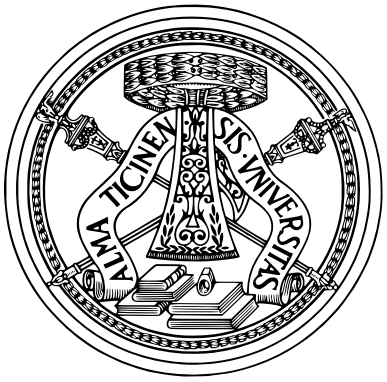

UNIVERSITÀ DEGLI STUDI DI PAVIA

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI

CORSO DI LAUREA MAGISTRALE IN SCIENZE FISICHE

Exploring the flavour dependence of unpolarized transverse-momentum-dependent distributions

Relazione per la laurea di

Andrea Signori

Supervisori

Alessandro Bacchetta
Università degli Studi di Pavia
Dipartimento di Fisica

Marco Radici
Università degli Studi di Pavia
INFN - sezione di Pavia

Anno Accademico 2011/2012

To my family.

Non quia difficilia sunt non audemus,
sed quia non audemus difficilia sunt.
L. A. Seneca

Nothing in life is to be feared, it is only to be understood.
Now is the time to understand more,
so that we may fear less.
Marie Curie

Contents

Chapter 1

Introduction

Protons and neutrons constitute the building blocks of matter, accounting for almost all the mass of our world. Even if we are still far from understanding their physical inner structure, many efforts have been made to deepen our knowledge about them.

Since a few years, thanks to a fruitful synergy of theoretical and experimental progress, we started opening the study of new multi-dimensional images of the distribution of quarks and gluons inside protons and neutrons. These pictures of the physics inside the nucleon challenge our comprehension of the underlying theory of quark and gluon dynamics (QCD - quantum chromodynamics) and at the same time put us face to face with fundamental questions [1], such as: "What is the shape of the nucleon?" "Where does the spin of the nucleon come from?"

When we look into the nucleons at extremely high resolution, we are in the regime of *perturbative* QCD: quarks and gluons appear almost free and we can explain this feature (asymptotic freedom) starting from the QCD Lagrangian. With the due caveats, we can compare the situation to observing water at extreme magnifications, and seeing quasi-free water molecules. As we reduce the magnification, we realize that the molecules clump together in heavier, composite droplets. Eventually, at low magnification they form a single object, like the nucleon. Pursuing the analogy, it is as if we were unable to describe water starting from the dynamics of molecules: currently we do not understand QCD in the *nonperturbative* regime. In fact *confinement*, the reason for quarks and gluons being inescapably bound into nucleons, is left without any mathematical explanation yet. Confinement is the most crucial characteristic of the theory and represents one of the hardest physics problems of today.

One of the best ways that we could follow to understand QCD and collect hints on confinement-induced physics is to study the inner structure of the nucleon through phenomenology, using experimental extractions of quantities that cannot be predicted with theoretical computation. Following this fruitful path, since a few years a systematic study of multi-dimensional pictures of the nucleon was started. The knowledge of the multi-dimensional structure allows the analysis of properties otherwise inaccessible: quark-gluon correlations, effects of final-state interactions, spin-orbit and spin-spin correlations, and much more. The situation may be compared to diagnostic studies: the electrocardiography, for example, gives us mono-dimensional information about the hearth activity. It is of fundamental importance, but it does not give detailed information about the multidimensional inner structure. Instead, more important for this purpose are multi-dimensional tomographies of heart activity (MRI, TAC and others). The enormous advantages of medical diagnostic imaging literally revolutionized medicine and surgery. In a similar way, the latest "multi-dimensional" pictures of the nucleon obtained with QCD phenomenology can improve the current status of hadronic physics and aim at better understanding particle physics in general.

Although one-dimensional (collinear) parton distribution functions (fig.1.1) are extremely useful for studying any process involving hadrons (including the proton-proton collisions taking place at the LHC), from the point of view of nucleon tomography they are rather limited, because they describe the distribution of partons in a single dimension. On the contrary, *Wigner distributions* represent the most complete information in the phase space about the parton. Collinear PDFs are projections of Wigner distributions integrated over all the physical variables apart from the light-cone momentum fraction x carried by the parton inside the hadron.

If we integrate Wigner distributions over all the coordinates, we obtain the so-called *transverse-momentum dependent distributions* (TMDs). They represent pictures of three-dimensional densities in momentum space. The distributions change depending on the energy scale at which they are probed (in a way that is calculable using *evolution equations* from perturbative QCD) and on the value of the longitudinal fractional momentum.

There are many nontrivial questions concerning TMDs that do not have an answer yet. First of all, these pictures are not spherically symmetric, because there is a clear distinction between the longitudinal dimension (the hard-probe direction) and the two transverse directions. In fact, the nucleon seen from the point of view of the probe does not look spherical at all, but rather like a flat dish, due to Lorentz contraction. We still do not have

sufficient information to discern if the parton density is higher in the center of the dish and decreasing monotonically as we move to the borders, or decreases in the center. At present, we know that experimental data (e.g. Drell-Yan) point towards *Gaussian* TMDs (in the case of unpolarized partons) with variance dependent on the collinear momentum fraction and the hard scale.

However many problems arise considering such a parametrization and corrections are needed. First of all, model calculations typically lead to non-Gaussian behaviour, also in unpolarized hadrons [2]. Significant deviations at higher values of the transverse momentum directly arise from perturbative QCD corrections (p_T-resummation effects); moreover, due to their non vanishing value in the limit of zero transverse momentum, Gaussian TMDs are compatible only with $s - wave$ functions. As a consequence, the *Gaussian ansatz* is a good mathematical tool to perform preliminary studies of the multi-dimensional structure of nucleons, but a more accurate functional form must be definitely found.

Another important question concerns the possible flavour dependence of transverse momentum dependent distributions. Is the up quark moving in the nucleon with greater mean square momentum than the down one (or viceversa)? What about sea quarks? The aim of this thesis is to perform an investigative study of the flavour dependence of unpolarized TMDs from Semi-Inclusive Deep Inelastic Scattering (SIDIS) data collected at the COMPASS experiment at CERN, *under the assumption of Gaussian parametrization*. We will use a statistical method that allows to extract *distributions* of the mean square transverse momentum for up, down and sea quarks and for hadrons fragmenting in *favoured* and *unfavoured* channels. COMPASS data are still preliminary: for this reason the results from the present analysis will be revised once official data will be available. Also, a comparison with results from similar investigations conducted at Jefferson Lab's facilities is required: these ones gave first feeble indications that the down-quark distribution could be larger than the up-quark one [3]. This behaviour is also supported by lattice QCD computations [4].

Results obtained through the present analysis, instead, indicate an up quark "wider" than the down one, contrary to the other explorations. This could be due to the limits of the present analysis or to the differences among existing studies.

The subject of parton intrinsic transverse momentum and its flavour dependence is of growing interest also to the spin structure community: the picture of partons moving collinear with the proton momentum is not sufficient to explain many spin features of the nucleons. The updated results

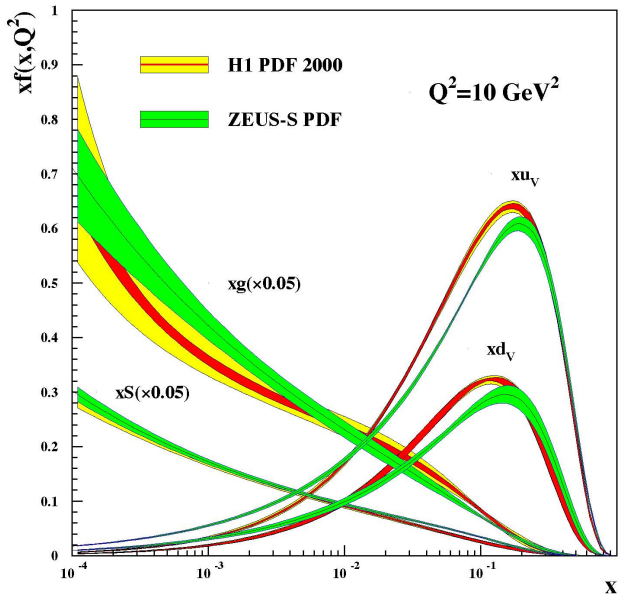

Figure 1.1: Collinear parton distribution functions for gluons, up, down and sea quarks measured by ZEUS and H1 collaboration at the HERA experiment (DESY).

of the study would be of great relevance not only for hadronic physics: they may have a significant impact on High Energy Physics in general. The statistical analysis of the huge amount of data collected at hadron colliders strongly relies on the detailed knowledge of parton distribution functions and fragmentation functions. For the moment being, the transverse momentum dependent part of these distributions is assumed flavour *independent*, but this assumption has never been properly tested and could be important to describe precisely the transverse momentum spectra of the the detected particles. TMDs play a crucial role also in the search for new physics Beyond the Standard Model: for example, an interesting method [5] has been proposed to determine the CP parity of the Higgs boson using transverse momentum dependent distributions of *polarized gluons*.

Chapter 2

Scattering of leptons off nucleons

In the early 1910s scientists probed the inner structure of the atom - which was at that time believed to be really *a-tomos*, i.e. without parts - with Rutherford's experiments. The scattering of an alpha-particle on thin gold foils revealed a hard core of matter (protons and neutrons) building the nucleus, empty space and electrons spread all around.

Within nearly fifty years this picture dramatically improved: in the 1960s scientists at SLAC and MIT collected data on deep inelastic scattering of leptons off atomic nucleons, accessing the inner structure of nucleons.

Deep inelastic scattering experiments are younger versions of Rutherford ones, with the same philosophy, but investigating Nature at the fundamental length-energy scale of quarks and gluons. The theory underlying our analysis is a non-abelian Quantum Field Theory with the SU(3) gauge group. It is known as Quantum Chromodynamics, QCD. Detailed and self−contained treatments can be found in [6, 7, 8, 9]. The accessed physical structure evolves with energy, as the theory underlying this framework and all the structure functions involved in the scattering run with the hard scale of the process.

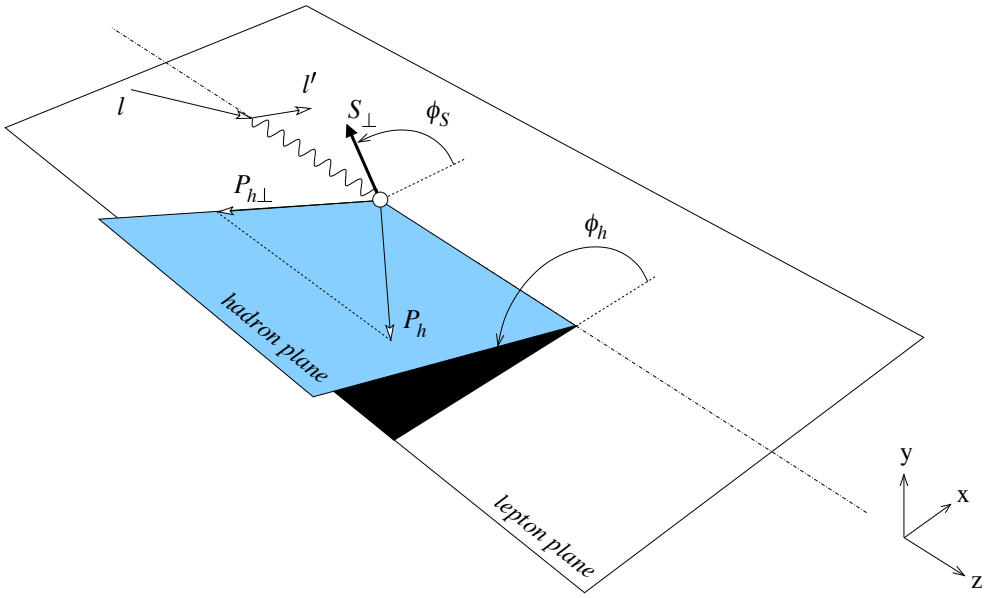

Figure 2.1: Kinematics of deep-inelastic scattering (DIS) of a lepton with 4-momentum l off a hadronic target at rest. ϕ_S is the azimuthal angle between the direction of the transverse target polarization $S_\perp$ and the lepton plane, formed by l and 4-momentum of the final lepton, l'. In inclusive DIS, no other particles are observed in the final state. For semi-inclusive DIS (SIDIS) a hadron with 4-momentum $\boldsymbol{P}_h$ is detected. ϕ_h is the azimuthal angle between the lepton and the hadron plane, formed by $\boldsymbol{P}_h$ and $\hat{z}$ axis [11].

2.1 Inclusive DIS

Let's consider the scattering of a lepton off a nucleon [10]:

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

where ℓ denotes the lepton beam, N the nucleon target, and where four-momenta are given in parentheses. The lepton mass can be neglected, M is the mass of the nucleon. We define the space-like four-momentum of the photon as $q = l - l'$ ($Q^2 = -q^2$ being the hard scale) and we introduce the variables

$$x_B = \frac{Q^2}{2\,P\cdot q}, \qquad y = \frac{P\cdot q}{P\cdot l} = \frac{Q^2}{x_B s} \qquad \gamma = \frac{2Mx}{Q} \tag{2.2}$$

x_B, the Bjorken variable, is the degree of inelasticity of the reaction, y is the fraction of energy transferred from the lepton to the proton and γ represents

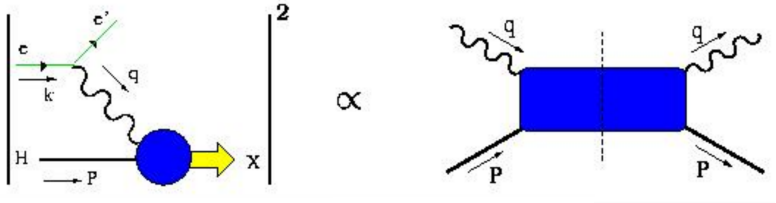

Figure 2.2: The cross section of DIS, built from the squared Feynman amplitude $|M|^2$; the r.h.s. diagram is a graphical representation of the hadronic part of $|M|^2$, i.e. the *hadronic tensor* [9].

mass corrections term.

S is the spin of the target nucleon and $S_\perp$ is its transverse part w.r.t. the virtual photon direction. ϕ_S is the azimuthal angle between the lepton plane and the direction of S (see fig. 2.1).

For later convenience we introduce normalized vectors:

$$\hat{q}^\mu = \frac{q^\mu}{Q}, \tag{2.3}$$

$$\hat{t}^\mu = \frac{2x_B}{Q\sqrt{1+\gamma^2}}\left(P^\mu - \frac{P\cdot q}{q^2}q^\mu\right), \tag{2.4}$$

$$\hat{l}^\mu = -\frac{g_\perp^{\mu\nu} l_\nu}{|g_\perp^{\mu\nu} l_\nu|} \tag{2.5}$$

with the projectors on the transverse space defined as

$$g_\perp^{\mu\nu} = g^{\mu\nu} + \hat{q}^\mu \hat{q}^\nu - \hat{t}^\mu \hat{t}^\nu, \tag{2.6}$$

$$\epsilon_\perp^{\mu\nu} = \epsilon^{\mu\nu\rho\sigma}\hat{t}_\rho \hat{q}_\sigma. \tag{2.7}$$

In this chapter we will systematically neglect all correction of order M/Q ("leading twist" (LT) approximation) and deal with totally inclusive processes where the only observed final state is the outgoing lepton. We will define "leading order" (LO approximation), instead, quantities of order $\alpha_S(Q^2)$, where α_s is the QCD coupling constant.

2.1.1 Cross section and tensors

The differential cross section for *polarized* lepton-nucleon scattering can be written in a general way as the contraction between a leptonic and a hadronic

tensor:

$$\frac{\mathrm{d}^3\sigma}{\mathrm{d}x_B\,\mathrm{d}y\,\mathrm{d}\phi_S} = \frac{\alpha^2}{2\,s\,x_B\,Q^2}\,L_{\mu\nu}(l,l',\lambda_e)\;2MW^{\mu\nu}(q,P,S), \tag{2.8}$$

where λ_e is the helicity of the lepton and s is the square of the energy in the CM frame. A pictorial representation of the hadronic tensor is given in fig. 2.2. The cross section is proportional to the squared modulus of the Feynman amplitude M and the hadronic tensor represents the hadronic part of the probability density. The given formula is valid in the approximation of single-photon exchange, i.e. in the so-called Born approximation.

Considering the lepton to be longitudinally polarized, in the massless limit the leptonic tensor is given by

$$\begin{aligned} L_{\mu\nu} &= \sum_{\lambda_e'} \Big(\bar{u}(l',\lambda_e')\,\gamma_\mu\,u(l,\lambda_e)\Big)^* \Big(\bar{u}(l',\lambda_e')\,\gamma_\nu\,u(l,\lambda_e)\Big) \\ &= -Q^2 g_{\mu\nu} + 2\left(l_\mu l'_\nu + l'_\mu l_\nu\right) + 2\mathrm{i}\,\lambda_e\,\epsilon_{\mu\nu\rho\sigma}\,l^\rho l'^\sigma. \end{aligned} \tag{2.9}$$

Some QED radiative corrections can be included without modifying this formula [12], but effects beyond Born approximation are left out [13]. The leptonic tensor contains all the information on the leptonic probe which can be described by means of Quantum Electrodynamics (QED), while the information on the hadronic target is contained in the hadronic tensor:

$$2MW^{\mu\nu}(q,P,S) = \frac{1}{2\pi}\sum_{\mathcal{X}}\int \frac{\mathrm{d}^3P_{\mathcal{X}}}{(2\pi)^3\,2P^0_{\mathcal{X}}}\,(2\pi)^4\,\delta^{(4)}\Big(q+P-P_{\mathcal{X}}\Big)\,H^{\mu\nu}(P,S,P_{\mathcal{X}}), \tag{2.10}$$

$$H^{\mu\nu}(P,S,P_{\mathcal{X}}) = \big\langle P,S\big|\,J^\mu(0)\,\big|\mathcal{X}\big\rangle\big\langle\mathcal{X}\big|\,J^\nu(0)\,\big|P,S\big\rangle. \tag{2.11}$$

The state $\mathcal{X}$ symbolizes any final state with total momentum $P_{\mathcal{X}}$ which can be reached by acting with the electromagnetic J^μ current operator on the target. We sum over all of them since in inclusive processes the final state goes undetected.

In general the expression of the hadronic tensor can be parametrized in terms of *structure functions* by requiring specific symmetry properties:

$$\text{hermiticity:} \qquad W^*_{\mu\nu}(q,P,S) = W_{\nu\mu}(q,P,S), \tag{2.12a}$$

$$\text{parity:} \qquad \Lambda^\rho_\mu\Lambda^\sigma_\nu W_{\rho\sigma}(q,P,S) = W_{\mu\nu}(\tilde{q},\tilde{P},-\tilde{S}), \tag{2.12b}$$

$$\text{time-reversal:} \qquad \Lambda^\rho_\mu\Lambda^\sigma_\nu W^*_{\rho\sigma}(q,P,S) = W_{\mu\nu}(\tilde{q},\tilde{P},\tilde{S}) \tag{2.12c}$$

where $\tilde{q}^\nu = \Lambda^\nu_\rho q^\rho$ and similarly for the other vectors and where Λ is the matrix representation of the parity-reversal transformation:

$$\Lambda^\rho_\sigma = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{2.13}$$

For the simplest case of unpolarized scattering the most general representation of the hadronic tensor consistent with previous constraints can be built using the metric $g^{\mu\nu}$ and symmetric combinations of P^μ and q^ν

$$2MW^{\mu\nu} = 2M\left[A\, g^{\mu\nu} + B\, q^\mu q^\nu + C\, \frac{P^\mu P^\nu}{M^2} + D\, \frac{P^\mu q^\nu + q^\mu P^\nu}{M^2}\right], \tag{2.14}$$

where each coefficient is a real function of Q^2 and $P \cdot q$.

