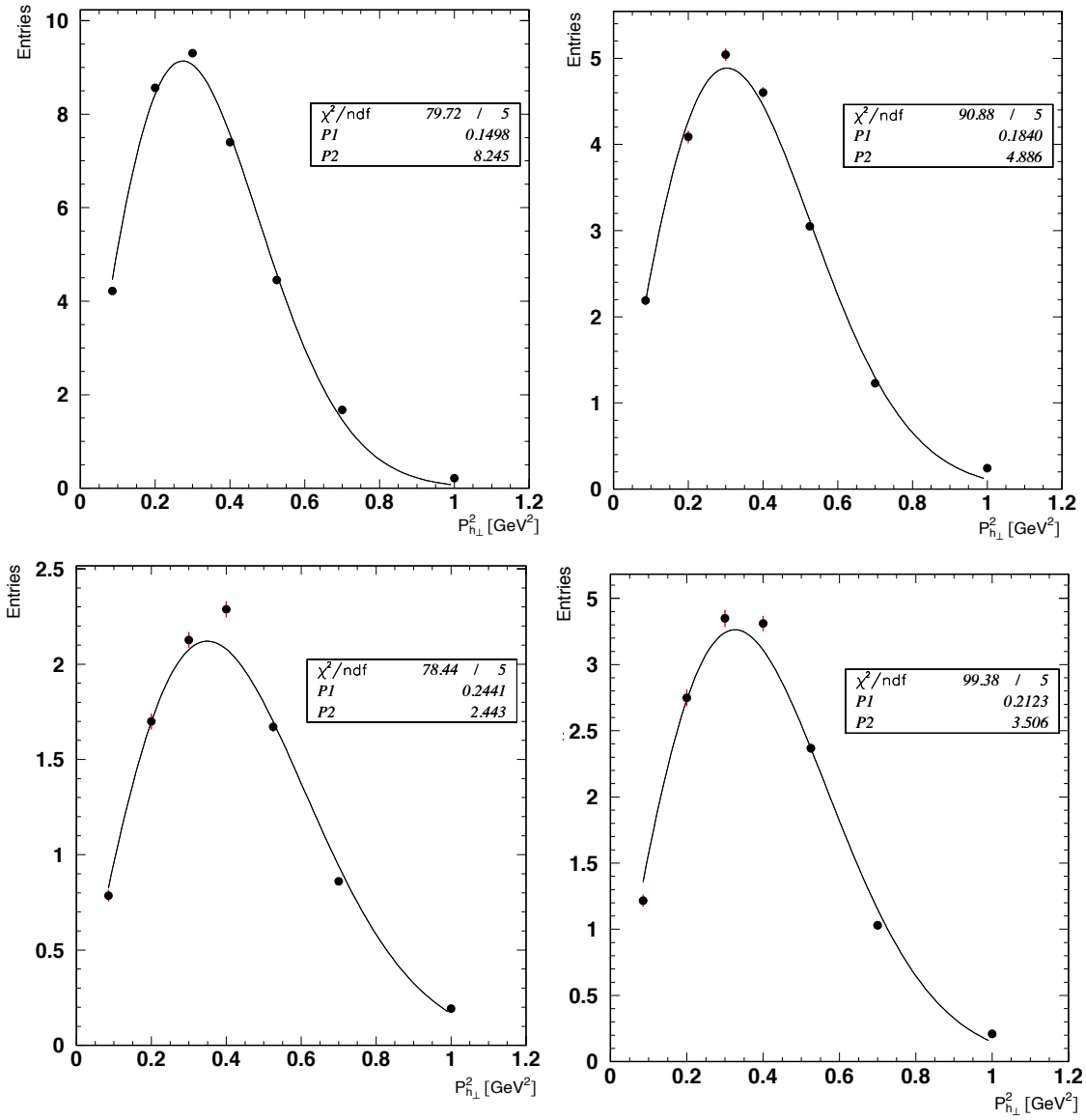

Figure 4.15: Gaussian fit of $P^2_{h\perp}$ for π^+ over unfolded real data for the first four z bins of the first x bin where P_1 and P_2 are the fitting output and χ^2 represents the accuracy of the fit.

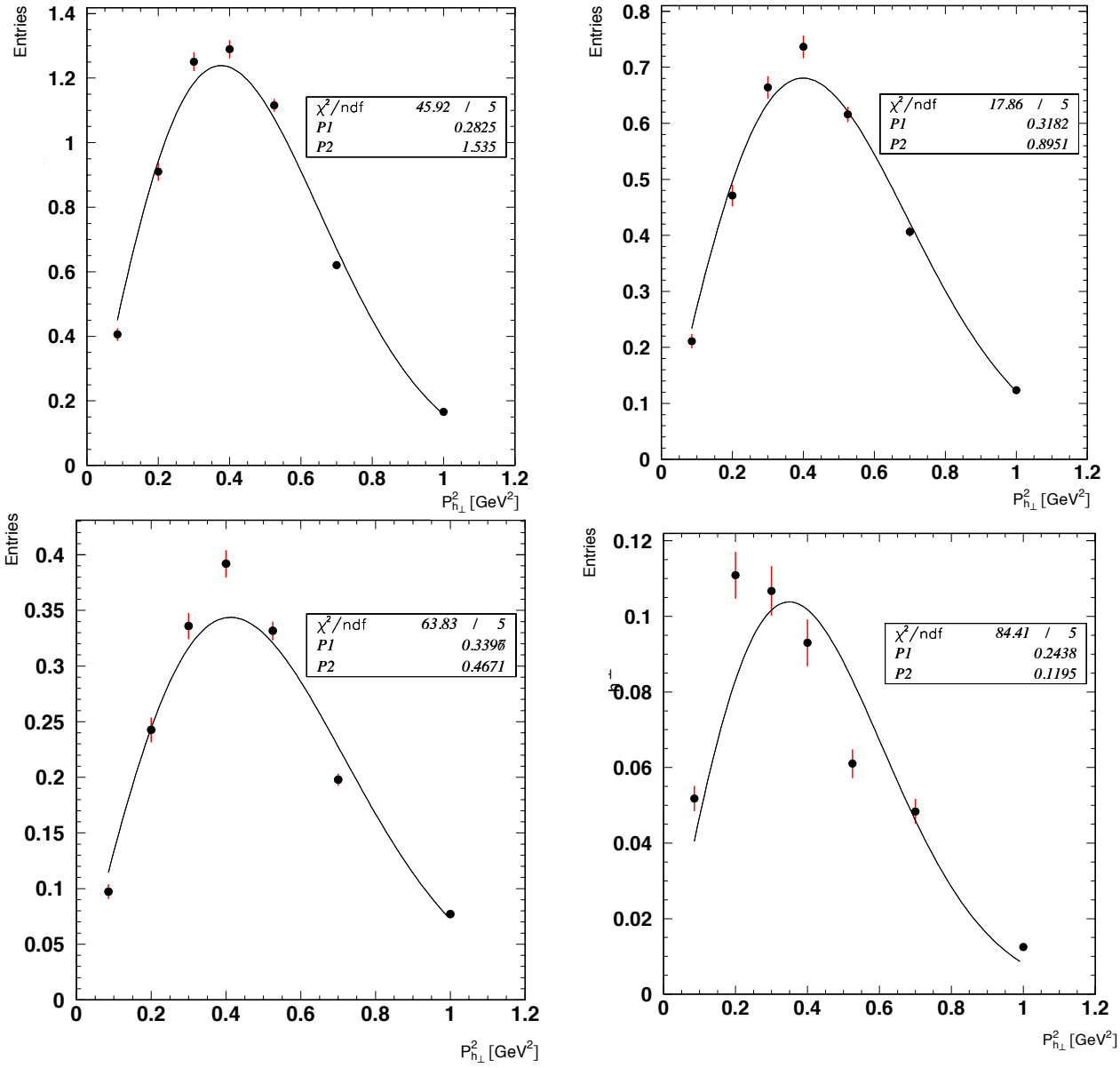

Figure 4.16: Gaussian fit of $P^2_{h\perp}$ for π^+ over unfolded real data for the last four z bins of the first x bin where P_1 and P_2 are the fitting output and χ^2 represents the accuracy of the fit.