Antisymmetric terms such as $\mathrm{i}\epsilon^{\mu\nu\rho\sigma} P_\rho q_\sigma$ are excluded by parity invariance and should be taken into account when considering weak interactions, e.g. neutrino scattering.

Finally, electromagnetic gauge invariance requires that

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

From this condition, it follows that

$$D = -\frac{P \cdot q}{q^2} B, \qquad\qquad C = \left(\frac{P \cdot q}{q^2}\right) B + \frac{M^2}{q^2} A \tag{2.16}$$

so that there are only two independent coefficients.

Therefore, for the *unpolarized* scattering we only have two structure functions, one for each polarization state of the virtual photon probing the nucleon:

$$2MW^{\mu\nu}(q, P, S) = \frac{1}{x_B}\left[-g^{\mu\nu}_\perp F_T(x_B, Q^2) + \hat{t}^\mu \hat{t}^\nu F_L(x_B, Q^2)\right]. \tag{2.17}$$

The connection with other frequently used unpolarized structure functions [6] is:

$$F_T(x_B, Q^2) = 2x_B F_1(x_B, Q^2), \tag{2.18}$$
$$F_L(x_B, Q^2) = (1+\gamma^2)F_2(x_B, Q^2) - 2x_B F_1(x_B, Q^2). \tag{2.19}$$

In the general case of inclusive DIS with beam and target *polarization*, we have the additional pseudo-vector S, which generates two more independent structures, anti-symmetric in Lorentz indices. Therefore, two more structure functions appear, the polarized g_1 and g_2 [9]

$$\begin{aligned} 2MW^{\mu\nu}(q,P,S) = \frac{1}{x_B}\Big[& - g_\perp^{\mu\nu} F_T(x_B,Q^2) + \hat{t}^\mu \hat{t}^\nu F_L(x_B,Q^2) \\ & + \mathrm{i} S_L \epsilon_\perp^{\mu\nu} 2x_B \left(g_1(x_B,Q^2) - \gamma^2 g_2(x_B,Q^2)\right) \\ & + \mathrm{i}\hat{t}^{[\mu}\epsilon_\perp^{\nu]\rho} S_\rho 2x_B\gamma \left(g_1(x_B,Q^2) + g_2(x_B,Q^2)\right)\Big] \end{aligned} \tag{2.20}$$

where S_L is the longitudinal component of the target spin vector. Finally, contracting leptonic and hadronic tensors we obtain the following expression for the inclusive DIS cross-section:

$$\begin{aligned} \frac{d\sigma}{dx_B\, dy\, d\phi_S} = \frac{2\alpha^2}{x_B y Q^2}\,\frac{y^2}{2\,(1-\varepsilon)}\Big\{ & F_T + \varepsilon F_L + S_L \lambda_e \sqrt{1-\varepsilon^2}\; 2x_B\,(g_1 - \gamma^2 g_2) \\ & - |\boldsymbol{S}_\perp| \lambda_e \sqrt{2\,\varepsilon(1-\varepsilon)}\,\cos\phi_S\, 2x_B\gamma\,(g_1+g_2)\Big\}, \end{aligned} \tag{2.21}$$

where the structure functions on the r.h.s. depend on x_B and Q^2. Moreover,

$$\varepsilon = \frac{1-y}{1-y+\frac{1}{2}y^2} \tag{2.22}$$

is the ratio ε of longitudinal and transverse photon flux.

2.1.2 Parton model picture of DIS

The phenomenology of DIS teaches us that at sufficiently high Q^2 we can assume that the scattering of the lepton takes place off an elementary constituent with mass m, the parton inside the nucleon (fig. 2.3).

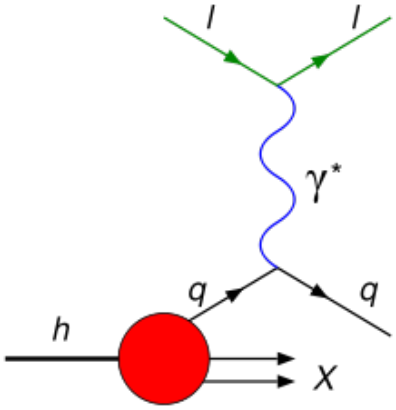

Figure 2.3: In the parton model picture the lepton scatters off a parton, an elementary constituent of the hadron

The *parton model* of hadron structure reproduces exactly this picture: hadrons are ensembles of asymptotically-free Dirac particles and the scattering can be factorized in a hard and a non-perturbative part. The first one deals with the electromagnetic interaction between the lepton and the struck parton and the latter refers to the probability of finding such a parton inside the nucleon.

Figure 2.4 represents the parton model interpretation of the hadronic tensor, the so-called handbag diagram. The hard electromagnetic interaction of the virtual photon with the struck parton (identified with a quark) is factorized from a soft function, the non-perturbative quark-quark correlator. A sum over all possible final states is understood in the cut of both the on-shell quark propagator and the soft correlator Φ.

In this model the final state $\mathcal{X}$ of the scattering can be split into a quark with momentum k and a residual state X with momentum P_X. Considering the electron-quark interaction at tree level only, the expression of the hadronic tensor is:

$$\begin{aligned} 2MW^{\mu\nu}(q,P,S) = \frac{1}{2\pi}\sum_q e_q^2 \sum_X \int \frac{\mathrm{d}^3 P_X}{(2\pi)^3\, 2P_X^0} \int \frac{\mathrm{d}^3 k}{(2\pi)^3\, 2k^0}\, (2\pi)^4 \\ \times \delta^{(4)}\Big(P+q-k-P_X\Big) \big\langle P,S\big|\, \bar{\psi}_i(0)\,\big|X\big\rangle\big\langle X\big|\,\psi_j(0)\,\big|P,S\big\rangle \\ \times \gamma^\mu_{ik}\,(\not{k}+m)_{kl}\,\gamma^\nu_{lj}, \end{aligned} \tag{2.23}$$

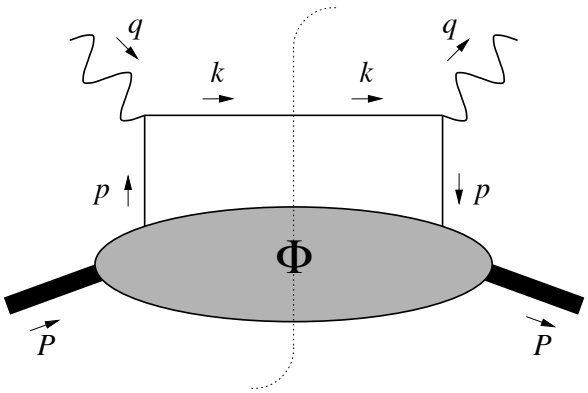

Figure 2.4: The handbag diagram describing the parton model picture of the hadronic tensor.

where the index q denotes the quark flavor and e_q is the fractional charge of the quark. The roman indices indicate the Dirac components of the various quantities, and a summation is understood on repeated indices.

Introducing the *quark-quark correlation function* Φ, the hadronic tensor can be cast in a more compact form:

$$\begin{aligned} 2MW^{\mu\nu}(q,P,S) = \sum_q e_q^2 \int \mathrm{d}^4p\, \delta\left(\left(p+q\right)^2 - m^2\right) \theta\left(p^0+q^0-m\right) \\ \times \mathrm{Tr}\left[\Phi(p,P,S)\gamma^\mu \left(\not{p}+\not{q}+m\right)\gamma^\nu\right] \end{aligned} \tag{2.24}$$

where Φ_{ji} is a bilinear matrix element of Dirac field operators [1]:

$$\begin{aligned} \Phi_{ji}(p,P,S) &= \frac{1}{(2\pi)^4}\int \mathrm{d}^4\xi\, \mathrm{e}^{-\mathrm{i}p\cdot\xi}\big\langle P,S\big|\,\bar\psi_i(\xi)\,\psi_j(0)\,\big|P,S\big\rangle \\ &= \sum_X \int \frac{\mathrm{d}^3P_X}{(2\pi)^3\,2P_X^0}\,\big\langle P,S\big|\,\bar\psi_i(0)\,\big|X\big\rangle\big\langle X\big|\,\psi_j(0)\,\big|P,S\big\rangle\,\delta^{(4)}\Big(P-p-P_X\Big). \end{aligned} \tag{2.25}$$

Using the Sudakov decomposition in terms of light-cone components we can parametrize the initial and final quark momenta p and k, as well as the

[1] the definition of this matrix elements must be corrected using *gauge links* in order to ensure color gauge invariance [9]

momentum transfer q, with

$$\begin{aligned} p^\mu &= \left[\tfrac{p^2+|p_T|^2}{2xP^+},\ xP^+,\ p_T\right], \\ k^\mu &= \left[\tfrac{P_h^-}{z},\ \tfrac{z(k^2+|k_T|)}{2P_h^-},\ k_T\right], \\ q^\mu &= \left[\tfrac{P_h^-}{z_h},\ -x_B P^+,\ q_T\right], \end{aligned} \tag{2.26}$$

where the transverse direction refers to the virtual photon momentum direction. In our approach we assume that neither the virtuality of the quark, p^2, nor its squared transverse momentum, $|p_T|^2$, can be large in comparison with the hard scale Q^2. Therefore, the dominant light-cone components for p and k are $p^+ = xP^+$ and k^-, respectively. In eq. (2.24), neglecting terms which are $1/Q$ suppressed, we can use an approximated expression for the delta function

$$\delta\left(\left(p+q\right)^2 - m^2\right) \approx \delta(p^+ + q^+) \approx P^+\,\delta(x - x_B) \tag{2.27}$$

and replace

$$\mathrm{d}^4p = \mathrm{d}^2p_T\ \mathrm{d}p^-\ P^+\ \mathrm{d}x \tag{2.28}$$

obtaining

$$\begin{aligned} 2MW^{\mu\nu}(q,P,S) &\approx \sum_q e_q^2 \int \mathrm{d}^2p_T\,\mathrm{d}p^-\,\mathrm{d}x\ \frac{P^+}{2P\cdot q}\,\delta\left(x - x_B\right) \\ &\quad \times \mathrm{Tr}\left[\Phi^q(p,P,S)\,\gamma^\mu\left(\not{p} + \not{q} + m\right)\gamma^\nu\right] \\ &= \sum_q e_q^2\,\frac{1}{2}\mathrm{Tr}\left[\Phi^q(x_B,S)\,\gamma^\mu\frac{P^+}{P\cdot q}\left(\not{p} + \not{q} + m\right)\gamma^\nu\right] \end{aligned} \tag{2.29}$$

where we introduced the *fully integrated* correlation function

$$\begin{aligned} \Phi^q_{ji}(x,S) &= \int \mathrm{d}^2p_T\,\mathrm{d}p^-\ \Phi^q_{ji}(p,P,S)\Big|_{p^+=xP^+} \\ &= \int \frac{\mathrm{d}\xi^-}{2\pi}\ \mathrm{e}^{-\mathrm{i}p\cdot\xi}\left\langle P,S\right|\bar\psi^q_i(\xi)\ \psi^q_j(0)\left|P,S\right\rangle\Big|_{\xi^+=\boldsymbol{\xi}_T=0}. \end{aligned} \tag{2.30}$$

Using the expressions for p, k, q we arrive at the final expression for the hadronic tensor :

$$2MW^{\mu\nu}(q,P,S) \approx \sum_q e_q^2\,\frac{1}{2}\mathrm{Tr}\left[\Phi^q(x_B,S)\,\gamma^\mu\gamma^+\gamma^\nu\right]. \tag{2.31}$$

2.1.3 Correlators and probability

At this point we need to analyze the structure of the correlation function Φ^q. We recall that Φ^q is a matrix of Dirac field operators. Therefore, we can decompose it in a general way on a basis of Dirac structures:

$$1,\ \gamma_5,\ \gamma^\mu,\ \gamma^\mu\gamma_5,\ \mathrm{i}\sigma^{\mu\nu}\gamma_5,$$

Each coefficient of the decomposition is the most general (pseudo)scalar that can be formed by combining the Lorentz vectors p, P and the pseudovector S (in case of spin-half hadrons).

Moreover, each term of the full expression has to satisfy the conditions of hermiticity and parity invariance:

$$\text{hermiticity:} \qquad \Phi(p,P,S) = \gamma^0\,\Phi^\dagger(p,P,S)\,\gamma^0, \tag{2.32a}$$

$$\text{parity:} \qquad \Phi(p,P,S) = \gamma^0\,\Phi(\tilde{p},\tilde{P},-\tilde{S})\,\gamma^0. \tag{2.32b}$$

In the fully *polarized* case, the most general decomposition of the correlation function Φ without taking into account also contributions from *gauge links* is [14, 15]

$$\begin{aligned}\Phi(p,P,S) = {} & M\,A_1\,\mathbf{1} + A_2\,\not{P} + A_3\,\not{p} + \frac{A_4}{M}\,\sigma_{\mu\nu}P^\mu p^\nu + \mathrm{i}A_5\;p\cdot S\,\gamma_5 \\ & + M\,A_6\,\not{S}\,\gamma_5 + A_7\,\frac{p\cdot S}{M}\,\not{P}\,\gamma_5 + A_8\,\frac{p\cdot S}{M}\,\not{p}\,\gamma_5 + \mathrm{i}A_9\,\sigma_{\mu\nu}\gamma_5 S^\mu P^\nu \\ & + \mathrm{i}A_{10}\,\sigma_{\mu\nu}\gamma_5 S^\mu p^\nu + \mathrm{i}A_{11}\,\frac{p\cdot S}{M^2}\,\sigma_{\mu\nu}\gamma_5 P^\mu p^\nu + A_{12}\,\frac{\epsilon_{\mu\nu\rho\sigma}\gamma^\mu P^\nu p^\rho S^\sigma}{M},\end{aligned} \tag{2.33}$$

where the amplitudes A_i are real scalar functions $A_i = A_i(p\cdot P, p^2)$ with dimension $1/[m]^4$. The complete expression of the correlation function including also gauge links contributions is described in [16].

In this general case the expression of the fully integrated correlation function, using the definition (2.30), becomes

$$\Phi(x,S) = \frac{1}{2}\left\{ f_1\not{n}_+ + S_L\,g_{1L}\gamma_5\not{n}_+ + h_1\,\frac{\left[\not{S}_T,\not{n}_+\right]\gamma_5}{2}\right\}, \tag{2.34}$$

where $n_+ = (0, 1, 0, 0)$ and where we introduced the *integrated parton distribution functions* (PDFs)

$$f_1(x) = \int \mathrm{d}^2 p_T \, \mathrm{d}p^2 \, \mathrm{d}(2p \cdot P) \, \delta\left(p_T^2 + x^2 M^2 + p^2 - 2xp \cdot P\right) \left[A_2 + xA_3\right], \tag{2.35a}$$

$$g_{1L}(x) = \int \mathrm{d}^2 p_T \, \mathrm{d}p^2 \, \mathrm{d}(2p \cdot P) \, \delta\left(p_T^2 + x^2 M^2 + p^2 - 2xp \cdot P\right) \times \left[-A_6 - \left(\frac{p \cdot P}{M^2} - x\right)(A_7 + xA_8)\right], \tag{2.35b}$$

$$h_1(x) = \int \mathrm{d}^2 p_T \, \mathrm{d}p^2 \, \mathrm{d}(2p \cdot P) \, \delta\left(p_T^2 + x^2 M^2 + p^2 - 2xp \cdot P\right) \times \left[-A_9 - xA_{10} + \frac{p_T^2}{2M^2} A_{11}\right]. \tag{2.35c}$$

The function g_{1L}^q is the *helicity distribution* of parton q and it can be denoted also by Δq. The function h_1 is known as the *transversity distribution* (in the literature it is sometimes denoted by δq, $\Delta_T q$).

As already stated, the correlation function is a 4×4 Dirac matrix. The probabilistic interpretation of these three PDFs depends on the basis on which we decompose the parton and the target spin: for example, on the helicity basis f_1 and g_{1L} are probability densities (namely, diagonal elements in the helicity space) whereas h_1 is an interference between right- and left-handed partons in transversely polarized nucleons.
The transversity distribution recovers a probability interpretation if we choose the *transversity basis* instead of the helicity basis for both quark and hadron [17, 18].

2.1.4 DIS Structure Functions in the QPM

Inserting the general integrated correlation function (2.34) in the *QPM dependent* expression of the hadronic tensor we can obtain the form of the four structure functions in the framework of quark parton model.

In the simplest unpolarized case the correlator reduces to

$$\Phi(x, S) = \frac{1}{2} \; f_1 \not{n}_+. \tag{2.36}$$

Inserting it in (2.31) and computing the trace of Dirac's gamma matrices we obtain an expression for the hadronic tensor compatible with (2.20); the unpolarized DIS structure functions become:

$$F_T = x_B \sum_q e_q^2 \, f_1^q(x_B) \ , \tag{2.37}$$

$$F_L = 0 \ . \tag{2.38}$$

In the polarized case we must substitute the full eq. (2.34) in the hadronic tensor. The resulting expression for the structure functions is:

$$F_1 = \frac{1}{2} \sum_q e_q^2 \, f_1^q(x_B), \tag{2.39}$$

$$F_L = 0, \tag{2.40}$$

$$g_1 = \frac{1}{2} \sum_q e_q^2 \, g_{1L}^q(x_B), \tag{2.41}$$

$$g_1 + g_2 = 0. \tag{2.42}$$

We can see that the transversity distribution h_1 does not occur in any of the above structure functions. This is due to the specific Dirac structure of the transversity term in the correlation function which causes that term to disappear. Moreover, the sum of polarized structure-functions $g_1 + g_2$ also vanishes. Finally, the longitudinal structure function F_L vanishes, and it remains zero also at the next-to-leading-twist (NLT) level. To obtain a nonzero result we have to consider QCD corrections of order α_S to the parton model (*improved* QPM) or go to the NNLT level.

2.2 Semi-inclusive DIS

In one-particle-inclusive deep inelastic scattering the kinematics is similar to the inclusive DIS case except that one hadron, produced after the collision, is detected. The cross section depends in particular on the azimuthal angle of the final state hadron about the virtual photon axis (see fig. 2.1).

In the kinematic region where the transverse momentum of the outgoing hadron is low compared to Q ($P_{h\perp}^2 \ll Q^2$), the process can be *factorized* in non-perturbative and hard parts and the cross section can be cast in terms of transverse-momentum-dependent parton distribution functions (TMD PDFs) and fragmentation functions (TMD FFs).

These partonic functions are generalizations of the distribution and fragmentation functions appearing in the standard collinear factorization framework. They are often referred to as *unintegrated* functions, as they are not integrated over the transverse momentum.