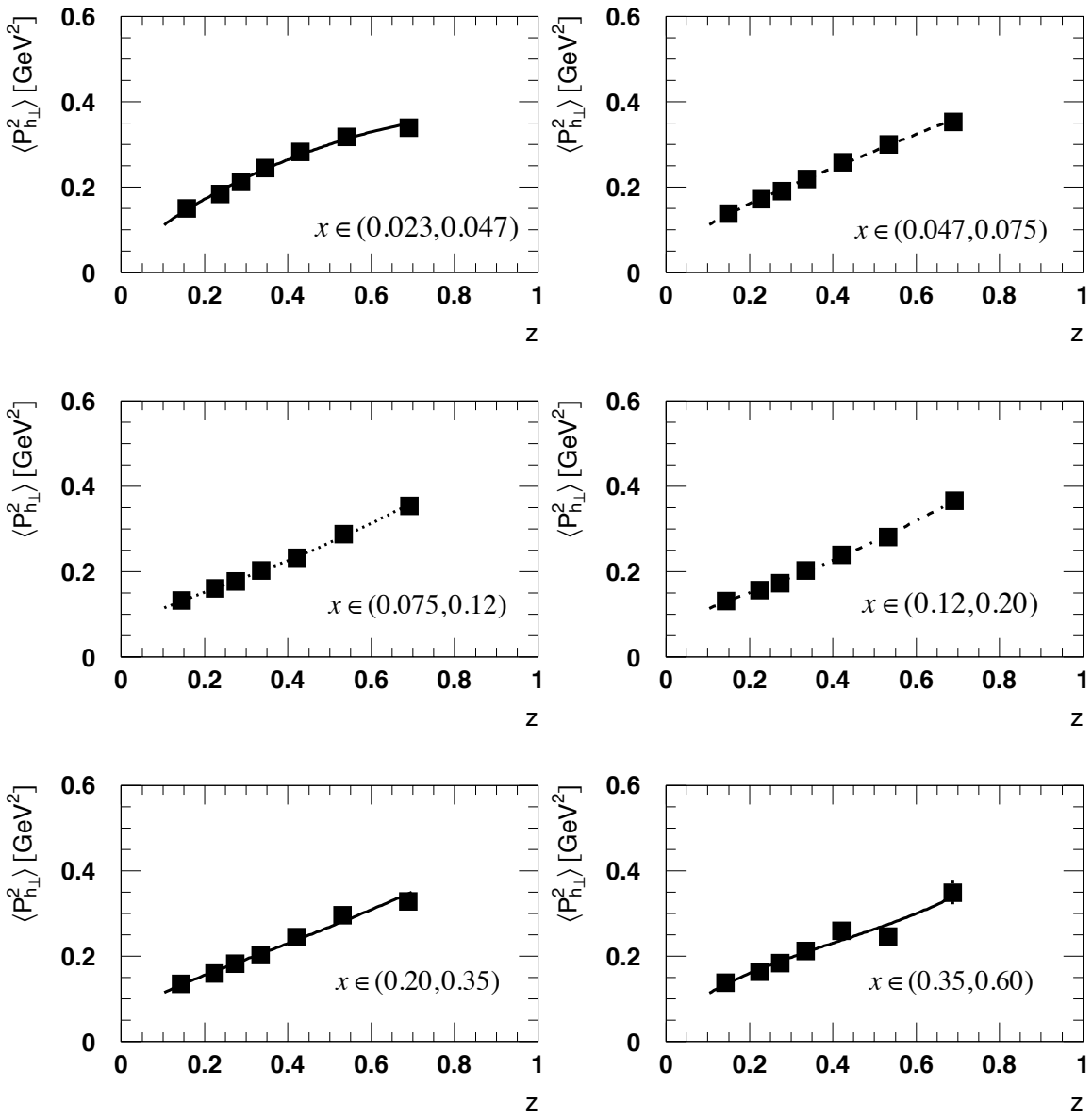

Figure 4.17: Fit of $\langle P_{h\perp}^2 \rangle$ using the "Gauss" method over unfolded real data for π^+ is presented.

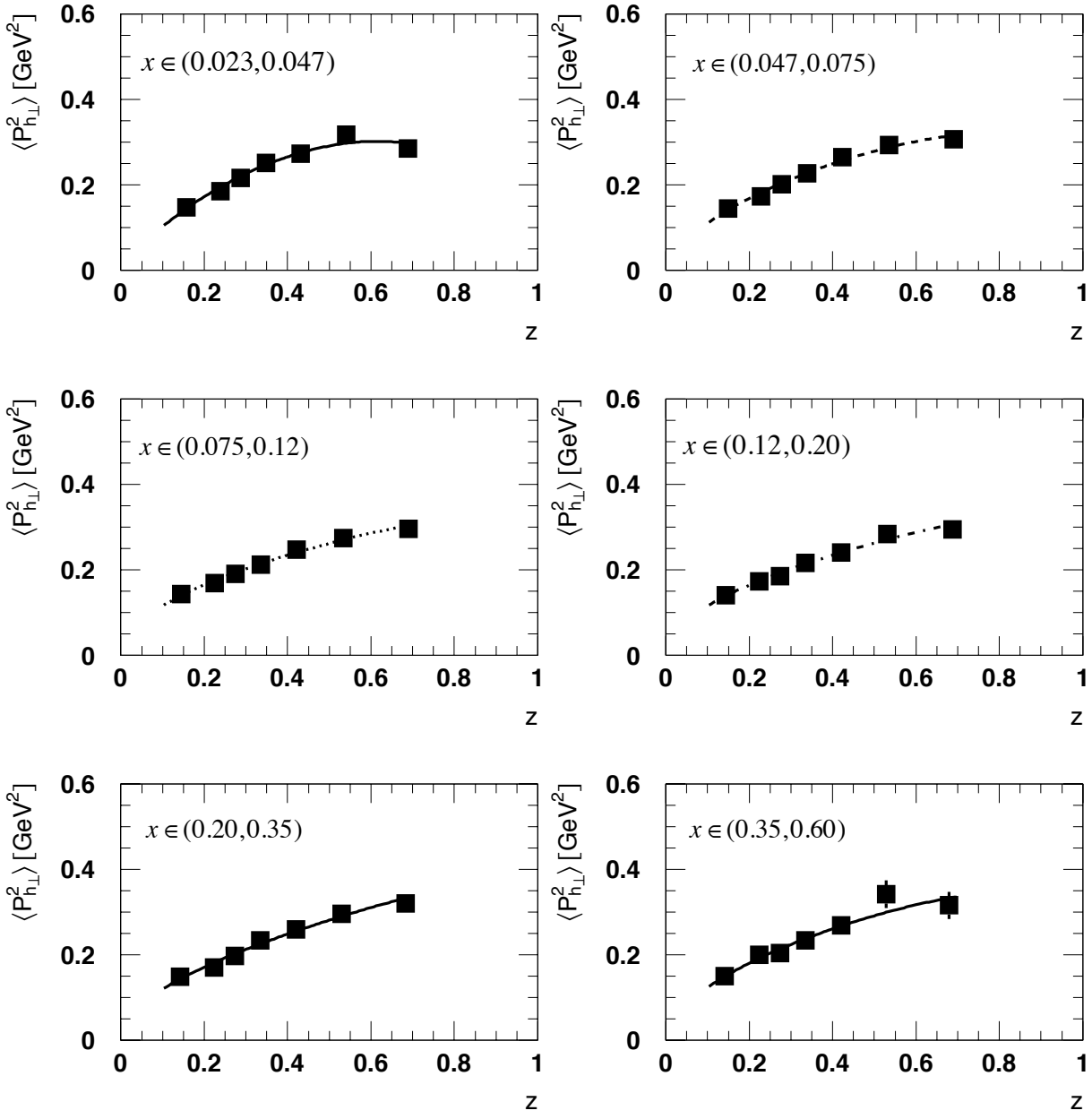

Figure 4.18: Fit of $\langle P_{h\perp}^2 \rangle$ using the "Gauss" method over unfolded real data for π^- is presented.

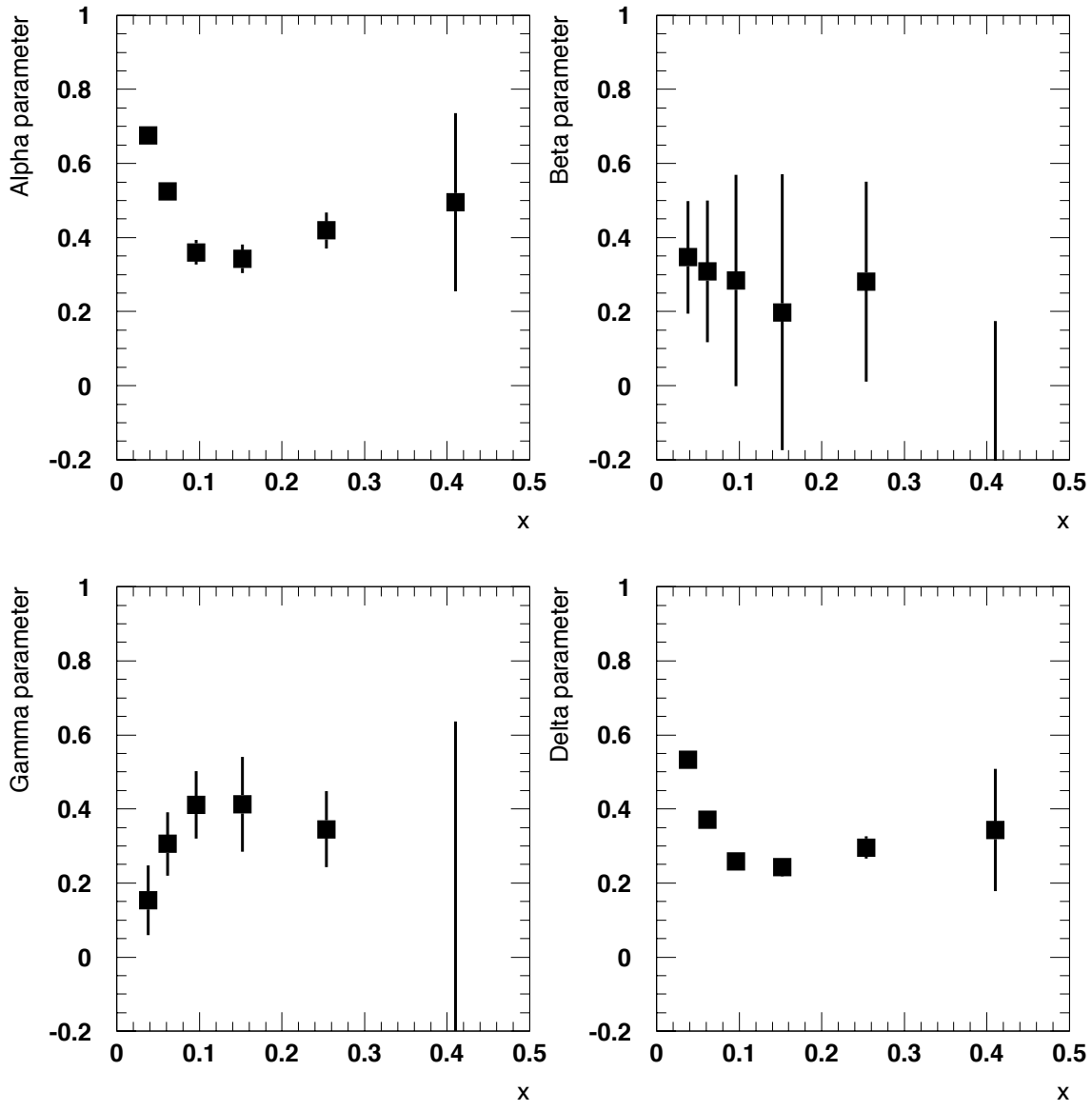

Figure 4.19: The dependence of the four fitted parameters in the "Gauss" method with respect to x for the case of π^+.