2.2.1 Cross section and tensors

Let's consider the process

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

where ℓ denotes the lepton beam, N the nucleon target, and h the detected hadron (four-momenta are given in parentheses). Let's denote by M and M_h the masses of the nucleon and of the hadron h. For sufficiently high energies we can neglect the lepton mass. We introduce the variable

$$z_h = \frac{P \cdot P_h}{P \cdot q} \ , \tag{2.44}$$

which is the fraction of energy transferred from the virtual photon to the detected hadron. $\boldsymbol{P}_{h\perp}$ is the transverse component of $\boldsymbol{P}_h$ with respect to the photon momentum. Let's consider the case where the detected hadron h has spin zero or where its polarization is not measured. In this case the cross section for one-particle inclusive electron-nucleon scattering can be written as

$$\frac{2E_h \, \mathrm{d}^6\sigma}{\mathrm{d}^3\boldsymbol{P}_h \, \mathrm{d}x_B \, \mathrm{d}y \, \mathrm{d}\phi_S} = \frac{\alpha^2}{2sx_BQ^2} \, L_{\mu\nu}(l, l', \lambda_e) \, 2MW^{\mu\nu}(q, P, S, P_h), \tag{2.45}$$

or equivalently (using $\mathrm{d}^3P_h/2E_h \approx \mathrm{d}z_h\,\mathrm{d}^2P_{h\perp}/2z_h$) as

$$\frac{\mathrm{d}^6\sigma}{\mathrm{d}x_B\,\mathrm{d}y\,\mathrm{d}z_h\,\mathrm{d}\phi_S\,\mathrm{d}^2P_{h\perp}} = \frac{\alpha^2}{4z_h s x_B Q^2}\,L_{\mu\nu}(l,l',\lambda_e)\;2MW^{\mu\nu}(q,P,S,P_h). \tag{2.46}$$

The hadronic tensor is

$$2MW^{\mu\nu}(q,P,S,P_h) = \frac{1}{(2\pi)^4}\sum_{\mathcal{X}'}\int\frac{\mathrm{d}^3P_{\mathcal{X}'}}{2P^0_{\mathcal{X}'}}\;2\pi\,\delta^{(4)}\Big(q+P-P_{\mathcal{X}'}-P_h\Big) \tag{2.47}$$

$$H^{\mu\nu}(P,S,P_{\mathcal{X}'},P_h) = \big\langle P,S\big|\,J^\mu(0)\,\big|P_h,\mathcal{X}'\big\rangle\big\langle P_h,\mathcal{X}'\big|\,J^\nu(0)\,\big|P,S\big\rangle, \tag{2.48}$$

where the state $\big|P_h,\mathcal{X}'\big\rangle$ symbolizes any final state formed by a detected hadron with momentum $\boldsymbol{P}_h$ and a residual state X' which can be reached by acting with the electromagnetic J^μ current operator on the target.

2.2.2 The unpolarized case

In SIDIS off unpolarized targets, like in the DIS case, the hadronic tensor can be parametrized in terms of structure functions, requiring specific symmetry properties

hermiticity:	$W^*_{\mu\nu}(q,P,S) = W_{\nu\mu}(q,P,S),$	(2.49a)
parity:	$L^\rho_\mu L^\sigma_\nu W_{\rho\sigma}(q,P,S) = W_{\mu\nu}(\tilde{q},\tilde{P},-\tilde{S}),$	(2.49b)
naive time-reversal:	$L^\rho_\mu L^\sigma_\nu W^*_{\rho\sigma}(q,P,S) = W_{\mu\nu}(\tilde{q},\tilde{P},\tilde{S})$	(2.49c)

Here the main difference with the DIS requirements is that the presence of final-state interactions on the detected final hadron prevent the time-reversal invariance to put any constraint, since the "in" and "out" states cannot be interchanged. In order to classify the various contributions, we introduce the naive time-reversal transformation, which is identical to the time-reversal one but for the interchange of initial and final states. In this respect, we will have naive time-reversal odd (naive T-odd) functions without violating the fundamental principle of time-reversal invariance.

Operating with these requirements we can parametrize the hadronic ten-

sor for unpolarized SIDIS in terms of 5 strucure functions:

$$
\begin{aligned}
2MW^{\mu\nu}(q,P,S) = \frac{2z_h}{x_B}\Bigg[& - g_\perp^{\mu\nu} F_{UU,T}(x_B, z_h, P_{h\perp}^2, Q^2) + \hat{t}^\mu \hat{t}^\nu F_{UU,L}(x_B, z_h, P_{h\perp}^2, Q^2) \\
& + \Big(\hat{t}^\mu \hat{h}^\nu + \hat{t}^\nu \hat{h}^\mu\Big) F_{UU}^{\cos\phi_h}(x_B, z_h, P_{h\perp}^2, Q^2) \\
& + \Big(\hat{h}^\mu \hat{h}^\nu + g_\perp^{\mu\nu}\Big) F_{UU}^{\cos 2\phi_h}(x_B, z_h, P_{h\perp}^2, Q^2) \\
& - \mathrm{i}\Big(\hat{t}^\mu \hat{h}^\nu - \hat{t}^\nu \hat{h}^\mu\Big) F_{UU}^{\sin\phi_h}(x_B, z_h, P_{h\perp}^2, Q^2)\Bigg],
\end{aligned}
\tag{2.50}
$$

where we introduced the normalized vector $\hat{h} = \boldsymbol{P}_{h\perp}/|\boldsymbol{P}_{h\perp}|$.

The notation for the structure functions include lower labels indicating the polarization state of the probe and of the target (in this case, U stands for unpolarized), followed when needed by the polarization component of the virtual exchanged photon (L for longitudinal, T for transverse). The upper labels indicate the specific angular dependence originating from the contraction between the various structures in the hadronic tensor and the leptonic tensor, e.g.

$$
\Big(\hat{t}_\mu \hat{l}_\nu + \hat{t}_\nu \hat{l}_\mu\Big)\Big(\hat{t}^\mu \hat{h}^\nu + \hat{t}^\nu \hat{h}^\mu\Big) = -g_\perp^{\mu\nu} \hat{l}_\mu \hat{h}_\nu \equiv \cos\phi_h \tag{2.51}
$$

$$
\Big(\hat{t}_\mu \epsilon_{\perp\nu\rho} \hat{l}^\rho + \hat{t}_\nu \epsilon_{\perp\mu\rho} \hat{l}^\rho\Big)\Big(\hat{t}^\mu \hat{h}^\nu + \hat{t}^\nu \hat{h}^\mu\Big) = -\epsilon_\perp^{\mu\nu} \hat{l}_\mu \hat{h}_\nu \equiv \sin\phi_h \tag{2.52}
$$

The above angles correspond to the definition given in the *Trento conventions* [19].

The resulting cross section is:

$$
\frac{d\sigma}{dx_B\, dy\, dz\, d\phi_h\, dP_{h\perp}^2} = \frac{2\pi\alpha^2}{x_B y\, Q^2}\, \frac{y^2}{2\,(1-\varepsilon)} \tag{2.53}
$$

$$
\begin{aligned}
\times \Bigg\{ & F_{UU,T} + \varepsilon F_{UU,L} + \sqrt{2\,\varepsilon(1+\varepsilon)}\, \cos\phi_h\, F_{UU}^{\cos\phi_h} \\
& + \varepsilon \cos(2\phi_h)\, F_{UU}^{\cos 2\phi_h} + \lambda_e \sqrt{2\,\varepsilon(1-\varepsilon)}\, \sin\phi_h\, F_{LU}^{\sin\phi_h} \Bigg\}
\end{aligned}
\tag{2.54}
$$

The last term drops if the lepton beam is unpolarized and the cos $n\phi_h$ dependent terms vanish if we integrate over the angle ϕ_h of the outgoing hadron. However, this may not correspond to reality, since acceptance effects must always be considered. In our analysis of unpolarized TMDs we will consider

eq. 2.54 integrated over the ϕ_h angle.

Integrating the cross section over $P_{h\perp}$ we obtain

$$\frac{d\sigma}{dx_B\, dy\, dz} = \frac{4\pi\alpha^2}{x_B y Q^2}\, \frac{y^2}{2\,(1-\varepsilon)} \Big(F_{UU,T} + \varepsilon F_{UU,L}\Big) \tag{2.55}$$

Now the structure functions are integrated versions of the previous ones, i.e.

$$F_{UU,T/L}(x_B, z_h, Q^2) = \int \mathrm{d}^2\boldsymbol{P}_{h\perp}\, F_{UU,T/L}(x_B, z_h, P_{h\perp}^2, Q^2). \tag{2.56}$$

We can recover the results for totally inclusive DIS

- integrating eq. (2.54) over the transverse momentum $\boldsymbol{P}_{h\perp}$ of the outgoing hadron
- integrating over the angle ϕ_S
- summing over all hadrons and energies in the final state:

$$\frac{d\sigma(\ell p \to \ell X)}{dx_B\, dy} = \sum_h \int \mathrm{d}z_h\, z_h\, \frac{d\sigma(\ell p \to \ell h X)}{dz\, dx_B\, dy} \tag{2.57}$$

This procedure leads to the result already given in eq. (2.21) for the unpolarized case ($S = \lambda_e = 0$), once we identify

$$\sum_h \int \mathrm{d}z_h\, z_h\, F_{UU,T}(x_B, z_h, Q^2) = F_T(x_B, Q^2), \tag{2.58}$$

$$\sum_h \int \mathrm{d}z_h\, z_h\, F_{UU,L}(x_B, z_h, Q^2) = F_L(x_B, Q^2). \tag{2.59}$$

The choice of a convenient frame to deal with SIDIS is not straightforward, due to the presence of $\boldsymbol{P}_h$. We have two choices:

- FRAME 1: Keep the photon and proton to be collinear, give a transverse component to $\boldsymbol{P}_h$. In terms of light-cone components this means

$$P^\mu = \left[\frac{x_B M^2}{Q\sqrt{2}},\ \frac{Q}{x_B\sqrt{2}},\ 0\right] \tag{2.60}$$

$$q^\mu = \left[\frac{Q}{\sqrt{2}},\ -\frac{Q}{\sqrt{2}},\ 0_T\right] \tag{2.61}$$

$$P_h^\mu = \left[\frac{z_h Q}{\sqrt{2}},\ \frac{M_h^2 + |P_{h\perp}|^2}{z_h Q\sqrt{2}},\ P_{h\perp}\right] \tag{2.62}$$

- FRAME 2: Keep the proton and outgoing hadron to be collinear, give a transverse component to q. In terms of light-cone vectors this means choosing

$$P^\mu = P^+ n_+^\mu + \frac{M^2}{2P^+}\, n_-^\mu \,, \tag{2.63}$$

$$P_h^\mu = P_h^- n_-^\mu + \frac{M_h^2}{2P_h^-}\, n_+^\mu \,. \tag{2.64}$$

In this frame, the photon momentum has a transverse component. If we further fix

$$xP^+ = P_h^-/z = Q/\sqrt{2} \tag{2.65}$$

we can explicitly write the vectors involved as follows

$$P^\mu = \left[\frac{x_B M^2}{Q\sqrt{2}},\ \frac{Q}{x_B\sqrt{2}},\ 0\right] \tag{2.66a}$$

$$P_h^\mu = \left[\frac{z_h Q}{\sqrt{2}},\ \frac{M_h^2}{z_h Q\sqrt{2}},\ 0\right] \tag{2.66b}$$

$$q^\mu = \left[\frac{Q}{\sqrt{2}},\ -\frac{(Q^2 - |q_T|^2)}{Q\sqrt{2}},\ q_T\right] \approx \left[\frac{Q}{\sqrt{2}},\ -\frac{Q}{\sqrt{2}},\ q_T\right] \tag{2.66c}$$

The first choice seems to be the most simple one, but in reality from the theoretical point of view it is better to choose the second one in order to preserve a symmetry between P and P_h. However in our phenomenological analysis of unpolarized TMDs we will use FRAME 1.

In any case, it turns out that if we neglect subleading twist corrections, all vectors in the two frames are approximately the same. The only difference is the presence of $P_{h\perp}$ in FRAME 1 and the presence of q_T in FRAME 2. The two are simply connected by

$$q_T = -zP_{h\perp}. \tag{2.67}$$

2.2.3 Parton model picture of unpolarized SIDIS

In parton model interpretation of SIDIS, the virtual photon strikes one of the quarks inside the nucleon.

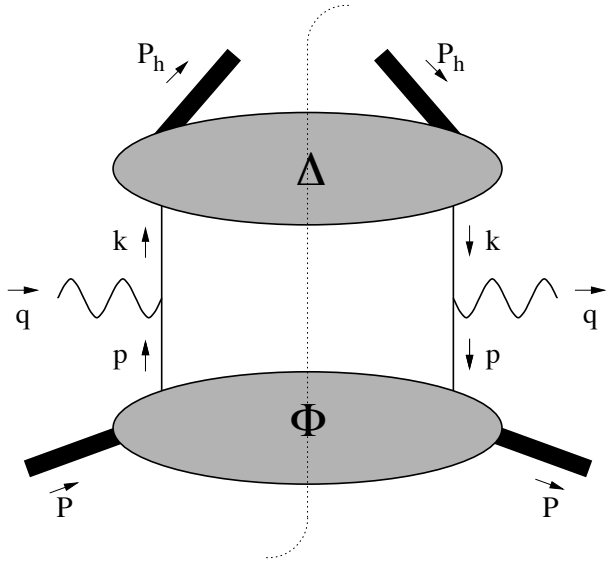

Figure 2.5: The diagram describing the hadronic tensor at tree level.

The parton model approach allows to factorize the scattering process in two *soft* hadronic parts connected by a *hard* scattering part (see fig. 2.5).

Considering only the tree-level contribution (LO processes) to the hard scattering, the hadronic tensor in the QPM picture can be written as

$$2MW^{\mu\nu}(q,P,P_h) = \sum_q e_q^2 \int \mathrm{d}^4p\,\mathrm{d}^4k\;\delta^{(4)}\left(p+q-k\right)\;\mathrm{Tr}\left(\Phi(p,P)\,\gamma^\mu\,\Delta(k,P_h)\,\gamma^\nu\right), \tag{2.68}$$

where p and k are the four-momenta of the struck parton before and after the interaction with the photon, respectively. The Φ function is the quark-quark correlator defined in eq.(2.25) for the inclusive case. The Delta function is a new soft correlator defined as

$$\begin{aligned}\Delta_{kl}(k,P_h) &= \frac{1}{(2\pi)^4}\int \mathrm{d}^4\xi\;\mathrm{e}^{\mathrm{i}k\cdot\xi}\left\langle 0\right|\psi_k(\xi)\left|P_h\right\rangle\left\langle P_h\right|\bar\psi_l(0)\left|0\right\rangle \\ &= \sum_Y \int \frac{\mathrm{d}^3P_Y}{(2\pi)^3\,2P_Y^0}\left\langle 0\right|\psi_k(0)\left|P_h,Y\right\rangle\left\langle P_h,Y\right|\bar\psi_l(0)\left|0\right\rangle \\ &\quad\times\delta^{(4)}\Big(k-P_h-P_Y\Big),\end{aligned} \tag{2.69}$$

where $\left|Y\right\rangle$ is the state (with momentum P_Y) collecting all the fragmentation products apart from the detected hadron (represented by $\left|P_h\right\rangle$).

A suitable parametrization in terms of light-cone coordinates [20] of partonic four-momenta is:

$$p^\mu = \left[\frac{p^2 + |p_T|^2}{2xP^+},\ xP^+,\ p_T\right], \tag{2.70a}$$

$$k^\mu = \left[\frac{P_h^-}{z},\ \frac{z\,(k^2 + |k_T|^2)}{2P_h^-},\ k_T\right]. \tag{2.70b}$$

With these two expressions, neglecting terms which are $1/Q$ suppressed, we can simplify the four-dimensional delta accounting for momentum conservation:

$$\begin{aligned}\delta^{(4)}\,(p+q-k) &\approx \delta(p^+ + q^+)\,\delta(q^- - k^-)\,\delta^{(2)}\Big(p_T + q_T - k_T\Big) \\ &\approx \frac{1}{P^+\,P_h^-}\,\delta(x - x_B)\,\delta(1/z - 1/z_h)\,\delta^{(2)}\Big(p_T + q_T - k_T\Big).\end{aligned} \tag{2.71}$$

Eventually, decomposing the differential element with light-cone coordinates

$$\mathrm{d}^4k = \mathrm{d}^2k_T\ \mathrm{d}k^+\,P_h^-\,\frac{\mathrm{d}\,z}{z^2} \tag{2.72}$$

we can cast the QPM hadronic tensor in compact form

$$2MW^{\mu\nu}(q,P,P_h) = 2z_h\,\mathcal{I}\Big[\mathrm{Tr}(\Phi(x_B,p_T)\,\gamma^\mu\,\Delta(z_h,k_T)\,\gamma^\nu)\Big], \tag{2.73}$$

where we used the shorthand notation

$$\begin{aligned}\mathcal{I}\Big[\cdots\Big] &\equiv \int \mathrm{d}^2p_T\,\mathrm{d}^2k_T\,\delta^{(2)}\Big(p_T + q_T - k_T\Big)\,\Big[\cdots\Big] \\ &= \int \mathrm{d}^2p_T\,\mathrm{d}^2k_T\,\delta^{(2)}\Big(p_T - \frac{P_{h\perp}}{z} - k_T\Big)\,\Big[\cdots\Big],\end{aligned} \tag{2.74}$$

and where the *unintegrated* or *transverse-momentum dependent* correlation functions have been introduced:

$$\Phi(x,p_T,S) \equiv \int \mathrm{d}p^-\Phi(p,P,S)\Big|_{p^+=xP^+}, \tag{2.75a}$$

$$\Delta(z,k_T) \equiv \frac{1}{2z}\int \mathrm{d}k^+\Delta(k,P_h)\Big|_{k^-=P_h^-/z}. \tag{2.75b}$$

Both Φ and Δ are mathematical objects describing the non perturbative physics generated by the color *confinement*:

- the correlation function Φ describes the probability density of finding a parton inside the hadron
- the correlation function Δ describes the way a off-shell parton *decays* into a hadron plus something else, i.e. $q^* \to hY$ (*hadronization* process).

Integrating the correlators over $\boldsymbol{P}_{h\perp}$ we get the fully-integrated functions:

$$2MW^{\mu\nu}(q,P) = 2z_h \,\mathrm{Tr}(\Phi(x_B)\,\gamma^\mu\,\Delta(z_h)\,\gamma^\nu), \tag{2.76a}$$

$$\Phi(x) \equiv \int \mathrm{d}p^-\,\mathrm{d}^2p_T\,\Phi(p,P)\Big|_{p^+=xP^+}, \tag{2.76b}$$

$$\Delta(z) \equiv \frac{z}{2}\int \mathrm{d}k^+\,\mathrm{d}^2k_T\,\Delta(k,P_h)\Big|_{k^-=P_h^-/z}. \tag{2.76c}$$

2.2.4 The unpolarized correlation functions

Following the same procedure developed in the DIS case and keeping only the leading-twist terms, the correlators of eq.(2.75a) can be parametrized as

$$\Phi(x,p_T) = \left\{ f_1(x,p_T^2) + \mathrm{i}h_1^\perp(x,p_T^2)\frac{\not{p}_T}{M}\right\}\frac{\not{n}_+}{2}. \tag{2.77}$$

Here we introduced the transverse-momentum dependent parton distribution functions (TMD PDFs):

$$f_1(x,p_T^2) = 2P^+\int \mathrm{d}p^-\,(A_2 + xA_3)\,, \quad h_1^\perp(x,p_T^2) = 2P^+\int \mathrm{d}p^-\,(-A_4)\,. \tag{2.78}$$

The function $f_1(x,p_T^2)$ is the TMD PDF (or TMD, for brevity). The function $h_1^\perp(x,p_T^2)$ is the *Boer-Mulders TMD* [21].