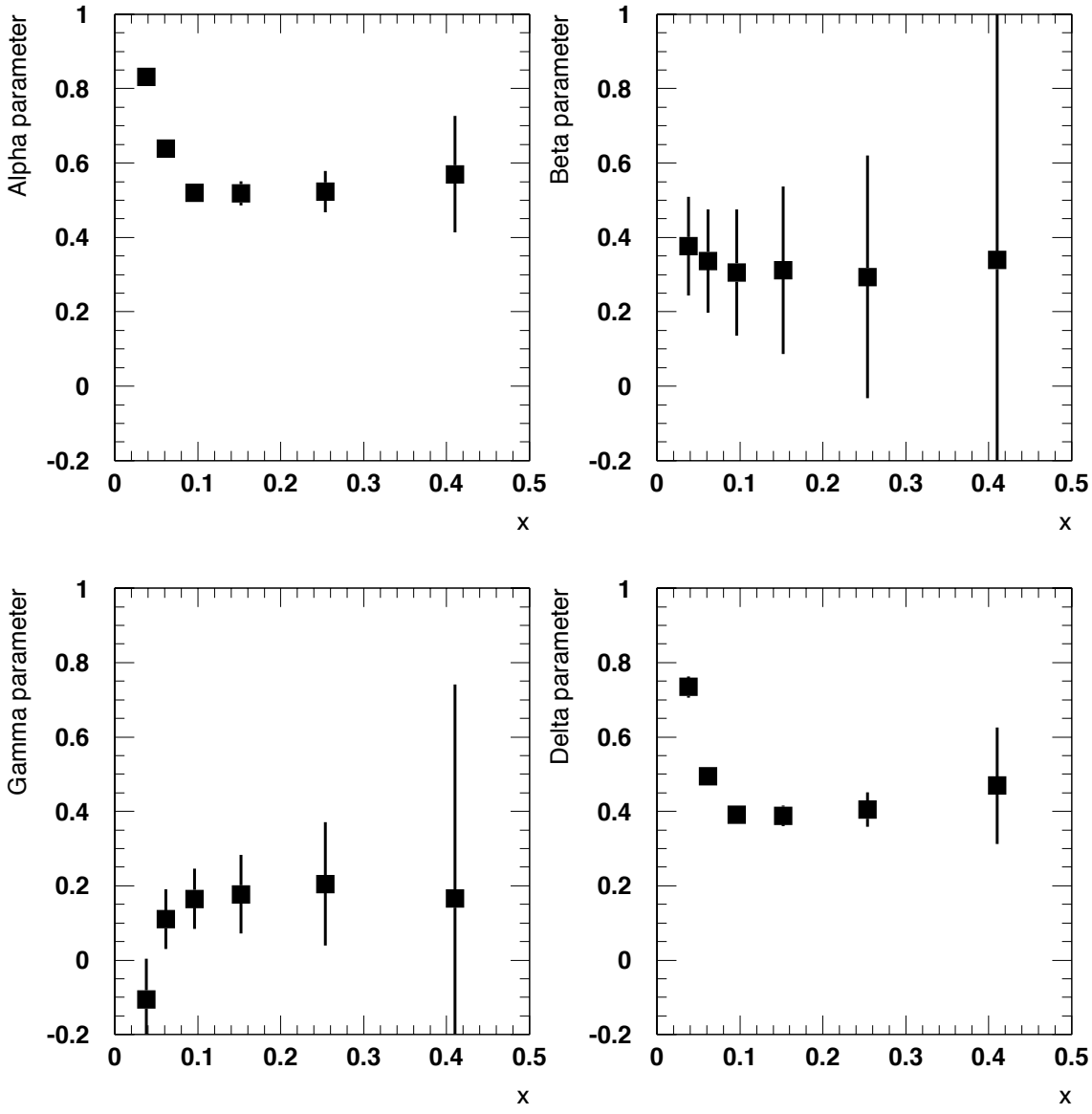

Figure 4.20: The dependence of the four fitted parameters in the "Gauss" method with respect to x for the case of π^-.