Analogously we obtain

$$\Delta(z,k_T) = \left(D_1(z,k_T^2) + i\,H_1^\perp(z,\boldsymbol{k}_T^2)\,\frac{\not{k}_T}{M_h}\right)\not{n}_-/2. \tag{2.79}$$

The function $D_1(z,k_T^2)$ is the *unpolarized transverse-momentum-dependent fragmentation function* (known as unpolarized TMD FF). The function $H_1^\perp(z,\boldsymbol{k}_T^2)$ is the *Collins function* [22].

The Boer-Mulders and Collins functions are particularly relevant because they give rise to nontrivial transverse-momentum dependences of cross sections already at the level of unpolarized processes.

Another important property of the Boer-Mulders and Collins functions is that they are T-odd (*naive* time-reversal odd):

$$\Phi^*(p,P,S) = -\mathrm{i}\gamma^1\gamma^3\,\Phi(\tilde{p},\tilde{P},\tilde{S})\,\mathrm{i}\gamma^1\gamma^3. \tag{2.80a}$$

As already stated, the T-odd functions do not violate *proper* time-reversal invariance. They simply give rise to observables that change sign when inverting momenta and angular momenta (e.g., single spin asymmetries [21]).

The correlation function is a 4×4 Dirac matrix. The leading-twist part of the correlator Φ can be projected out using the projector

$$\mathcal{P}_+ = \frac{1}{2}\,\gamma^-\gamma^+, \tag{2.81}$$

Let's also define $F = \Big(\mathcal{P}_+\,\Phi\gamma^+\Big)^T$, i.e. the Dirac transpose of the leading-twist part of the correlation function. It describes the probability to find the dominant component of given chirality of a quark in the target (in jargon, the "good" component [23]). For any Dirac spinor $\big|a\big\rangle$, the expectation value $\big\langle a\,\big|\,F\,\big|a\big\rangle$ must be non-negative. This means that the matrix is *positive semidefinite*, i.e. the determinant of all the principal minors of the matrix has to be positive or zero, imposing constraints on matrix elements. Since the matrix elements are the TMDs themselves, this property has deep phenomenological consequences.

Using the chiral (Weyl) representation, the leading-twist projection of the correlation function reads

$$\Big(\mathcal{P}_+\,\Phi(x,p_T)\gamma^+\Big)_{ji} = \begin{pmatrix} f_1 & 0 & 0 & \mathrm{i}e^{\mathrm{i}\phi_p}\,\dfrac{|p_T|}{M}\,h_1^\perp \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\mathrm{i}e^{-\mathrm{i}\phi_p}\,\dfrac{|p_T|}{M}\,h_1^\perp & 0 & 0 & f_1 \end{pmatrix}. \tag{2.82}$$

The 2×2 relevant part of the Dirac space is the sub-matrix corresponding to good quark fields. This correlation matrix in the good quark chirality space is

$$\Big(\mathcal{P}_+\,\Phi\gamma^+\Big)_{\chi_1'\chi_1} = \begin{pmatrix} f_1 & \mathrm{i}e^{\mathrm{i}\phi_p}\,\dfrac{|p_T|}{M}\,h_1^\perp \\ -\mathrm{i}e^{-\mathrm{i}\phi_p}\,\dfrac{|p_T|}{M}\,h_1^\perp & f_1 \end{pmatrix}. \tag{2.83}$$

The distribution matrix must be Hermitean and positive definite. The latter requirement, as already stated, allows to derive a positivity bound for its elements. For example, requiring a non-negative determinant for the matrix we are left with:

$$\frac{|p_T|}{M} |h_1^\perp(x, p_T^2)| \leq f_1(x, p_T^2). \tag{2.84}$$

2.2.5 The general polarized case

The most complete treatment up to date of one-particle-inclusive deep inelastic scattering at small transverse momentum is given in [11, 15]. In the general polarized case the hadronic tensor can be parametrized in terms of *18 structure functions* [11].

The complete cross section resulting from the contraction of the hadronic and the leptonic tensor is:

$$\frac{d\sigma}{dx_B\, dy\, d\phi_S\, dz\, d\phi_h\, dP_{h\perp}^2} = \frac{\alpha^2}{x_B y\, Q^2}\, \frac{y^2}{2\,(1-\varepsilon)} \tag{2.85}$$

$$\begin{aligned}
&\times \Bigg\{ F_{UU,T} + \varepsilon F_{UU,L} + \sqrt{2\,\varepsilon(1+\varepsilon)}\, \cos\phi_h\, F_{UU}^{\cos\phi_h} \\
&\quad + \varepsilon \cos(2\phi_h)\, F_{UU}^{\cos 2\phi_h} + \lambda_e \sqrt{2\,\varepsilon(1-\varepsilon)}\, \sin\phi_h\, F_{LU}^{\sin\phi_h} \\
&\quad + S_\parallel \left[\sqrt{2\,\varepsilon(1+\varepsilon)}\, \sin\phi_h\, F_{UL}^{\sin\phi_h} + \varepsilon \sin(2\phi_h)\, F_{UL}^{\sin 2\phi_h} \right] \\
&\quad + S_\parallel\, \lambda_e \left[\sqrt{1-\varepsilon^2}\, F_{LL} + \sqrt{2\,\varepsilon(1-\varepsilon)}\, \cos\phi_h\, F_{LL}^{\cos\phi_h} \right] \\
&\quad + |\boldsymbol{S}_\perp| \Bigg[\sin(\phi_h - \phi_S) \left(F_{UT,T}^{\sin(\phi_h-\phi_S)} + \varepsilon\, F_{UT,L}^{\sin(\phi_h-\phi_S)} \right) \\
&\qquad + \varepsilon\, \sin(\phi_h + \phi_S)\, F_{UT}^{\sin(\phi_h+\phi_S)} + \varepsilon\, \sin(3\phi_h - \phi_S)\, F_{UT}^{\sin(3\phi_h-\phi_S)} \\
&\qquad + \sqrt{2\,\varepsilon(1+\varepsilon)}\, \sin\phi_S\, F_{UT}^{\sin\phi_S} + \sqrt{2\,\varepsilon(1+\varepsilon)}\, \sin(2\phi_h - \phi_S)\, F_{UT}^{\sin(2\phi_h-\phi_S)} \Bigg] \\
&\quad + |\boldsymbol{S}_\perp| \lambda_e \Bigg[\sqrt{1-\varepsilon^2}\, \cos(\phi_h - \phi_S)\, F_{LT}^{\cos(\phi_h-\phi_S)} + \sqrt{2\,\varepsilon(1-\varepsilon)}\, \cos\phi_S\, F_{LT}^{\cos\phi_S} \\
&\qquad + \sqrt{2\,\varepsilon(1-\varepsilon)}\, \cos(2\phi_h - \phi_S)\, F_{LT}^{\cos(2\phi_h-\phi_S)} \Bigg] \Bigg\}.
\end{aligned} \tag{2.86}$$

More details about the analytic expression of the various (polarized) structure functions can be found in [11].

2.2.6 Parton model picture of polarized SIDIS

When including also target polarization, starting from the general decomposition presented in eq. (2.33), the *leading order* and *leading twist* part of the TMD correlator becomes

$$\Phi(x, p_T, S) = \frac{1}{2}\left\{ f_1 \not{n}_+ + f_{1T}^{\perp} \frac{\epsilon_T^{\rho\sigma} S_{T\rho} p_{T\sigma}}{M} \not{n}_+ + g_{1s}\gamma_5 \not{n}_+ \right. \\ \left. + h_{1T} \frac{\left[\not{S}_T, \not{n}_+\right]\gamma_5}{2} + h_{1s}^{\perp} \frac{\left[\not{p}_T, \not{n}_+\right]\gamma_5}{2M} + \mathrm{i}\, h_1^{\perp} \frac{\left[\not{p}_T, \not{n}_+\right]}{2M} \right\}$$

where we introduced

$$\epsilon_T^{\alpha\beta} = \epsilon^{\alpha\beta\rho\sigma}\, n_{+\rho} n_{-\sigma}, \tag{2.87}$$

The distribution functions on the r.h.s. depend on x and p_T^2. Here we used the shorthand notation [15]

$$g_{1s}(x, p_T) = S_L\, g_{1L}(x, p_T^2) - \frac{S_T \cdot p_T}{M}\, g_{1T}(x, p_T^2) \tag{2.88}$$

and similarly for the other functions.

Like the Boer-Mulders and the Collins function, $f_{1T}^{\perp}$ (Sivers function) is T-odd [21, 16]. The notation for the distribution functions follows closely that of [15], sometimes referred to as *Amsterdam notation*.

The following names are in common use for the TMDs:

- f_1: unpolarized TMD;
- g_{1L}: helicity TMD;
- h_1: transversity TMD;
- $f_{1T}^{\perp}$: Sivers TMD;
- $h_1^{\perp}$: Boer-Mulders TMD;
- $g_{1T}^{\perp}$: worm-gear TMD, or transversal helicity TMD;
- $h_{1L}^{\perp}$: worm-gear TMD, or Kotzinian-Mulders TMD, or longitudinal transversity TMD;
- $h_{1T}^{\perp}$: pretzelosity TMD or quadrupole TMD.

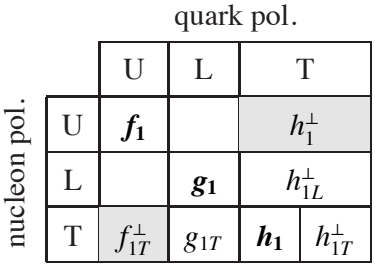

Figure 2.6: Twist-2 transverse-momentum-dependent distribution functions. The U,L,T correspond to unpolarized, longitudinally polarized and transversely polarized nucleons (rows) and quarks (columns). Functions in boldface survive transverse momentum integration. Functions in gray cells are T-odd. At higher twist the number of distribution functions considerably increases [15].

The table in fig. 2.6 lists the twist-2 TMDs with their connection to quark and target polarizations.

The *probabilistic meaning* of TMDs is summarized in fig. 2.7. Assuming that the target nucleon is moving towards the reader, for example the Sivers function $f_{1T}^\perp$ is given by the distortion of the transverse momentum distribution of unpolarized quarks in a transversely polarized nucleon, while the Boer-Mulders function $h_1^\perp$ describes the same distortion but for transversely polarized quarks in unpolarized nucleons. Following this interpretation $f_{1T}^{\perp q} > 0$ corresponds to a preference of the quark to move to the right if the proton is moving towards the observer and the proton spin is pointing upwards. At the same time, $h_1^{\perp q} > 0$ corresponds to a preference of the quark to move to the right if the proton is moving towards the observer and the quark spin is pointing upwards.

2.2.7 Collinear vs TMD PDFs

The connection between collinear and TMD distribution functions defined in (2.35) is

$$f_1(x) = \int \mathrm{d}^2 p_T \; f_1(x, p_T^2), \tag{2.89a}$$

$$g_{1L}(x) = \int \mathrm{d}^2 p_T \; g_{1L}(x, p_T^2), \tag{2.89b}$$

$$h_1(x) = \int \mathrm{d}^2 p_T \; h_1(x, p_T^2). \tag{2.89c}$$

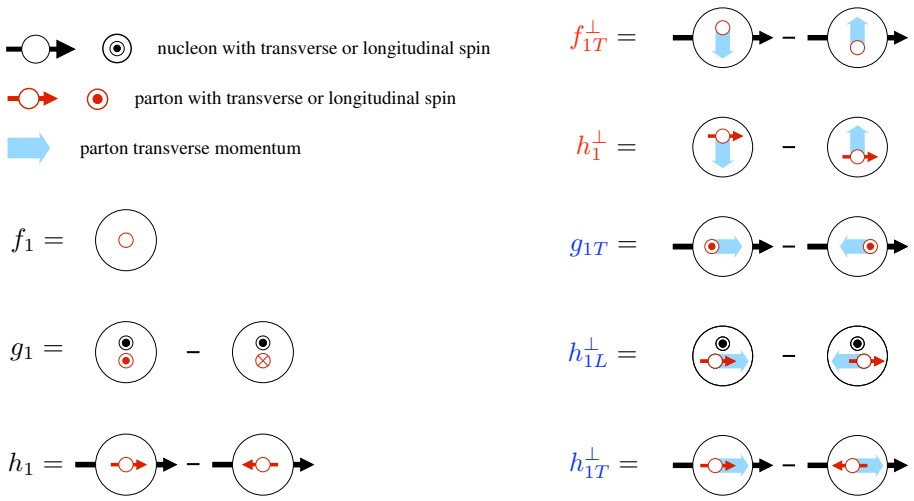

Figure 2.7: Probabilistic interpretation of twist-2 transverse-momentum-dependent distribution functions. To avoid ambiguities, it is necessary to indicate the directions of quark's transverse momentum, target spin and quark spin, specifying that the proton is moving out of the page, or alternatively the photon is moving into the page.

However something dangerous is hidden in the above integrations. In fact, at the parton model level we can assume that TMDs are integrable with total integral approximately equal to their integral in the region up to $p_T^2 \ll Q^2$. This is the mathematical translation of the physical assumption that partons have some *intrinsic* transverse momentum of order $p_T^2 \sim M^2$ or $p_T^2 \sim \Lambda_{QCD}^2$.

This framework is *no more* valid when we take into account *perturbative QCD corrections*. In fact, in this case partons acquire transverse momentum also because of gluon radiation.

The foundational hypothesis used to derive all the results related to SIDIS is t $P_{h\perp}^2 \ll Q^2$. Instead, this approximation have never been used to introduce structure functions in the framework of inclusive DIS. As stated in [11], DIS structure functions are obtained from an integration over $P_{h\perp}^2$ up to a limit of Q^2. They contain therefore *also contributions that are outside the reach of the TMD formalism*. Therefore, we should not expect in general that the integral of TMDs gives back collinear PDFs: other contributions are required to recover the collinear framework.

Is there a *true* relationship between collinear and transverse momentum dependent PDFs? The full QCD formalism tells us that such a relation exists and involves PDFs and the perturbative tail of TMDs at high-transverse-momentum [24].

2.3 Experimental investigations

The physical investigation of the nucleon structure relies on the theoretical framework of Quantum Chromodynamics. Due to its non−abelian character, the coupling constant of the theory has its Landau pole in the limit of small Q^2. Since the mass of the nucleon is ~ 1 GeV (low in comparison to scales involved in high energy physics), we cannot make predictions on its structure starting from the QCD Lagrangian.

For these reasons parton distribution functions and fragmentations functions are not directly accessible through perturbative techniques in quantum field theory: they must be *fitted* on experimental data. There are different kinds of experiments available in order to gain information on the non−perturbative structure of the nucleon and each one has its own peculiarity:

1. **Drell−Yan**

$$\frac{\mathrm{d}\sigma}{\mathrm{d}q_T^2} \sim \sum_q e_q^2 f_1^q(x, p_T^2) \otimes f_1^{\bar{q}}(x, p_T^2) \tag{2.90}$$

   Since the cross section is proportional to a sum of convolutions between distribution functions for the annihilating parton and anti−parton, Drell−Yan processes (hadron collisions with leptonic final states) are suitable to access information on parton and antiparton TMD PDFs. Here x_B is the Bjorken variable, $\boldsymbol{p}_T$ is the transverse momentum of the parton w.r.t. to the direction of the parent hadron momentum and $\boldsymbol{q}_T$ is the transverse momentum of the detected lepton w.r.t. the direction of the vector boson momentum.

2. l^+l^- **annihilation**

$$\frac{\mathrm{d}\sigma}{\mathrm{d}q_T^2} \sim \sum_q e_q^2 D_1^q(z, k_T^2) \otimes D_1^{\bar{q}}(z, \bar{k}_T^2) \tag{2.91}$$

   In l^+l^- annihilations elementary particles collide, so there is no need for any probability function to describe the structure of the interacting

particles. Since many hadrons can be observed in the final states, we can access experimentally TMD FFs, describing the probability density for quarks and antiquarks fragmenting into composite hadrons (confinement-induced hadronization). k_T is related to the transverse momentum of the detected hadron w.r.t. to the jet axis (similarly for $\bar{k}_T$). It holds $q_T = k_T + \bar{k}_T$.

3. **Semi−inclusive deep inelastic scattering**

$$\frac{\mathrm{d}\sigma}{\mathrm{d}q_T^2} \sim \sum_q e_q^2 f_1^q(x, p_T^2) \otimes D_1^q(z, k_T^2) \tag{2.92}$$

Since it deals with leptons scattering off hadrons with one or more hadrons detected in the final state, its cross sections involves convolutions of TMD PDFs and FFs. This is the most complete process in order to gain information on the non perturbative structure of hadrons. Detailed information on both the process (from hadrons to quarks and from quarks to hadrons) are available.

In this thesis we will perform a phenomenological analysis of SIDIS data, collected at the COMPASS experiment. Extending the analysis also to data available from the HERMES collaboration and from Jefferson Lab's facilities would be of great interest to challenge the validity of the results.

In the next sections we will give some information about COMPASS and HERMES experiments, stressing the differences that could have an impact on the phenomenology of distribution and fragmentation functions. For the aim of exploring the flavour dependence of *unpolarized* TMDs, we are interested in unpolarized semi-inclusive deep inelastic scattering data. Two collaborations are about to publish data related to such analyses: COMPASS (at CERN) and HERMES (at DESY). In the present thesis we will analyze *preliminary* results by COMPASS. Data analysis in the HERMES group, instead, is still ongoing. We will investigate *Gaussian fits* of hadron multiplicities, not the experimental data. Possible extensions and upgrades of the analysis presented in this thesis are:

- repeating the analysis on data delivered by COMPASS, once they will be published (expected time 1 month, approximately);
- analyzing unpolarized SIDIS events collected at the HERMES experiment: its apparatus, in fact, have some advantages in comparison with the one at COMPASS:

1. *hadron identification system*:
   HERMES can establish (given a confidence level) the nature of the detected hadrons. This feature has not been used by the COMPASS collaboration in the preliminary analysis (even if available, in principle)
2. *gas target*:
   the gaseous nature of the target allows to skip nuclear corrections necessary in COMPASS analysis.

2.3.1 COMPASS experiment

COMPASS stands for COmmon Muon and Proton Apparatus for Structure and Spectroscopy. It originates from the association of two projects. The first one, HMC, was a proposal for polarized DIS to investigate the structure of the nucleons, in particular the contribution of the gluons to the spin of the nucleon. The second one, CHEOPS, was a proposal for the study of hadronic structure, semi-leptonic decay of charmed baryons and search for exotic baryon states. The goals of those two projects could be reached by one fixed target spectrometer in the North Area of the CERN SPS. The joint project COMPASS was proposed in order to reach the goals of both HMC and CHEOPS [25]. In its first years of data taking (2002-2007) and during the last two (2010-2011), the nucleon spin structure was studied with a polarized muon beam scattering off a polarized target. The years 2008 and 2009, instead, were dedicated to hadron spectroscopy using hadron beams.
In the case of the nucleon structure studies, it is crucial to detect with high precision the incoming beam muon (160 GeV), the scattered muon and the produced hadrons. The large amount of high quality data accumulated provides access to the *unpolarized* and *polarized parton distributions* of the nucleon and to the functions describing the *hadronization* process. Subtle differences between polarized cross sections (*asymmetries*) have been predicted for hadron production from polarized muon-nucleon interaction for COMPASS. It is based on these differences that the polarized parton distributions can be measured. On the contrary, from the analysis of unpolarized semi-inclusive differential cross sections (or the closely related differential multiplicities) it is possible to access unpolarized parton distribution functions.