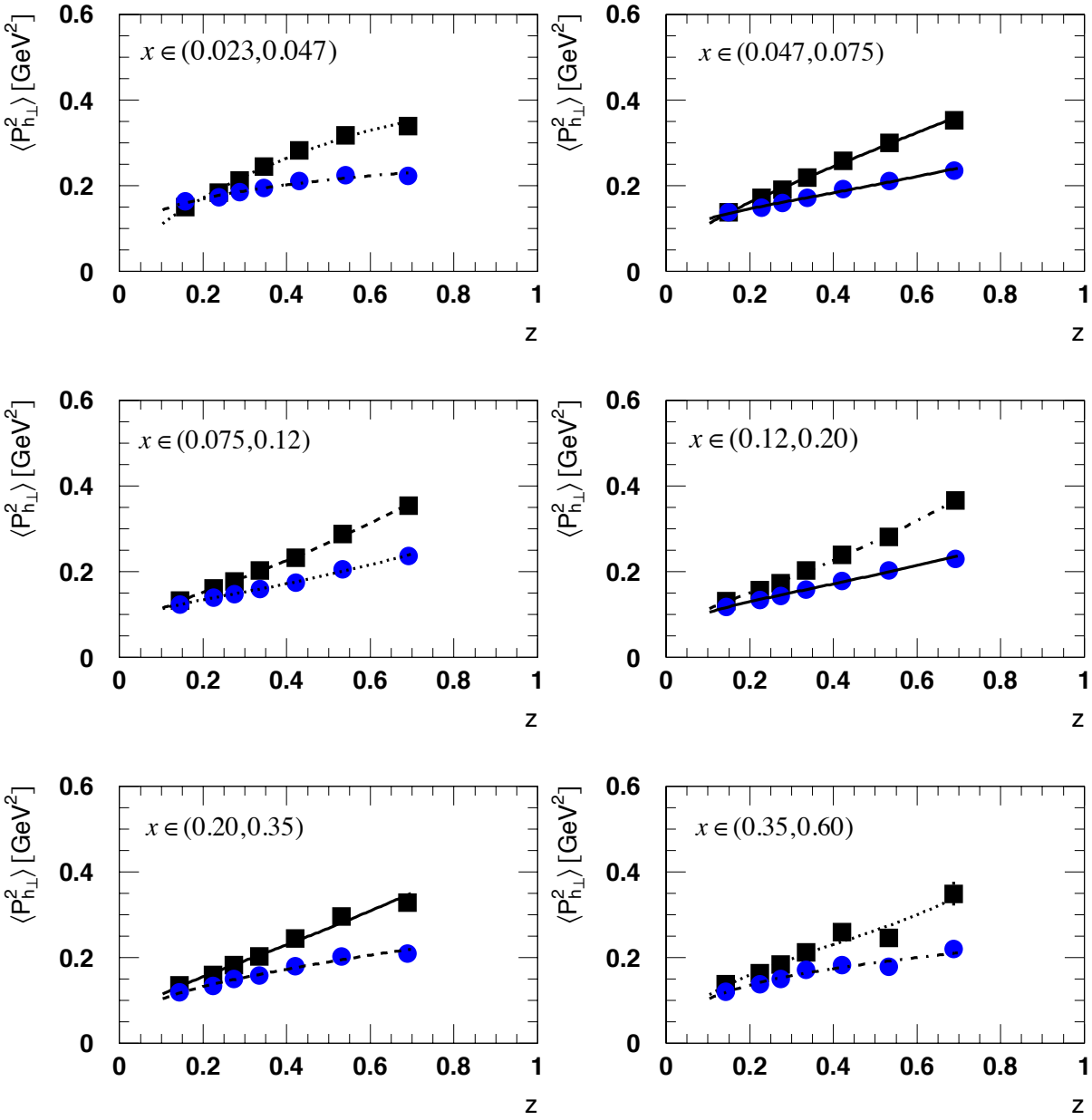

Figure 4.21: Fits of $\langle P_{h\perp}^2 \rangle$ for π^+ over unfolded real data using both methods.