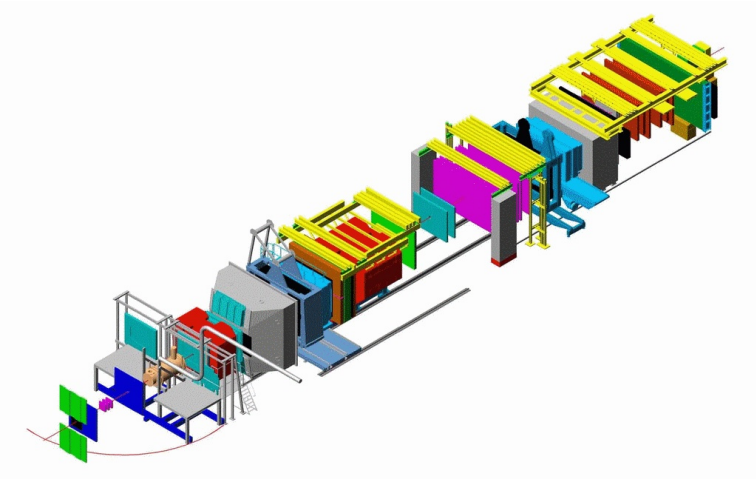

Figure 2.8: Artistic view of the 60 m long COMPASS two-stage spectrometer. The two dipole magnets are indicated in red.

Chapter 3

Flavour dependence of unpolarized TMDs

From now on, we will focus on totally unpolarized SIDIS data, in order to extract hints on the *flavour dependence* of *unpolarized* TMDs.

The process of semi-inclusive deep inelastic lepton-nucleon scattering has been shown to factorize [26] in the low transverse momentum regime ($P_{h\perp}^2 \ll Q^2$), into lepton-quark scattering followed by quark hadronization. In the unpolarized case and at leading twist the probability of producing a hadron with a transverse momentum $\boldsymbol{P}_{h\perp}$ relative to the virtual photon direction is described by a *convolution* involving the TMD parton distribution functions f_1^a and parton fragmentation functions D_1^a. The dependence of these functions on kinematical variables, and the related physical meaning, will be investigated.

Our goal is to explore the flavour dependence of the convolution by investigating preliminary results on hadron multiplicities delivered by the COMPASS collaboration. Since we are interested mainly in non-perturbative contributions to transverse motion of partons, we will analyze events with $P_{h\perp}^2 < 0.7$ GeV^2.

3.1 Relevance of TMDs

In the previous chapter we described the theory of semi-inclusive lepton-nucleon scattering introducing the transverse-momentum-dependent quark-quark correlators, from which distribution functions and fragmentation functions are extracted. From the phenomenological point of view much is known about the dependence on the light-cone momentum fraction x and virtuality scale Q^2. In contrast, very little is presently known about the dependence of these functions on the transverse momentum.

Usually, a Gaussian behaviour is assumed for this dependence. This assumption is known to be inadequate at $P_{h\perp}^2 \gg M^2$: in this regime, in fact, perturbative QCD effects produce significant deviations. However, the Gaussian ansatz can effectively describe data at $P_{h\perp}^2 \approx M^2$, a regime where there are plenty of experimental data available to conduct interesting phenomenological analysis.

The knowledge of details on just the simplest TMDs, the unpolarized one, can have a significant impact on high-energy physics. As en example, let's consider fig. 3.1, where the cross section for Z boson production at the Tevatron is plotted [27]. The difference between the curves originates from *different models of the non perturbative component of the TMDs.* Also, in [28] the CDF collaboration discusses several ways to fit the W mass. According to the analysis, TMDs uncertainties generate an error of 3.9 MeV on the W mass determination (the total systematic error is about 34 MeV, so TMDs account for $\sim 10\%$ of the systematic error, a non-negligible percentage).

This comparison clearly shows that the knowledge of TMDs is essential for *precision* studies in high energy physics: the determination of two of the fundamental free parameters of the Standard Model, the mass of the electroweak gauge bosons, is affected by the uncertainties related to the non perturbative structure of *unpolarized* TMDs.

An important ingredient of the analysis is the capability of separating the contribution of different flavors. As a rule of thumb, since collinear PDFs and FFs have a clear dependence on flavour, we would expect that their 3-dimensional generalizations also display a dependence on flavour. Also, in this thesis we will demonstrate that the Gaussian ansatz with flavour dependence fits data better than the same model without flavour dependence. Reliable experimental results are needed in order to study in detail the flavor structure of TMDs. Interesting data are coming from the ongoing analysis at Jefferson Lab, or performed by the COMPASS and HERMES collaborations.

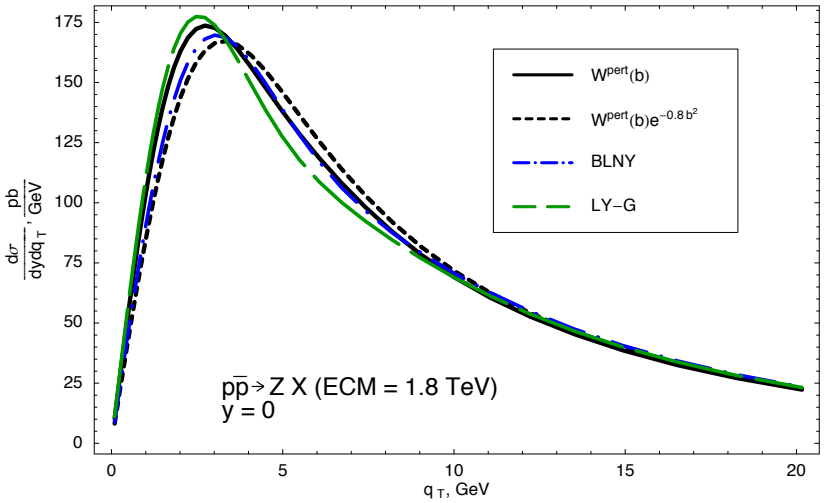

Figure 3.1: The cross sections of Z boson production at the Tevatron [27]. The difference between the curves shows the impact of choosing different non perturbative models for the TMDs.

3.2 Flavour independent Gaussian ansatz

The flavour independent *Gaussian ansatz* consists in assuming *flavour-independent* Gaussian isotropic behaviour in the transverse plane, both for PDFs and FFs:

$$f_1^a(x,p_T^2) = \frac{f_1^a(x)}{\pi\sigma_f^2} e^{-p_T^2/\sigma_f^2}, \qquad D_1^a(z,K_T^2) = \frac{D_1^a(z)}{\pi\sigma_D^2} e^{-K_T^2/\sigma_D^2}, \tag{3.1}$$

where $\sigma_f^2 = \langle p_T^2 \rangle$ and $\sigma_D^2 = \langle K_T^2 \rangle$ are the variances of the transverse momentum dependence of PDFs and FFs, respectively (for the definition of $\boldsymbol{p}_T$ and $\boldsymbol{k}_T$ see next section). The factorization between collinear and transverse kinematics is only apparent: $\langle p_T^2 \rangle$ and $\langle K_T^2 \rangle$, in fact, depend in turn on the kinematic variables x and Q^2.

The Gaussians in the square modulus of the transverse momenta are the result of the multiplication of two Gaussian functions one for each component of $\boldsymbol{p}_T$ and $\boldsymbol{K}_T$. In the case of the PDF, for example, assuming isotropy in the transverse plane ($\sigma_x = \sigma_y = \sigma$) [1] we have:

$$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{p_{T,x}^2}{2\sigma^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{p_{T,y}^2}{2\sigma^2}} = \frac{1}{\pi\sigma_f^2} e^{-\frac{p_T^2}{\sigma_f^2}}, \tag{3.2}$$

[1] This is certainly true in the unpolarized case.

where $\sigma_f^2 = \sigma_x^2 + \sigma_y^2 = 2\sigma^2$ and $p_T^2 = p_{T,x}^2 + p_{T,y}^2$. The same holds for the unpolarized fragmentation function. Currently, in phenomenological applications, unpolarized quark TMDs are parametrized by flavor-independent Gaussians with $\sqrt{\langle p_T^2 \rangle} \approx 0.4 - 0.8$ GeV, depending on the kinematic region explored.

There are some advantages concerning the Gaussian assumption. First of all, Gaussians are mathematically easy to handle. Moreover, since the convolution of two Gaussians is still a Gaussian, this framework allows to easily extract the mean squared transverse momenta of the struck parton $\langle p_T^2 \rangle$ and the fragmenting hadron $\langle K_T^2 \rangle$ in SIDIS experiments from $\boldsymbol{P}_{h\perp}$ measurements (see Appendix).
The Gaussian ansatz agrees well with experimental data for low values of the transverse momentum [29]. As already stated, for higher values perturbative corrections are needed, leading to significant deviations from the Gaussian behaviour (high momentum *tails*).
Experimental distributions in the transverse momentum of detected hadrons clearly show also a sensitivity to the flavor of active partons which is not included in eq.(3.1). But summing upon different flavor dependent convolutions of TMD PDFs and FFs is equivalent to summing different Gaussians, which results in a non-Gaussian distribution. Moreover, in the limit of zero transverse momentum Gaussians are compatible only with the shape of a s-wave function, which seems not a natural constraint. Finally, many popular models show non Gaussian analytic dependences in the transverse momentum, with important p-wave components [2].

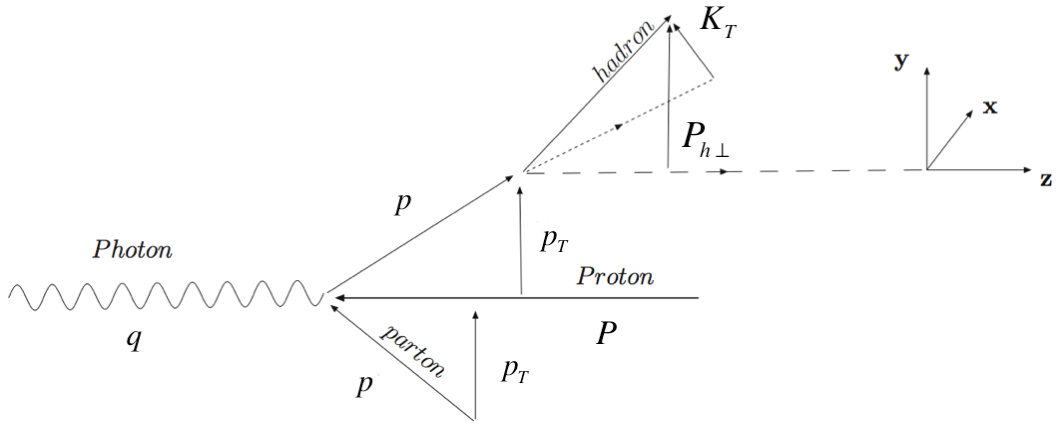

Figure 3.2: Diagram describing the amplitude for a SIDIS event: a virtual photon with four-momentum q strikes a parton with transverse momentum $\boldsymbol{p}_T$. The struck quark fragments in a hadron with transverse-momentum $\boldsymbol{P}_{h\perp}$.

3.3 Totally unpolarized SIDIS

Let's consider SIDIS with unpolarized beam and unpolarized target. Transverse momenta involved in our analysis are (see fig. 3.2):

- $\boldsymbol{p}_T$, the transverse momentum of the parton w.r.t. the nucleon impulse direction
- $\boldsymbol{K}_T$, transverse momentum of the detected hadron, acquired during the fragmentation process, w.r.t. the struck quark impulse direction
- $\boldsymbol{P}_{h\perp}$, transverse momentum of the detected hadron w.r.t. the photon impulse direction.

Another useful variable is $\boldsymbol{k}_T$, the transverse momentum of the struck parton (after the interaction with the photon) w.r.t. the detected hadron impulse direction. Describing the process in terms of $\boldsymbol{p}_T$ and $\boldsymbol{k}_T$ would be better from the theoretical point of view, since the two are both related to the parton, before and after the interaction. However in our analysis we will use the phenomenological variable $\boldsymbol{K}_T$. In order to switch to the $\boldsymbol{k}_T$-dependent description it's sufficient to use the relation

$$-z\boldsymbol{k}_T \;=\; \boldsymbol{K}_T \;. \tag{3.3}$$

The differential cross section in $x_B,\ y,\ z,\ \boldsymbol{P}_{h\perp}$ can be computed integrating 2.54 on ϕ_h and disregarding the last term, depending on the lepton beam polarization λ_e. We obtain:

$$\frac{d^{(4)}\sigma}{dx_B\,dy\,dz\,dP_{h\perp}^2} = \frac{2\pi\alpha^2}{xyQ^2}\,\frac{y^2}{2\,(1-\varepsilon)}\left(F_{UU,T}(x,z,P_{h\perp}^2,Q^2) + \varepsilon F_{UU,L}(x,z,P_{h\perp}^2,Q^2)\right). \tag{3.4}$$

Defining the convolution between unpolarized distribution functions and fragmentation functions as in [11] [2]

$$f\otimes D = x_B\int d^2\boldsymbol{p}_T\,d^2\boldsymbol{K}_T\,\delta^{(2)}\big(z\boldsymbol{p}_T+\boldsymbol{K}_T-\boldsymbol{P}_{h\perp}\big)\,f(x_B,p_T^2)\,D(z,K_T^2)\ , \tag{3.5}$$

we can express the unpolarized structure functions in compact form:

$$F_{UU,T} = \sum_a e_a^2 f_1^a\otimes D_1^a, \qquad\qquad F_{UU,L} = \mathcal{O}\left(\frac{M^2}{Q^2},\frac{P_{h\perp}^2}{Q^2}\right). \tag{3.6}$$

where a is the flavour index. Before going further, we need to investigate the physical meaning of the *delta* distribution inside the convolution.

As a first step, let's define $\hat{z}$ as the light-cone fraction of the parton momentum carried by the detected hadron. The relationship between $\hat{z}$ and the previously defined z variable is:

$$\hat{z} = z + \mathcal{O}\left(\frac{p_T^2}{Q^2}\right)\ . \tag{3.7}$$

The relationship ensuring transverse momentum conservation is:

$$\boldsymbol{P}_{h\perp} = \hat{z}\boldsymbol{p}_T + \boldsymbol{K}_T\ . \tag{3.8}$$

When $p_T^2 \ll Q^2$, we can approximate eq.(3.8) using eq.(3.7):

$$\boldsymbol{P}_{h\perp} = z\boldsymbol{p}_T + \boldsymbol{K}_T + \mathcal{O}\left(\frac{p_T^2}{Q^2}\right)\ . \tag{3.9}$$

The equation emerging in this limit is exactly the statement enclosed in the 2-dimensional delta function:

$$\boldsymbol{P}_{h\perp} \approx z\boldsymbol{p}_T + \boldsymbol{K}_T\ . \tag{3.10}$$

[2] With respect to [11], we change variable from $\boldsymbol{k}_T$ to $\boldsymbol{K}_T$: $d^{(2)}\boldsymbol{K}_T = z^2 d^{(2)}\boldsymbol{k}_T$ and $\delta^{(2)}(a\boldsymbol{x}) = \frac{1}{|a|^2}\,\delta^{(2)}(\boldsymbol{x})$

The Gaussian ansatz allows to break the convolution and to access the values of the transverse momenta of the struck parton ($\boldsymbol{p}_T$) and of the fragmenting hadron ($\boldsymbol{K}_T$). Since the convolution of two Gaussian functions is a Gaussian function the unpolarized structure function $F_{UU,T}(x, z, Q^2, P_{h\perp}^2)$ is a Gaussian in $\boldsymbol{P}_{h\perp}$ [3]:

$$F_{UU,T}(x, z, Q^2, P_{h\perp}^2) = x\left[\sum_a e_a^2 f_1^a(x, Q^2) D_1^a(z, Q^2)\right] \frac{1}{\pi(z^2\langle p_T^2\rangle + \langle K_T^2\rangle)} e^{-\frac{P_{h\perp}^2}{z^2\langle p_T^2\rangle + \langle K_T^2\rangle}} \tag{3.11}$$

$F_{UU,T}$ factorizes in the product of a collinear part and a transverse-momentum-dependent part. The first one is flavour dependent, whereas the second one is not.

Since we are investigating SIDIS phenomenology in the framework of the Gaussian ansatz, we will work at $P_{h\perp}^2 \ll Q^2$.
For this reason in our analysis we can neglect the longitudinal structure function $F_{UU,L}$, which is suppressed as $1/Q^2$ (leading twist approximation). Also, since $M^2 \ll Q^2$, we can neglect mass corrections. Along the same line, the term in eq.(3.4) involving the depolarization factor ε becomes:

$$\frac{y^2}{2(1-\varepsilon)} \;=\; \frac{1}{1+y^2}(1 - y + \frac{1}{2}y^2 + \frac{1}{4}\gamma^2 y^2) \;\approx\; (1 - y + \frac{1}{2}y^2). \tag{3.12}$$

In the same approximation y is related to x and Q^2 trough the relation

$$xy(s - M^2) = Q^2, \tag{3.13}$$

where s is the square of the energy in the center of mass frame and M is the mass of the target hadron. Putting all together, the leading twist cross section for unpolarized SIDIS results:

$$\frac{d^{(4)}\sigma}{dx_B\; dQ^2\; dz\; dP_{h\perp}^2} = \frac{\pi\alpha^2}{xQ^4}\left[1 + \left(1 - \frac{Q^2}{x(s - M^2)}\right)\right]\; F_{UU,T}(x, z, Q^2, P_{h\perp}^2)\;. \tag{3.14}$$

[3] see Appendix for the explicit calculation.

3.3.1 Isospin symmetry and charge conjugation

In the preliminary data presented by the COMPASS collaboration no particle identification systems have been used. For this reason, we can only distinguish positive hadrons from negative ones. Let's assume that these detected hadrons are *pions*. Since muons at COMPASS scatter off protons we have the following fragmentation processes:

$$\begin{aligned} u, \bar{u} &\rightarrow \pi^{+/-} \\ d, \bar{d} &\rightarrow \pi^{+/-} \\ \text{sea} &\rightarrow \pi^{+/-} \end{aligned}$$

Here we separated up and down valence quarks from a generic flavour-symmetric sea contribution. In order to significantly simplify our study, our analysis will rely on two *approximations*, simultaneously applied to fragmentation functions: *isospin symmetry* and *charge conjugation*.
Fragmentation functions will be invariant under transformation from one quark to its isospin-doublet partner (e.g. $u \leftrightarrow d$) and from a positive (negative) hadron to its negative (positive) counterpart (e.g. $\pi^+ \leftrightarrow \pi^-$).
According to these approximations we are left with *two* independent fragmentation functions only. The *favoured* one:

$$D_1^{u\to\pi^+} = D_1^{\bar{d}\to\pi^+} = D_1^{\bar{u}\to\pi^-} = D_1^{d\to\pi^-} \doteq D_1^{fav} \,, \tag{3.15}$$

and the *unfavoured* one:

$$D_1^{d\to\pi^+} = D_1^{\bar{u}\to\pi^+} = D_1^{\bar{d}\to\pi^-} = D_1^{u\to\pi^-} = D_1^{sea\to\pi^+} = D_1^{sea\to\pi^-} \doteq D_1^{unf} \,. \tag{3.16}$$

Fragmentation processes for which the flavour of the fragmenting parton is the same as the valence content of the detected hadron are defined to be *favoured*. Otherwise the process is classified as *unfavoured*.