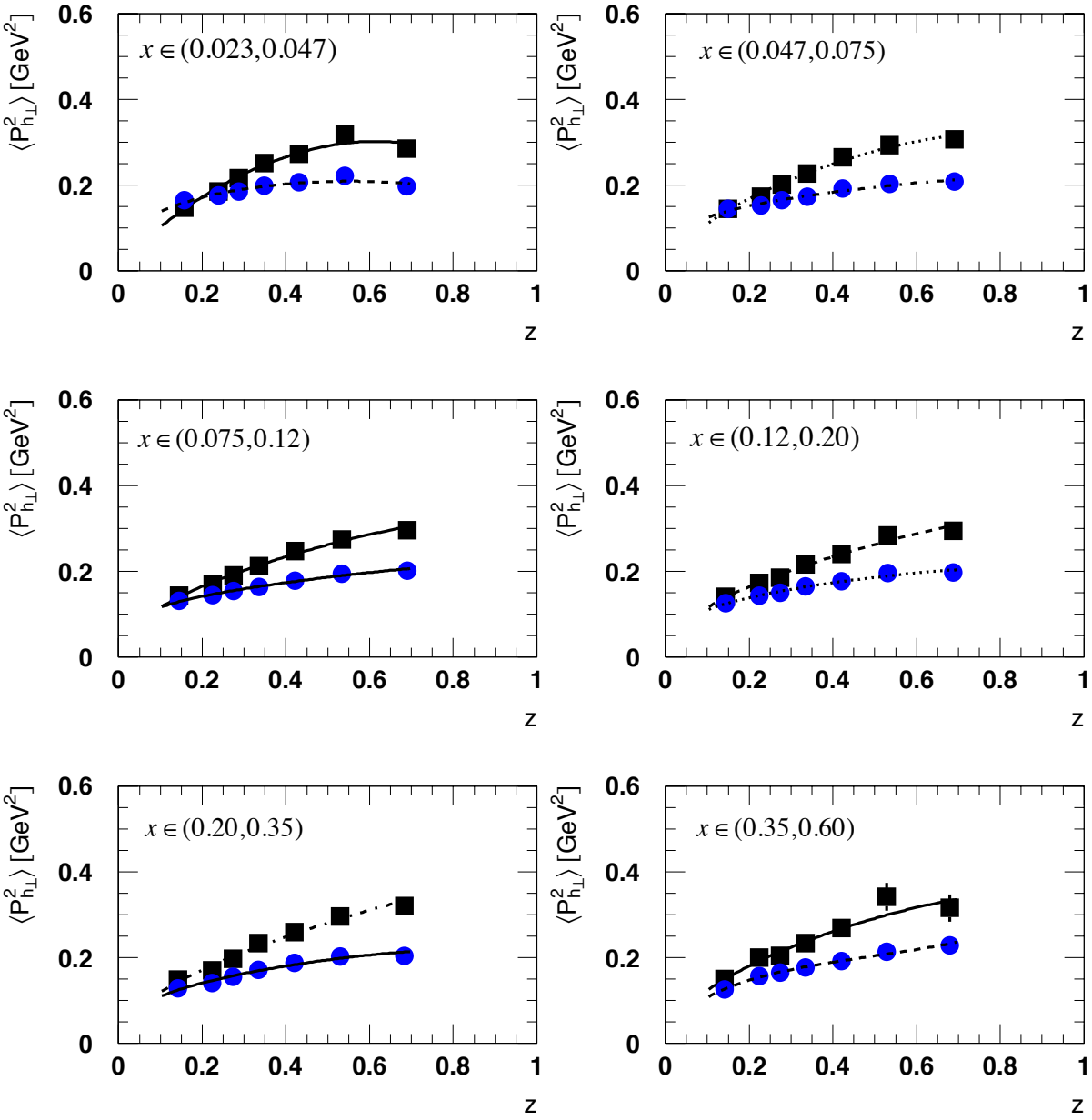

Figure 4.22: Fits of $\langle P_{h\perp}^2 \rangle$ for π^- over unfolded real data using both methods.

5
Conclusions

From the comparison between real data and results from PYTHIA Monte Carlo there are several things we have to take into account. First of all, as we stated on the results, the agreement is very low for all the particles analysed. Both the Monte Carlo data and the real data are under acceptance effects. PYTHIA simulates SIDIS events, so in principle its approach to real data should be quite precise. However, we can perceive a big disagreement that can be due to a bad simulation of the high HERMES acceptance effects. It could also be that PYTHIA is not so good at describing our concrete analysis, since the tuning parameters available for transverse momentum are not very flexible. It was not designed for the study of TMDs and therefore, it can be that it does not give a good description of our variables.

In order to make a more realistic study of reality, GMC_Trans Monte Carlo generator was used. In this case, a better agreement between Monte Carlo results and reality is expected, because GMC_Trans was specifically designed for the measurement of TMDs. Under the hypothesis done, the agreement between both "mean method" and "Gaussian method" would have to be high. The study of figure (4.10) shows a nice agreement for the Monte Carlo results when comparing the two methods. That is not the case of real data: plots 4.21 and 4.22 and the numerical results obtained for the fit (α, β, γ and δ) show a significant disagreement between both methods.

This disagreement could be understood in different ways, whose investigation would go beyond the scope of this thesis:

- real data is not Gaussian. The equivalence of the two methods is based on a Gaussian distribution. However, it has been pointed out that several contributions can lead to deviations from a Gaussian at large values of $P_{h\perp}$, such as NLO QCD effects (gluon radiation resulting in power-law corrections). These deviations can influence both methods differently. It would be interesting to implement such corrections in the Monte Carlo simulation and test its impact.

- real data p_T distribution is not only one Gaussian but the convolution of numerous Gaussians. In principle, each quark flavor could contribute to the transverse momentum distributions of the final hadron with a different Gaussian width. That again would deviate from the simple Gaussian Ansatz and might affect differently the two methods.

- insufficient number of data points. The methods were tested with a high-statistics Monte Carlo sample, finely binned in transverse momentum. For the real data only a very limited number of $P_{h\perp}$ bins were available, for which binning effects (integration over to large regions in $P_{h\perp}$ and in fact of other kinematic variables) may become important. In principle also this could be an extension of the current analysis by binning the available Monte Carlo data in a similar way as the real data.

In conclusion one can state that in principle the extraction of transverse-momentum parameters is possible for a Gaussian distribution, giving results for Monte Carlo data that are in excellent agreement with the input to the Monte Carlo simulation, real data appears to be much more difficult to fit. Two different methods gave non-consistent results. The reasons for the disagreement need to be explored before reliable parameterizations for the TMD PDFs and FFs can be extracted.

Acknowledgements

First of all I would like to thank my supervisors Gunar and Charlotte because of their never ending patience and the effort and time invested in me. They made it possible to perform this work in a very familiar and pleasant environment and gave me the chance to study in different research centres among lots of experts in the field. These facts substantially enriched myself, not only in the intellectual aspect, launching my future career, but also in the personal one, offering their advice when needed.

I would also like to thank my master friends because of their never-failing humour and support. Specially the ones with whom I shared the department seminar during the hardest thesis days. Having lunch with them was always a nice way of forgetting about work and resting the mind. As well, I do not want to forget about my closest friends, always there to cheer me up when the work was too much, and particularly the one that everyday had a word to make me smile.

I also do not want to forget some professors from the department, offering their help every time they could, either giving us a place to sit when our office was full or simply with a homely word every morning. As well, I would like to thank the people from the HERMES collaboration because of their help and contribution to the results of this thesis. Their always kind suggestions and comments were essential to give this analysis a much wider point of view. Finally, without the suggestions and the help of the tribunal members, many little details would have been missing in this master thesis.

Finalmente, pero no por eso menos relevante, me gustaría dedicar esta tesis de máster a mis Aitas, por haberme dado fuerzas desde que era una niña para ir siempre un paso más allá y por enseñarme a creer en mí y en mi capacidad de obtener siempre lo que se desea mediante el trabajo y la dedicación, sabiendo siempre lo

que uno quiere y no dejando que el miedo nos haga pensar que somos incapaces. Sin su incansable apoyo y energía creo que nunca hubiera tomado la decisión de ir por este camino y seguir adelante con mi carrera de investigadora. Por eso ahora, no puedo más que estarles por siempre agradecida y saber, que todo lo que consiga de ahora en adelante, así como todo el camino recorrido hasta aquí se lo debo plenamente a ellos, y siempre se lo agradeceré.

Bibliography

[1] J.L.Sanchez-Gomez Editors: F.Barreiro. *Lepton Nucleon Interactions at High Energies*. Wold Scientific.

[2] U. Elschenbroich. *Transverse Spin Structure of the Proton Studied in Semi-inclusive DIS*. PhD thesis, Universiteit Gent, 2006. DESY - THESIS - 06-059.

[3] `http://hermes.desy.de/notes/pub/trans-public-subject.html`.

[4] M. Anselmino, A. Efremov, and E. Leader. The Theory and phenomenology of polarized deep inelastic scattering. *Phys.Rept.*, 261:1–124, 1995.

[5] Alessandro Bacchetta, Markus Diehl, Klaus Goeke, Andreas Metz, Piet J. Mulders, et al. Semi-inclusive deep inelastic scattering at small transverse momentum. *JHEP*, 0702:093, 2007.

[6] Ringaile Placakyte. Parton Distribution Functions. 2011.

[7] Marco Stratmann. Lectures on qcd at work, May 2012. ECT* Doctoral Training Programme.

[8] D. J. Gross C. G. Callan Jr. *Phys. Rev. Lett.*, 22:156, 1969.

[9] `https://hermes-wiki.desy.de/Multiplicity_Report`.

[10] Alessandro Bacchetta. Transverse-momentum-dependent parton distributions (TMDs). *AIP Conf.Proc.*, 1374:29–34, 2011.

[11] A. Bacchetta U. Elschenbroich Y. Miyachi G. Schnell T.-A. Shibata Y. Takubo H. Tanaka R. Seidl. On the way to a decomposition of the flavor

dependent sivers functions: Transverse asymmetries, the decomposition and the transverse mc generator gmctrans. HERMES Internal Report 04-039.

[12] A.A. Sokolov and I.M. Ternov. On Polarization and spin effects in the theory of synchrotron radiation. *Sov.Phys.Dokl.*, 8:1203–1205, 1964.

[13] Jean Buon and Klaus Steffen. HERA VARIABLE ENERGY 'MINI' SPIN ROTATOR AND HEADON E P COLLISION SCHEME WITH CHOICE OF ELECTRON HELICITY. *Nucl.Instrum.Meth.*, A245:248, 1986.

[14] `https://hermes-wiki.desy.de/Triggers`.

[15] D. Allasia et al. Inelastic J / psi production in deep inelastic scattering from hydrogen and deuterium and the gluon distribution of free nucleons. *Phys.Lett.*, B258:493–498, 1991.

[16] J. Breitweg et al. D* production in deep inelastic scattering at HERA. *Phys.Lett.*, B407:402–418, 1997.

[17] K. Rith. Spin asymmetries in deep-inelastic electron-nucleon scattering - selected hermes results. *Prog. Part. Nucl. Physics*, 49:245–324, 2002.

[18] F. Meissner. *Measurement of the J/psi Cross Section and Double-Spin Asymmetries in Vector Meson Production in Polarised Lepton-Nucleon Scattering at HERMES*. PhD thesis, Humboldt-Universität Berlin, 2001. DESY - THESIS - 00-009.

[19] `https://hermes-wiki.desy.de/Productions`.

[20] "`http://www-hermes.desy.de/groups/daqlgrp/OFFLINE_DQ/uDST/index.html`".

[21] P. Jung. Pion - kaon separation with the hermes rich detector, 2000.

[22] C.Scholz F.Zetsche B.Povh, K. Rith. *Particles and Nuclei. An introduction to the Physical Concepts*. Springer-Verlag Berlin Heidelberg.

[23] M. Beckmann. *Extraction of polarised quark distributions of the nucleon from deep inelastic scattering at the HERMES experiment*. PhD thesis, Albert-Ludwigs-Universität Freiburg, 2000. DESY - THESIS - 00-020.