The biggest difference between the two classes is the number of $q\bar{q}$ pairs excited from the vacuum in order to produce the detected hadron. Favored processes require the creation of one $q\bar{q}$ pair only to make the struck parton hadronizing in a pion, unfavoured processes require more than one (see fig. 3.3).

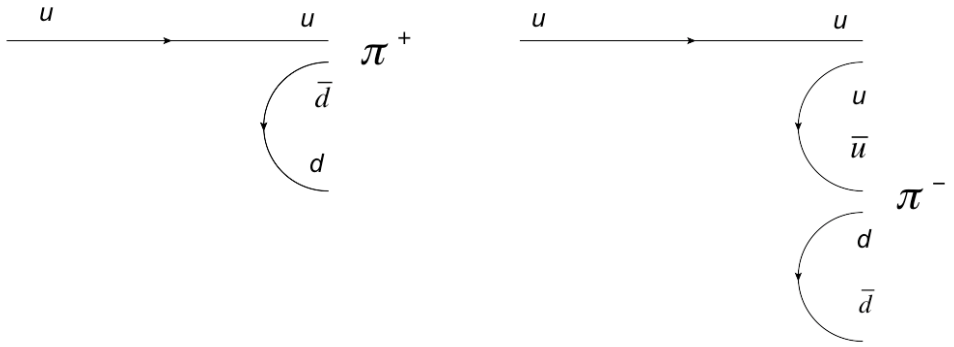

Figure 3.3: Favored fragmentation processes (left) vs unfavoured ones (right). The number of $q\bar{q}$ pairs excited from the sea discriminates the two sets.

3.3.2 Hadron multiplicities

We will study the flavour dependence of unpolarized TMDs trough a phenomenological analysis of *hadron multiplicities* for SIDIS.

The hadron multiplicity is the differential number of hadrons produced per DIS scattering event:

$$m_h(x,z,Q^2,P_{h\perp}^2) = \frac{\frac{dN^h}{dxdQ^2dzdP_{h\perp}^2}}{\frac{dN_{DIS}}{dxdQ^2}} \ . \tag{3.17}$$

First of all we need to compute the cross section for the totally inclusive DIS process. According to [11], the inclusive DIS cross section is:

$$\frac{d^{(2)}\sigma}{dxdy} = \frac{\alpha^2}{xyQ^2}\Big[1+(1-y)^2\Big]F_T \ , \tag{3.18}$$

or

$$\frac{d^{(2)}\sigma}{dxdQ^2} = \frac{\alpha^2}{xQ^4}\left[1+\left(1-\frac{Q^2}{x(s-M^2)}\right)^2\right]F_T \ . \tag{3.19}$$

$F_T(x,Q^2)$ is the transverse structure function, integrated version of $F_{UU,T}$:

$$F_T(x,Q^2) = \sum_h \int_0^1 dz\ z\ \int d^{(2)}\boldsymbol{P}_{h\perp} F_{UU,T}(x,z,Q^2,P_{h\perp}^2) \ . \tag{3.20}$$

Since

$$\int d^{(2)}\boldsymbol{P}_{h\perp}\frac{1}{\pi\langle P_{h\perp}^2\rangle}\exp\left\{-\frac{P_{h\perp}^2}{\langle P_{h\perp}^2\rangle}\right\} = 1 \ , \tag{3.21}$$

$$\sum_h \int_0^1 dz z D_1^{a\to h}(z) = 1 \ , \tag{3.22}$$

we have

$$F_T(x,Q^2) = \sum_h \int_0^1 dzz \int d^{(2)}\boldsymbol{P}_{h\perp} F_{UU,T}(x,z,Q^2,P_{h\perp}^2) = x\sum_a e_a^2 f_1^a(x)\ . \tag{3.23}$$

The number of the detected particles is related to the cross section of the process through luminosity $\mathcal{L}$ and acceptance $\mathcal{A}$ factors [4]:

$$\frac{d^{(n)}N}{d\boldsymbol{x}} = \mathcal{L}\ \mathcal{A}\ \frac{d^{(n)}\sigma}{d\boldsymbol{x}}. \tag{3.24}$$

At COMPASS acceptance factors for SIDIS are factorized in two parts [30]:

$$\mathcal{A}_{SIDIS} = \mathcal{A}_{incl}\ \mathcal{A}_{h_{+/-}}\ . \tag{3.25}$$

Using eq. 3.24 and 3.25 we can express hadron multiplicities in eq. 3.17 as

$$m_h(x,z,Q^2,P_{h\perp}^2) = \frac{\mathcal{L}\ \mathcal{A}_{incl}\ \mathcal{A}_{h_{+/-}}\ \frac{d^{(4)}\sigma^h_{SIDIS}}{dxdQ^2dzdP_{h\perp}^2}}{\mathcal{L}\ \mathcal{A}_{incl}\ \frac{d^{(2)}\sigma_{DIS}}{dxdQ^2}} = \mathcal{A}_{h_{+/-}}\ \frac{d^{(4)}\sigma^h_{SIDIS}}{d^{(2)}\sigma_{DIS}}\ . \tag{3.26}$$

In this way we can express hadron multiplicities as ratios of cross sections corrected by the hadron production-related acceptance factor.

Working in the framework of Gaussian approximation, Jean-Francois Rajotte extracted Gaussian fits of hadron multiplicities for SIDIS at COMPASS:

$$m_h(x,z,Q^2,P_{h\perp}^2) = \mathcal{A}_{h_{+/-}}\ \frac{d^{(4)}\sigma^h_{SIDIS}}{d^{(2)}\sigma_{DIS}} \sim A_{+/-}\ \exp\left\{-\frac{P_{h\perp}^2}{\langle P_{h\perp}^2\rangle}\right\}\ , \tag{3.27}$$

where $A_{+/-}(x,z,Q^2)$ and $\langle P_{h\perp}^2\rangle(x,z,Q^2)$ are *best-fit* parameters, whose values are functions of the kinematical variable (x,z,Q^2). In [30] best-fit values for $A_{+/-}$ and $\langle P_{h\perp}^2\rangle$ are collected, each with its *statistical* error.

The kinematic intervals covered by COMPASS are [3]:

$$0.0045 < x_B < 0.12 \tag{3.28}$$

$$0.05 < \langle z^2\rangle < 0.56 \tag{3.29}$$

$$1 < Q^2 < 10\ \mathrm{GeV}^2\ , \tag{3.30}$$

[4]Correction for acceptance are required in order to take into account the limited solid angle ($< 4\pi$) covered by the detector and effects due to limited efficiencies.

grouped in 184 3-dimensional bins in (x, z, Q^2) (see fig. 3.4).

From eq.(3.26) the expression of $m_h(x, z, Q^2, P_{h\perp}^2)$ with the flavour independent Gaussian ansatz is [5]:

$$m_h(x, z, Q^2, P_{h\perp}^2) = \pi \frac{\sum_a \left[e_a^2 f_1^a(x, Q^2) D_1^a(z, Q^2) \right]}{\sum_a e_a^2 f_1^a(x, Q^2)} \frac{e^{-\frac{P_{h\perp}^2}{z^2\langle p_T^2\rangle + \langle K_T^2\rangle}}}{\pi(z^2\langle p_T^2\rangle + \langle K_T^2\rangle)} \ . \quad (3.31)$$

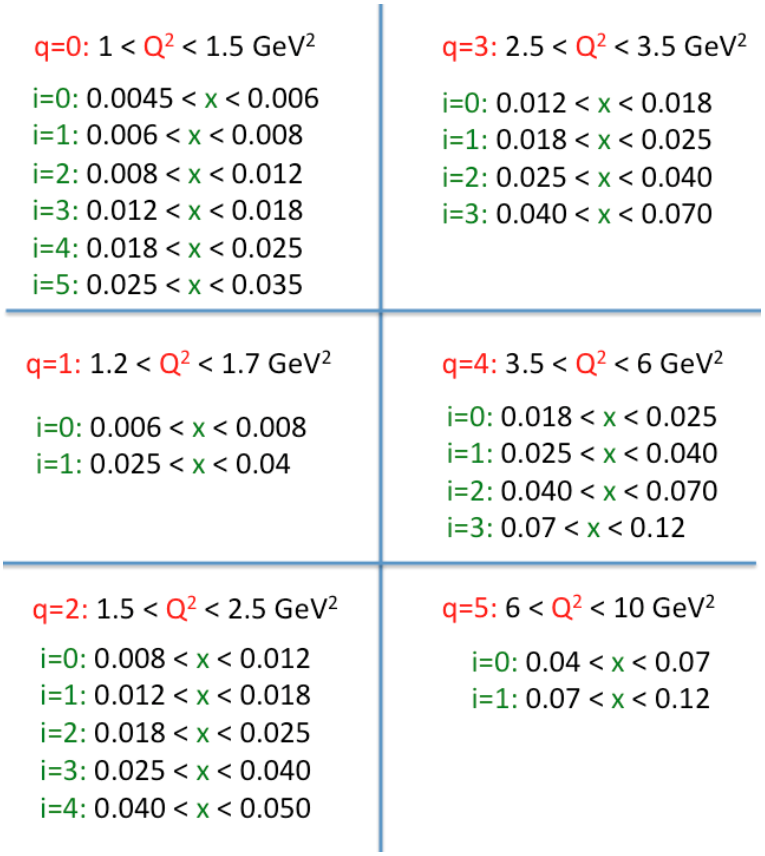

Figure 3.4: The 184 kinematic bins in (x, z, Q^2) analyzed by COMPASS. We will group them in six Q^2 bins (red ones), each with a sub-structure in x (green ones) and z bins (not represented). For every x-bin there are eight z-bins for positive hadrons and eight z-bins for negative hadrons.

[5] For explicit calcuations see Appendix.

3.4 Flavour dependent multiplicities

The goal of our work consists in generalizing eq.(3.31) by introducing the flavor dependence also in the transverse momentum dependence:

$$m_h(x,z,Q^2,P_{h\perp}^2) = \frac{1}{z^2\langle p_T^2\rangle + \langle K_T^2\rangle} \frac{\sum_a e_a^2 f_1^a(x,Q^2) D_1^a(z,Q^2) e^{-\frac{P_{h\perp}^2}{z^2\langle p_{T,a}^2\rangle + \langle K_{T,a}^2\rangle}}}{\sum_a e_a^2 f_1^a(x,Q^2)} , \tag{3.32}$$

where $\langle p_{T,a}^2\rangle$ is the mean square transverse momentum of the parton a. Moreover, $\langle K_{T,a}^2\rangle$ is the mean square transverse momentum of the detected hadron, fragmenting from parton a. Depending on the flavour, we will classify it as $\langle K_{T,fav}^2\rangle$ (favored) or $\langle K_{T,unf}^2\rangle$ (unfavoured).

We separate the summation on flavours in *valence* and *sea* sectors, considering up, down and strange valence quarks and a *flavour-blind* sea. The strange quark is added to the valence sector with the constraint on its valence distribution

$$\int_0^1 dx(st(x) - \overline{st}(x)) = 0 , \tag{3.33}$$

namely the probability of finding a valence strange quark in a nucleon is zero. Since the Q^2 ranges between 1 and 10 GeV^2 we are in open-strange, open-charm and open-bottom regimes. Bottom and strange quarks are suppressed by a factor 4 (due to electric charge) with respect to the charm ones. However, the most non-negligible probability density function (evaluated for mean values of x, Q^2) is related to strange quarks, so we will neglect charm and bottom flavors.

For convenience, let's introduce the notations [6]:

$$\langle p_{T,up}^2\rangle \doteq u , \tag{3.34}$$
$$\langle p_{T,down}^2\rangle \doteq d , \tag{3.35}$$
$$\langle p_{T,sea}^2\rangle \doteq s , \tag{3.36}$$
$$\langle K_{T,favoured}^2\rangle \doteq F , \tag{3.37}$$
$$\langle K_{T,unfavoured}^2\rangle \doteq U . \tag{3.38}$$

The $\langle p_{T,sea}^2\rangle$ variance will be used also to describe the TMD Gaussian of the strange quarks, in order to reduce the number of parameters in the fit.

[6]Here s is not to be confused with the square of energy in the centre of mass system.

When the detected hadron is a positive pion the hadron multiplicity becomes:

$$
\begin{aligned}
m_{\pi^+}(x,z,Q^2,P_{h\perp}^2) = \frac{\pi n}{\sum_a e_a^2 f_1^a(x,Q^2)}\cdot \\
\cdot \Bigg\{ & e_u^2 f_1^{u_v}(x,Q^2) D_1^{fav}(z,Q^2) \frac{e^{-\frac{-P_{h\perp}^2}{z^2 u+F}}}{\pi(z^2u+F)} + \\
& + e_d^2 f_1^{d_v}(x,Q^2) D_1^{unf}(z,Q^2) \frac{e^{-\frac{-P_{h\perp}^2}{z^2 d+U}}}{\pi(z^2d+U)} + \\
& + e_{st}^2 f_1^{st_v}(x,Q^2) D_1^{unf}(z,Q^2) \frac{e^{-\frac{-P_{h\perp}^2}{z^2 s+U}}}{\pi(z^2s+U)} + \\
& + e_u^2 f_1^{u-u_v}(x,Q^2) D_1^{fav}(z,Q^2) \frac{e^{-\frac{-P_{h\perp}^2}{z^2 s+F}}}{\pi(z^2s+F)} + \\
& + e_{\bar{u}}^2 f_1^{\bar{u}}(x,Q^2) D_1^{unf}(z,Q^2) \frac{e^{-\frac{-P_{h\perp}^2}{z^2 s+U}}}{\pi(z^2s+U)} + \\
& + e_d^2 f_1^{d-d_v}(x,Q^2) D_1^{unf}(z,Q^2) \frac{e^{-\frac{-P_{h\perp}^2}{z^2 s+U}}}{\pi(z^2s+U)} + \\
& + e_{\bar{d}}^2 f_1^{\bar{d}}(x,Q^2) D_1^{fav}(z,Q^2) \frac{e^{-\frac{-P_{h\perp}^2}{z^2 s+F}}}{\pi(z^2s+F)} + \\
& + e_{st}^2 f_1^{st-st_v}(x,Q^2) D_1^{unf}(z,Q^2) \frac{e^{-\frac{-P_{h\perp}^2}{z^2 s+U}}}{\pi(z^2s+U)} + \\
& + e_{\overline{st}}^2 f_1^{\overline{st}}(x,Q^2) D_1^{unf}(z,Q^2) \frac{e^{-\frac{-P_{h\perp}^2}{z^2 s+U}}}{\pi(z^2s+U)} \Bigg\} ,
\end{aligned}
\tag{3.39}
$$

where n is a global normalization parameter involved in fitting our theoretical expression to experimental data, accounting for acceptance effects and higher-order corrections. The first three terms in the formula are related to valence quarks, the following ones contain sea contributions.

For the case of detected π^- we have:

$$
\begin{aligned}
m_{\pi^-}(x,z,Q^2,P_{h\perp}^2) = \frac{\pi n}{\sum_a e_a^2 f_1^a(x,Q^2)}\cdot \\
\cdot \Bigg\{ & e_u^2 f_1^{u_v}(x,Q^2) D_1^{unf}(z,Q^2) \frac{e^{-\frac{-P_{h\perp}^2}{z^2 u+U}}}{\pi(z^2u+U)} + \\
& + e_d^2 f_1^{d_v}(x,Q^2) D_1^{fav}(z,Q^2) \frac{e^{-\frac{-P_{h\perp}^2}{z^2 d+F}}}{\pi(z^2d+F)} + \\
& + e_{st}^2 f_1^{st_v}(x,Q^2) D_1^{unf}(z,Q^2) \frac{e^{-\frac{-P_{h\perp}^2}{z^2 s+U}}}{\pi(z^2s+U)} + \\
& + e_u^2 f_1^{u-u_v}(x,Q^2) D_1^{unf}(z,Q^2) \frac{e^{-\frac{-P_{h\perp}^2}{z^2 s+U}}}{\pi(z^2s+U)} + \\
& + e_{\bar{u}}^2 f_1^{\bar{u}}(x,Q^2) D_1^{fav}(z,Q^2) \frac{e^{-\frac{-P_{h\perp}^2}{z^2 s+F}}}{\pi(z^2s+F)} + \\
& + e_d^2 f_1^{d-d_v}(x,Q^2) D_1^{fav}(z,Q^2) \frac{e^{-\frac{-P_{h\perp}^2}{z^2 s+F}}}{\pi(z^2s+F)} + \\
& + e_{\bar{d}}^2 f_1^{\bar{d}}(x,Q^2) D_1^{unf}(z,Q^2) \frac{e^{-\frac{-P_{h\perp}^2}{z^2 s+U}}}{\pi(z^2s+U)} + \\
& + e_{st}^2 f_1^{st-st_v}(x,Q^2) D_1^{unf}(z,Q^2) \frac{e^{-\frac{-P_{h\perp}^2}{z^2 s+U}}}{\pi(z^2s+U)} + \\
& + e_{\overline{st}}^2 f_1^{\overline{st}}(x,Q^2) D_1^{unf}(z,Q^2) \frac{e^{-\frac{-P_{h\perp}^2}{z^2 s+U}}}{\pi(z^2s+U)} \Bigg\} .
\end{aligned}
\tag{3.40}
$$

It's important to notice that $m_{\pi^{+/-}}$ are *no more* Gaussian functions in $\boldsymbol{P}_{h\perp}$, since the sum of Gaussian functions is not a Gaussian one.

3.4.1 Kinematic dependence in $\langle K_T^2 \rangle$

As already stated, behind eq. 3.1 there is not a perfect factorization between collinear and TMD part, since $\langle p_T^2 \rangle$ and $\langle K_T^2 \rangle$ depend on x, z and Q^2. In

particular $\langle p_T^2 \rangle$ depends on (x, Q^2), whereas $\langle K_T^2 \rangle$ depends on (z, Q^2).

Let's focus on $\langle K_T^2 \rangle$. Neglecting the flavour index, the assumed functional form is:

$$\langle K_T^2 \rangle(z, Q^2) = N \; z^\alpha \; (1-z)^\beta \left[1 + \ln\left(\frac{Q^2}{Q_0^2} \right) \right] , \tag{3.41}$$

where N, α, β are best-fit parameters and Q_0^2 will be fixed to 1 GeV^2. N is the normalization factor. Concerning the z-dependence, instead:

- $z \to 0$
  We assume $\alpha > 0$; as a consequence, the $\langle K_T^2 \rangle$ becomes negligible in the limit of $z \to 0$ (meaning that the hadron could not have gain much transverse momentum during fragmentation when there is little energy available). This hypothesis, however, could be relaxed allowing also negative values for α.

- $z \to 1$
  The hadron carries almost all of the virtual photon momentum; it probably did not go through many fragmentation steps. The fewer fragmentation processes reduce the chances of the hadrons to acquire transverse momentum during fragmentation. This behavior is reproduced by introducing a factor $(1-z)^\beta$.

The logarithmic factor accounts for the evolution of the TMD part of the functions with the hard-scale Q^2. Similarly to DGLAP equations [6], TMD evolution equations describe the evolution of TMD PDFs and FFs with the hard-scale Q^2. Since the detailed study of unpolarized TMD evolution equations is not considered in our analysis, we will insert this logarithmic factor only as a *template* of the proper behaviour with Q^2 for the mean square transverse momenta.

3.4.2 Kinematic dependence in $\langle p_T^2 \rangle$

Let's now turn to the analysis of $\langle p_T^2 \rangle$. We parameterize it in the same way as in eq. 3.41:

$$\langle p_T^2 \rangle(x, Q^2) = M \; x^\delta \; (1-x)^\gamma \left[1 + \ln\left(\frac{Q^2}{Q_0^2} \right) \right] . \tag{3.42}$$

Here, however, the justification is less intuitive. After the interaction with the virtual photon, the *on-shell* proton splits in one *off-shell* struck parton and an *on-shell* spectator state. M is the normalization factor, while the logarithmic

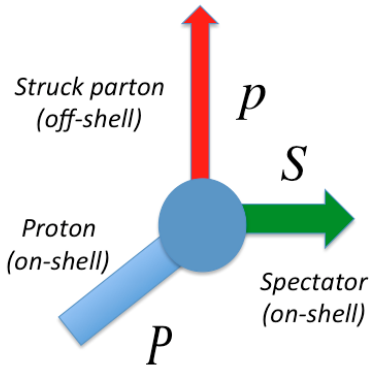

Figure 3.5: Physical states and their momenta before and after the interaction: P is the four-momentum of the on-shell proton, p is the one related to the struck parton and S is the four-momentum of the residual (spectator quarks) state.

term in the hard-scale simulates TMD evolution. In order to explain the x-dependence of $\langle p_T^2 \rangle$ we need to introduce the light-cone [7] expansion of the involved four-momenta. As explained in fig. 3.5, P is the four-momentum of the on-shell proton, p is the one related to the struck parton and S [8] is the four-momentum of the residual (spectator) state. The expression of P in the target rest frame is:

$$P = (P^-, P^+, \boldsymbol{P}_T) \equiv (\frac{M}{\sqrt{2}}, \frac{M}{\sqrt{2}}, \boldsymbol{0}_T) \ . \tag{3.43}$$

Let's define x [9] as the fraction of the light-cone plus component of the proton momentum carried by the parton. We have:

$$p = (p^-, p^+, \boldsymbol{p}_T) \equiv (p^-, xP^+, \boldsymbol{p}_T) \ . \tag{3.44}$$

The four-momentum S can be built using the *approximated* [10] Altarelli-Parisi kinematics. For example, we choose to conserve the plus and the transverse components of its momentum:

$$S = (S^-, P^+ - p^+, \boldsymbol{P}_T - \boldsymbol{p}_T) \equiv (S^-, (1-x)\frac{M}{\sqrt{2}}, -\boldsymbol{p}_T) \ . \tag{3.45}$$

[7] For a concise explanation of *light-cone coordinates* see [20].

[8] Not to be confused with the spin of the target.

[9] At leading order $x \sim x_B$.

[10] The approximated kinematics refers to a framework without complete momentum conservation due to impossibility of the three states to be on-shell at the same time. In our case the parton will be off-shell. (see [6]).

Since the physical state of the spectator must be *real*, the on-shellness condition holds:

$$0 \leqslant S^2 = 2S^+S^- - |\boldsymbol{S}_T|^2 = \sqrt{2}MS^-(1-x) - |\boldsymbol{p}_T|^2 \ , \ \ \forall (x, |\boldsymbol{p}_T|^2, S^-) \ . \tag{3.46}$$

Let's analyze eq.(3.46) by considering the upper bound for x:

- $x \to 1$:

$$x \to 1 \Longrightarrow |\boldsymbol{p}_T| \to 0 \ , \tag{3.47}$$

  since S^2 must be non-negative. The transverse momentum of the parton vanishes in the limit $x \to 1$.

Moreover, we assume that

$$x \to 0 \Longrightarrow |\boldsymbol{p}_T| \to 0 \ . \tag{3.48}$$

This statement is supported by considering eq.(3.46) with the lower bound for x:

- $x \to 0$:

$$x \to 0 \Longrightarrow |\boldsymbol{p}_T|^2 \leqslant \sqrt{2}MS^- \ . \tag{3.49}$$

  The on-shellness condition holds for any $(x, |\boldsymbol{p}_T|^2, S^-)$. Choosing $S^- = 0$ we have:

$$|\boldsymbol{p}_T| \to 0 \ , \tag{3.50}$$

  namely the transverse momentum of the parton vanishes in the limit of zero collinear momentum.

The terms x^δ and $(1-x)^\gamma$ in eq. 3.42 account for the above behaviour.

Here the theoretical study of flavour dependence of unpolarized TMDs ends. In the following chapter we will deepen the analysis from a phenomenological point of view.

3.5 Appendix

3.5.1 Convolution of two Gaussians

In this section we will outline the steps required to break the convolution in $F_{UU,T}$ and obtain eq.(3.11):

$$F_{UU,T} = x \sum_a e_a^2 \int d^{(2)}\boldsymbol{p}_T \int d^{(2)}\boldsymbol{K}_T \, \delta^{(2)}(z\boldsymbol{p}_T + \boldsymbol{K}_T - \boldsymbol{P}_{h\perp}) f_1^a(x, p_T^2) D_1^a(z, K_T^2) \, . \tag{3.51}$$

Inserting the flavour independent Gaussian parametrization of TMDs:

$$F_{UU,T} = x \left[\sum_a e_a^2 f_1^a(x) D_1^a(z) \right] \int d^{(2)}\boldsymbol{p}_T \int d^{(2)}\boldsymbol{K}_T \; \delta^{(2)}(z\boldsymbol{p}_T + \boldsymbol{K}_T - \boldsymbol{P}_{h\perp}) \cdot \\ \cdot \frac{1}{\pi \langle p_T^2 \rangle} e^{-\frac{p_T^2}{\langle p_T^2 \rangle}} \frac{1}{\pi \langle K_T^2 \rangle} e^{-\frac{K_T^2}{\langle K_T^2 \rangle}} \tag{3.52}$$

Let's consider the TMD part, performing the integration on $\boldsymbol{K}_T$ with the delta distribution:

$$\begin{aligned} &\int d^{(2)}\boldsymbol{p}_T \int d^{(2)}\boldsymbol{K}_T \; \delta^{(2)}(z\boldsymbol{p}_T + \boldsymbol{K}_T - \boldsymbol{P}_{h\perp}) \frac{1}{\pi^2 \langle p_T^2 \rangle \langle K_T^2 \rangle} e^{-\frac{p_T^2}{\langle p_T^2 \rangle} - \frac{K_T^2}{\langle K_T^2 \rangle}} = \\ &= \frac{1}{\pi^2 \langle p_T^2 \rangle \langle K_T^2 \rangle} \int d^{(2)}\boldsymbol{p}_T \; e^{-\frac{p_T^2}{\langle p_T^2 \rangle}} \exp\left(- \frac{|\boldsymbol{P}_{h\perp} - z\boldsymbol{p}_T|^2}{\langle K_T^2 \rangle} \right) = \\ &= \frac{1}{\pi^2 \langle p_T^2 \rangle \langle K_T^2 \rangle} \left[\int d^{(2)}\boldsymbol{p}_T \; \exp\left(- \frac{p_T^2}{\langle p_T^2 \rangle} - \frac{z^2 p_T^2}{\langle K_T^2 \rangle} + \frac{2z}{\langle K_T^2 \rangle} (\boldsymbol{P}_{h\perp} \cdot \boldsymbol{p}_T) \right) \right] \; \exp\left(- \frac{P_{h\perp}^2}{\langle K_T^2 \rangle} \right) \, . \end{aligned} \tag{3.53}$$

Let's now focus on the integral in square brackets:

$$\begin{aligned} \Big[\dots \Big] &= \int d^{(2)}\boldsymbol{p}_T \; \exp\left[-p_T^2 \left(\frac{1}{\langle p_T^2 \rangle} + \frac{z^2}{\langle K_T^2 \rangle} \right) + \frac{2z}{\langle K_T^2 \rangle} \boldsymbol{P}_{h\perp} \cdot \boldsymbol{p}_T \right] = \\ &= \int d^{(2)}\boldsymbol{p}_T \; \exp\left[-p_T^2 \frac{z^2 \langle p_T^2 \rangle + \langle K_T^2 \rangle}{\langle p_T^2 \rangle \langle K_T^2 \rangle} + \frac{2z}{\langle K_T^2 \rangle} \boldsymbol{P}_{h\perp} \cdot \boldsymbol{p}_T \right] \, . \end{aligned} \tag{3.54}$$

Completing the argument of the integral to a perfect square term:

$$\int d^{(2)}\boldsymbol{p}_T \ \exp\left[-\frac{z^2\langle p_T^2\rangle+\langle K_T^2\rangle}{\langle p_T^2\rangle\langle K_T^2\rangle}\left(p_T^2-\frac{2z}{\langle K_T^2\rangle}\frac{\langle p_T^2\rangle\langle K_T^2\rangle\boldsymbol{P}_{h\perp}\cdot\boldsymbol{p}_T}{z^2\langle p_T^2\rangle+\langle K_T^2\rangle}\pm\frac{z^2(\langle p_T^2\rangle)^2P_{h\perp}^2}{(z^2\langle p_T^2\rangle+\langle K_T^2\rangle)^2}\right)\right]=$$
$$=\int d^{(2)}\boldsymbol{p}_T \ \exp\left[-\frac{z^2\langle p_T^2\rangle+\langle K_T^2\rangle}{\langle p_T^2\rangle\langle K_T^2\rangle}\left(\left|\boldsymbol{p}_T-\frac{z\langle p_T^2\rangle}{z^2\langle p_T^2\rangle+\langle K_T^2\rangle}\boldsymbol{P}_{h\perp}\right|^2-\frac{z^2\langle p_T^2\rangle^2P_{h\perp}^2}{(z^2\langle p_T^2\rangle+\langle K_T^2\rangle)^2}\right)\right]=$$
$$=\exp\left[\frac{z^2\langle p_T^2\rangle P_{h\perp}^2}{(z^2\langle p_T^2\rangle+\langle K_T^2\rangle)\langle K_T^2\rangle}\right]\int d^{(2)}\boldsymbol{p}_T \ \exp\left[-\frac{z^2\langle p_T^2\rangle+\langle K_T^2\rangle}{\langle p_T^2\rangle\langle K_T^2\rangle}\left|\boldsymbol{p}_T-\frac{z\langle p_T^2\rangle}{z^2\langle p_T^2\rangle+\langle K_T^2\rangle}\boldsymbol{P}_{h\perp}\right|^2\right] . \tag{3.55}$$

Let's collect all the exponential factors involving $\boldsymbol{P}_{h\perp}$:

$$\begin{aligned}&\exp\left\{-\frac{P_{h\perp}^2}{\langle K_T^2\rangle}+\frac{z^2\langle p_T^2\rangle P_{h\perp}^2}{(z^2\langle p_T^2\rangle+\langle K_T^2\rangle)\langle K_T^2\rangle}\right\}=\\&=\exp\left\{-\frac{P_{h\perp}^2}{z^2\langle p_T^2\rangle+\langle K_T^2\rangle}\right\} .\end{aligned} \tag{3.56}$$

Let's define

$$\boldsymbol{p}'_T=\boldsymbol{p}_T-\frac{z\langle p_T^2\rangle}{z^2\langle p_T^2\rangle+\langle K_T^2\rangle}\boldsymbol{P}_{h\perp} . \tag{3.57}$$

The integral in $\boldsymbol{p}_T$ can be resolved using properties of Gaussian functions [11]. Since the integration on $\boldsymbol{p}_T$ is performed at constant $\boldsymbol{P}_{h\perp}$:

$$d^{(2)}\boldsymbol{p}'_T=d^{(2)}\boldsymbol{p}_T . \tag{3.58}$$

and

$$\begin{aligned}&\frac{1}{\pi^2\langle p_T^2\rangle\langle K_T^2\rangle}\int d^{(2)}\boldsymbol{p}'_T \ \exp\left(-\frac{z^2\langle p_T^2\rangle+\langle K_T^2\rangle}{\langle p_T^2\rangle\langle K_T^2\rangle}{p'_T}^2\right)=\\&=\frac{1}{\pi^2\langle p_T^2\rangle\langle K_T^2\rangle}\frac{\pi}{\frac{z^2\langle p_T^2\rangle+\langle K_T^2\rangle}{\langle p_T^2\rangle\langle K_T^2\rangle}}=\\&=\frac{1}{\pi(z^2\langle p_T^2\rangle+\langle K_T^2\rangle)} .\end{aligned} \tag{3.59}$$

Collecting all the terms, we obtain the expression of the flavour independent $F_{UU,T}$ in the Gaussian ansatz framework:

$$F_{UU,T}=\left\{x\sum_a e_a^2 f_1^a(x)D_1^a(z)\right\}\frac{1}{\pi(z^2\langle p_T^2\rangle+\langle K_T^2\rangle)}\exp\left[-\frac{P_{h\perp}^2}{z^2\langle p_T^2\rangle+\langle K_T^2\rangle}\right] . \tag{3.60}$$

[11] $\int_{\mathbb{R}} dx\ e^{-ax^2}=\sqrt{\frac{\pi}{a}}\ , Re\ a>0$.

Chapter 4

Phenomenological exploration of the flavour dependence

In this chapter we will describe the *statistical methods* used to explore the flavour dependence of unpolarized TMDs from SIDIS data. However, there are two main *caveat* that we want to stress since the very beginning.

First of all, the COMPASS experiment did not deliver official data yet. Only preliminary ones have been discussed by the collaboration. As a consequence, all the phenomenological results obtained with our analysis should be *revised* when the official version of SIDIS data will be available (we expect them to be available in the next months). Moreover, the present work is founded on *Gaussian fits* of the preliminary data. More precise statements about the flavour dependence of unpolarized TMDs could be made once that official SIDIS data will be available from COMPASS, HERMES and Jefferson Lab's experiments.
Our goal is to *fit* SIDIS experimental data with flavour *dependent* multiplicities. Since we have available flavour *independent* Gaussian fits of experimental data, the first challenge is to reproduce good pseduo-experimental data from these fit functions and, then, fit these "fake" data with flavour dependent multiplicities (see fig. 4.1).

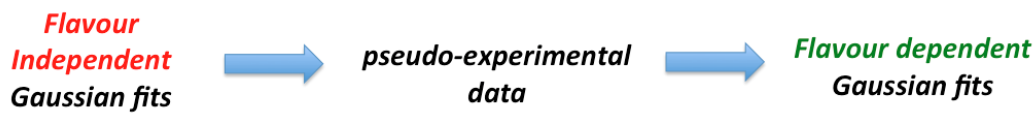

Figure 4.1: Steps involved in the analysis: first of all generate pseudo-experimental data from flavour independent fits of preliminary data; then, fit them with flavour dependent Gaussian multiplicities.

We will name the statistical technique used to generate pseudo-experimental data as the "method of random multiplicities". We will describe it in detail in the next section.

$\langle z^2_{h^+}\rangle$	$\langle P^2_{h^+\perp}\rangle$	A_+	$\chi^2_{h^+}/dof$
0.30	0.35 ± 0.010	1.27 ± 0.012	2.0

$\langle z^2_{h^-}\rangle$	$\langle P^2_{h^-\perp}\rangle$	A_-	$\chi^2_{h^-}/dof$
0.30	0.32 ± 0.013	0.82 ± 0.009	1.1

Table 4.1: The χ^2/dof, $\langle z^2\rangle$, best-fit values and errors for best-fit parameters, for positive and negative hadrons. Results are related to $0.07 < x_B < 0.12$ and $3.5 < Q^2 < 6$ GeV^2 bins.

4.1 Method of Random Multiplicities

The method of *random multiplicities* consists in generating a set of pseudo-experimental points adding *Gaussian noise* to the COMPASS fit function. Let's denote as $g_C(P^2_{h\perp}; x, z, Q^2)$ the Gaussian fit of the hadron multiplicity (see eq. 3.27):

$$g_C(P^2_{h\perp}; x, z, Q^2) = A_{+/-} \ \exp\left\{-\frac{P^2_{h\perp}}{\langle P^2_{h\perp}\rangle}\right\}, \tag{4.1}$$

where the best-fit parameters $A_{+/-}$ and $\langle P^2_{h\perp}\rangle$ are functions of (x, z, Q^2). For every kinematic bin (x, z, Q^2) best-fit values and statistical errors are given ([30]). By propagating the errors of the two parameters to the Gaussian fit function we can calculate the statistical error on the fit function itself and build 1σ-bands centered around g_C:

$$\Delta g = \sqrt{\left(\frac{\partial g_C}{\partial A_{+/-}}\right)^2 \left(\Delta A_{+/-}\right)^2 + \left(\frac{\partial g_C}{\partial \langle P^2_{h\perp}\rangle}\right)^2 \left(\Delta \langle P^2_{h\perp}\rangle\right)^2}, \tag{4.2}$$

where $\Delta A_{+/-}$ and $\Delta\langle P^2_{h\perp}\rangle$ are the statistical errors on the best-fit values. As an example, tab. 4.1 contains best-fit values and errors related to the kinematic bin 216 (see [30]). As we can deduce from eq. 4.2, the statistical error Δg is a function of the transverse momentum of the hadron. As a consequence, the width of 1σ-bands is a function of the transverse momentum $\boldsymbol{P}_{h\perp}$ (see fig. 4.2 and 4.3).

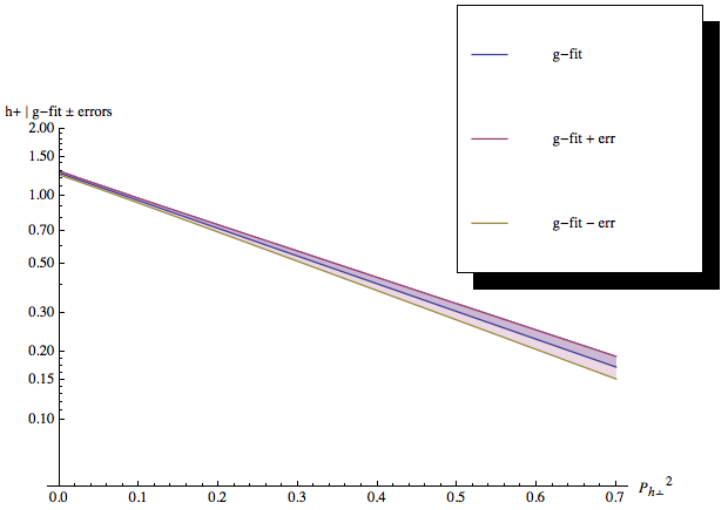

Figure 4.2: 1σ band centered around the fit function for bin 216 (see tab. 4.1 and [30]) for *positive* hadrons.

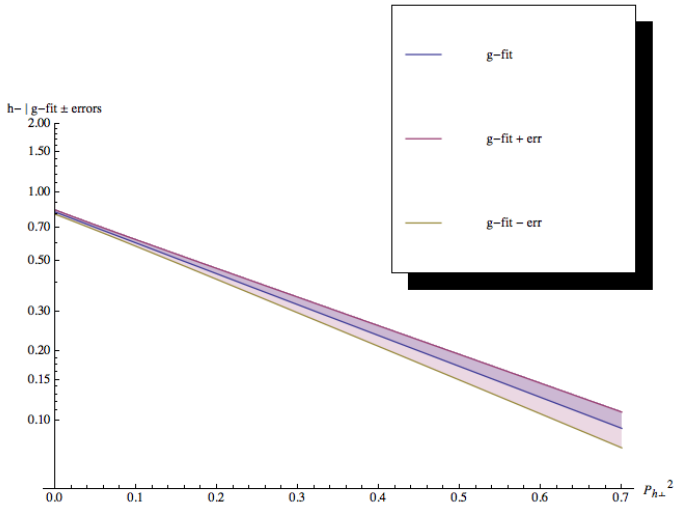

Figure 4.3: 1σ band centered around the fit function for bin 216 (see tab. 4.1 and [30]) for *negative* hadrons.

For every kinematic bin, let's select 30 values of $P_{h\perp}^2$ between 0 and 0.7 GeV^2: $\{P_{h\perp,i}^2\}_{i:1\to30}$ (this corresponds to the number of real data considered by the collaboration). The main step in the method consists in building sets of pseudo-experimental data adding numerical realizations of a random Gaussian variable to the fit function g_C evaluated in $P_{h\perp,i}^2$. Collecting all the kinematic variables in $\boldsymbol{y}$, we will build pseudo-experimental points as:

$$\left\{ P_{h\perp,i}^2;\ g_{exp}(P_{h\perp,i}^2;\ \boldsymbol{y}) \right\}_{i:1\to30} , \tag{4.3}$$

where g_{exp} is:

$$g_{exp}(P_{h\perp,i}^2;\ \boldsymbol{y}) = g_C(P_{h\perp,i}^2;\boldsymbol{y}) +\ \mathcal{N}(0,\mathcal{E}) \ . \tag{4.4}$$

$\mathcal{N}(0,\mathcal{E})$ is a random Gaussian variable with mean zero and standard deviation $\mathcal{E}$. The latter is defined as:

$$\mathcal{E}(\boldsymbol{y};P_{h\perp}^2) = \sqrt{b^2\Delta g_C(P_{h\perp}^2;\boldsymbol{y})^2 + \lambda^2 g_C(P_{h\perp}^2;\boldsymbol{y})^2} \ . \tag{4.5}$$

The first addend in the square root is the *statistical* component of the error, whereas the second one represents an estimate of the *systematic* error (the two components are added in quadrature). b is a multiplicative factor, used to increase the statistical error. λ represents the percentage of g_C to be considered as systematic error. For example, $\lambda = 0.10$ means that the systematic error related to the point $\{P_{h\perp,i}^2; g_{exp}(P_{h\perp,i}^2;\boldsymbol{y})\}$ is $0.10 g_C(P_{h\perp,i}^2;\boldsymbol{y})$. The b factor and the systematic component are needed because if we calculate the error on pseudo-data using error propagation on the fit function only, we would underestimate it. The COMPASS collaboration did not quote systematic errors in the preliminary data, but, presumably, there should be many sources of such errors.
In order to take into account all these effects, we replicated the analysis choosing b between 2 and 3 and introducing a systematic error between the 10% and 20% (different combinations have been investigated).
With our method, we add to all $g_C(P_{h\perp,i}^2;\boldsymbol{y})$ values a Gaussian number spread around the fit function with mean zero and dispersion $\mathcal{E}$. As a consequence, the set of generated pseudo-experimental data looks like a set of points whose best-fit curve is g_C (see fig. 4.4 and 4.5) [1].

The *fidelity* of the generated points with respect to the Gaussian fit function is self-evident. Since we add noise with mean zero, the average behaviour of the points is exactly the same as the Gaussian fit function. The problem

[1] As explained in the next section, fig. 4.4 and 4.5 contain no systematic component of the error: $\lambda = 0$.

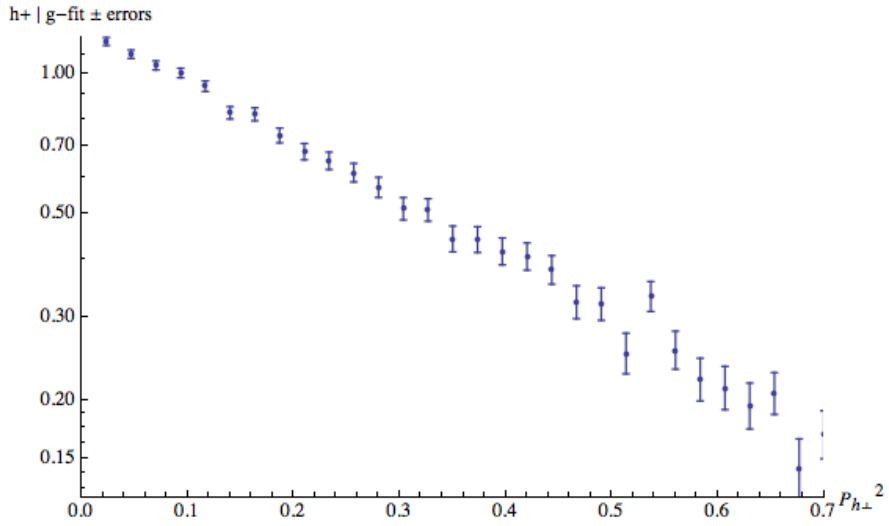

Figure 4.4: A generation of pseudo-experimental points with the kinematics of bin 216 (see tab. 4.1 and [30]), for *positive* hadrons.

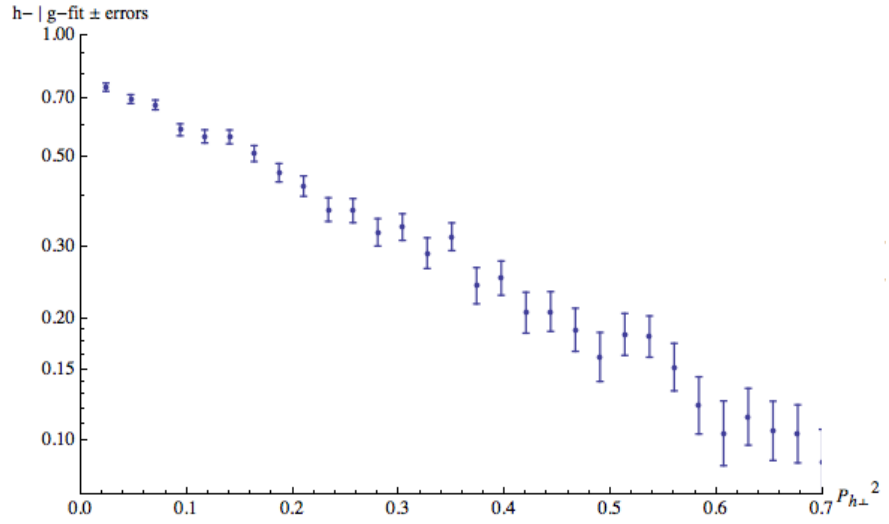

Figure 4.5: A generation of pseudo-experimental points with the kinematics of bin 216 (see tab. 4.1 and [30]), for *negative* hadrons.

is that the agreement between Gaussian fits of COMPASS preliminary data and data themselves is not perfect: the related χ^2/dof are always greater than 1 [30]. As a consequence, we are using sets of data that are quite far from the physical ones.
In order to generate pseudo-data not too far from the preliminary ones, we

will disregard the kinematic bins in the COMPASS analysis with *high* values of χ^2/dof, i.e. > 4. This happens for the first two Q^2 bins ($1 < Q^2 < 1.5$ GeV^2 and $1.2 < Q^2 < 1.7$ GeV^2), the first two x bins in the third Q^2 bin ($1.5 < Q^2 < 2.5$ GeV^2, $0.008 < x < 0.025$) and all the kinematic bins with $\langle z^2 \rangle = 0.05$ (see fig. 3.4).

4.1.1 Chi-square function

The fit of pseudo-data with the flavour dependent multiplicity is a *combined* one. It is performed by minimizing the following chi-square function

$$\chi^2 = \sum_{Q^2,x,\langle z^2 \rangle} \sum_{h=\pi^+,\pi^-} \sum_{i:0}^{29} \frac{\left[g_{exp}(P^2_{h\perp,i}; x, z, Q^2) - m_h(x, z, Q^2; \boldsymbol{v}; P^2_{h\perp,i}) \right]^2}{\mathcal{E}(x, z, Q^2; P^2_{h\perp,i})^2} \ , \tag{4.6}$$

with respect to $\boldsymbol{v}$, the set of best-fit parameters. The summation is performed over the COMPASS kinematic bins (see fig. 3.4) and over the detected hadrons, assumed to be pions. The last summation is over the selected values of $P^2_{h\perp}$, equally separated between 0 and 0.7 GeV^2:

$$P^2_{h\perp,i} \doteq \frac{0.7}{30} \cdot i \ , \ \{i : 0 \to 29\} \ . \tag{4.7}$$

The peculiarity of our analysis consists in *replicating* all the fitting procedure a fixed number $\mathcal{M}$ of times, with $\mathcal{M}$ different sets of pseudo-data (all generated according to eq. 4.4), obtaining $\mathcal{M}$ different sets of best-fit values [2]. In this way, we can build *histograms* of best-fit values for each parameter, evaluating their mean values and statistical errors. In the *frequentist* limit $\mathcal{M} \to \infty$ histograms approach *probability density distributions*.
For each set of best-fit values we can plot the related best-fit multiplicities $m_h(\boldsymbol{y}, \boldsymbol{v}; P^2_{h\perp,i})$, obtaining a *band* of best-fit models. Plotting one set of pseudo-experimental data (with 1σ error bands) vs the *band of models* we can qualitatively evaluate the agreement between COMPASS preliminary data and *flavour dependent* hadron multiplicity based on a Gaussian ansatz. Moreover, the agreement is quantitatively evaluated through the χ^2/dof, where χ^2 is defined in eq. 4.6 and the number of degrees of freedom is:

$$dof = 2 \cdot 30 \cdot N_z \cdot (N_x(q_0) + \cdots + N_x(q_5)) - N_{par} \ . \tag{4.8}$$

[2] Pseudo-data and related best-fit values will be uncorrelated since $\mathcal{N}$ is a random variable.

Here, 2 accounts for the summation over detected π^+ and π^-; 30 is the number of selected $P^2_{h\perp,i}$, N_z is the number of z bins, $N_x(q_i)$ is the number of x bins related to the ith Q^2 bin and N_{par} is the number of parameters in the fit.
Replicating the fitting procedure $\mathcal{M}$ times we can also build the histogram of χ^2/dof values. In our simulations we obtained histograms resembling the probability density distribution of chi-square random variables (see fig. 4.7). Since good agreement between data and model corresponds to $\chi^2/dof \sim 1$, we decided to cut the right tail of the histogram, namely to exclude replicas with $\chi^2/dof > 5$. This selection allows to build histograms for best-fit values of parameters including only good replicas of pseudo-data.

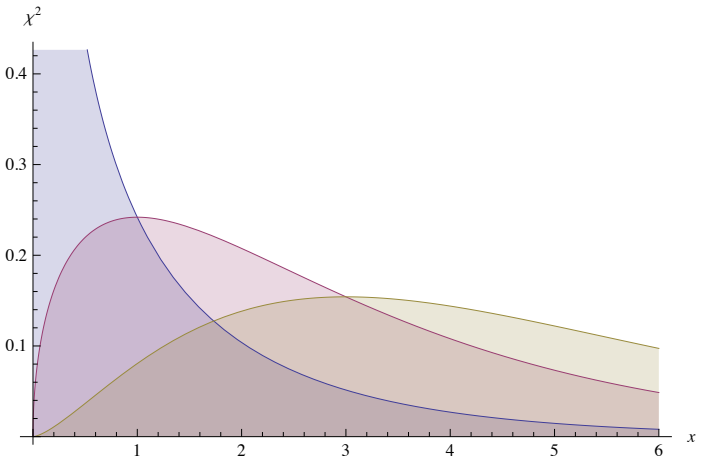

Figure 4.6: The probability density distribution for χ^2 random variables with 1 (purple), 3 (yellow) and 5 (blue) degrees of freedom.

The minimization algorithm is written using the C++ programming language, with ROOT interface to MINUIT. MSTW2008 parton distribution functions for the proton (available in LHAPDF [31] package) and DSS [32] parameterizations of fragmentation functions for pions are used. One of the 100 MINUIT output of the minimization algorithm is shown:

```
************************************
*       LHAPDF Version 5.8.7       *
*   Configured for the following:  *
*             All PDFs             *
*          LOW MEMORY option       *
*    Maximum  3 concurrent set(s)  *
************************************

>>>>>> PDF description: <<<<<<
MSTW 2008 LO (68% C.L.)
```

```
Reference:
 A.D. Martin, W.J. Stirling, R.S. Thorne and G. Watt
 "Parton distributions for the LHC"
 arXiv:0901.0002 [hep-ph]
 http://projects.hepforge.org/mstwpdf/
This set has 41 member PDFs.
 mem=0 --> central value
 mem=1-40 --> 20 eigenvector sets (+/- directions)
 See Section 6 of paper for error calculation.
>>>>>>                  <<<<<<

**********
**  499 **MIGRAD        5000           1
**********
FIRST CALL TO USER FUNCTION AT NEW START POINT, WITH IFLAG=4.
START MIGRAD MINIMIZATION.  STRATEGY  1.  CONVERGENCE WHEN EDM .LT. 1.00e-03
FCN=1.27441e+11 FROM MIGRAD    STATUS=INITIATE      359 CALLS      495241 TOTAL
                    EDM= unknown      STRATEGY= 1      NO ERROR MATRIX
 EXT PARAMETER               CURRENT GUESS       STEP         FIRST
 NO.   NAME      VALUE            ERROR          SIZE      DERIVATIVE
  1  n           1.35629e+00   1.35629e-01   0.00000e+00   1.75632e+11
  2  N_u         1.49995e+01   4.90477e-01   0.00000e+00   1.38693e+08
  3  a_u         6.00539e-04   5.71342e-01   0.00000e+00   5.79065e+07
  4  d_u         3.01498e+00   6.62711e-01   0.00000e+00   6.53610e+11
  5  N_d         1.50000e+01   1.05471e-01   0.00000e+00   2.01467e+02
  6  a_d         2.27041e+00   2.27041e-01   0.00000e+00   3.90533e+04
  7  d_d         1.86571e-02   7.57158e-01   0.00000e+00   5.31142e+05
  8  N_s         9.38777e+00   9.38777e-01   0.00000e+00  -4.39802e+10
  9  a_s         2.73211e+00   2.73211e-01   0.00000e+00   9.32673e+09
 10  d_s         3.28354e+00   6.35309e-01  -3.63127e-01   1.66521e+12
 11  N_fv        5.81315e+00   5.81315e-01   0.00000e+00  -1.20507e+11
 12  b_fv        8.43612e+00   8.43612e-01   0.00000e+00   4.28104e+11
 13  c_fv        7.96659e+00   7.96659e-01   0.00000e+00   4.19950e+11
 14  N_un        2.86374e+00   2.86374e-01   0.00000e+00  -4.87979e+10
 15  b_un        7.09612e+00   7.09612e-01   0.00000e+00   1.26283e+11
 16  c_un        7.21693e+00   7.21693e-01   0.00000e+00   9.11094e+10
MIGRAD MINIMIZATION HAS CONVERGED.
MIGRAD WILL VERIFY CONVERGENCE AND ERROR MATRIX.
```

4.1.2 Illustrative partial analysis

In this paragraph we will presents results related to the analysis of a subset of all the available kinematic bins, used as a template for more complete studies. We will investigate six z bins, with fixed x and Q^2 values [30]:

$$3.5 < Q^2 < 6 \text{ GeV}^2 \ , \ \ 0.07 < x < 0.12 \ , \qquad 0.11 < \langle z^2 \rangle < 0.56 \ . \tag{4.9}$$

The parameters involved in the fit are:

- $\boldsymbol{n}$: the overall normalization
- $\boldsymbol{u}$: $\langle p_{T,up}^2 \rangle$, the variance of the Gaussian TMD PDF for up (valence) quarks

- $\boldsymbol{d}$: $\langle p^2_{T,down}\rangle$, the variance of the Gaussian TMD PDF for down (valence) quarks
- $\boldsymbol{s}$: $\langle p^2_{T,sea}\rangle \equiv 0.2$ GeV^2. We chose to *fix* the variance of the Gaussian TMD PDF for sea quarks .
- $\boldsymbol{N_1}$: the normalization coefficient of $\langle K^2_{T,fav}\rangle$
- $\boldsymbol{\alpha}$: the exponent of the z-dependent term in $\langle K^2_{T,fav}\rangle$
- $\boldsymbol{\beta}$: the exponent of the $(1-z)$-dependent term in $\langle K^2_{T,fav}\rangle$.
- $\boldsymbol{N_2}$: the normalization coefficient of $\langle K^2_{T,unf}\rangle$
- $\boldsymbol{\gamma}$: the exponent of the z-dependent term in $\langle K^2_{T,unf}\rangle$
- $\boldsymbol{\delta}$: the exponent of the $(1-z)$-dependent term in $\langle K^2_{T,unf}\rangle$.

We do not use the complete parametrization for $\langle p^2_T\rangle$. Since the values of x and Q^2 are fixed, we will use only one parameter. The expression of the variance for the favoured fragmentation process is:

$$\langle K^2_{T,fav}\rangle = N_1 z^\alpha (1-z)^\beta \ . \tag{4.10}$$

The normalization N_1 involves the logarithmic term from TMD evolution (which is however constant, since Q^2 is fixed).
As in the previous case, the expression of the unfavoured variance is:

$$\langle K^2_{T,unf}\rangle = N_2 z^\gamma (1-z)^\delta \ , \tag{4.11}$$

In this analysis the number of degrees of freedom is (see tab. 3.4 and [30]:

$$dof = 2 \cdot 30 \cdot 6 \ - 9 = 351 \ . \tag{4.12}$$

For this *partial* analysis we set $\lambda = 0$ and $b = 2$ in eq.(4.5), avoiding the introduction of systematic components in the definition of the error $\mathcal{E}(\boldsymbol{y}; P^2_{h\perp})$.

$$\mathcal{E}(\boldsymbol{y}; P^2_{h\perp}) = 2 \cdot \Delta g_C(\boldsymbol{y}; P^2_{h\perp}) \ . \tag{4.13}$$

Nevertheless, the fit converges to physical minima of the χ^2 function, since we have a reasonable number of parameters (9) for fitting 360 data. On the contrary, the introduction of systematic components will be mandatory when performing the *global* analysis: in that case there will be 16 parameters for $\sim 5 \cdot 10^3$ data points and other error components must be introduced beyond

Table 4.2: Mean values and statistical errors for best-fit parameters over the $\mathcal{M} = 100$ replicas (partial analysis).

	n	$\langle p^2_{T,up}\rangle$	$\langle p^2_{T,down}\rangle$	N_1	α	β	N_2	γ	δ
mean	1.430	**0.467**	**0.062**	0.948	1.088	0.902	3.798	1.883	1.516
σ	0.005	**0.040**	**0.034**	0.414	0.190	0.214	2.274	0.426	0.445

the statistical one.
Let's now show results related to the minimization of χ^2 for $\mathcal{M} = 100$ replicas. For each set of pseudo-data we obtain a set of best-fit values for the nine parameters, whose *mean* values and *statistical error* are grouped in tab. 4.2: The histogram of χ^2/dof values calculated over all the 100 replicas resembles the proper reduced chi-square profile: In the next figures (4.8, 4.9, 4.10

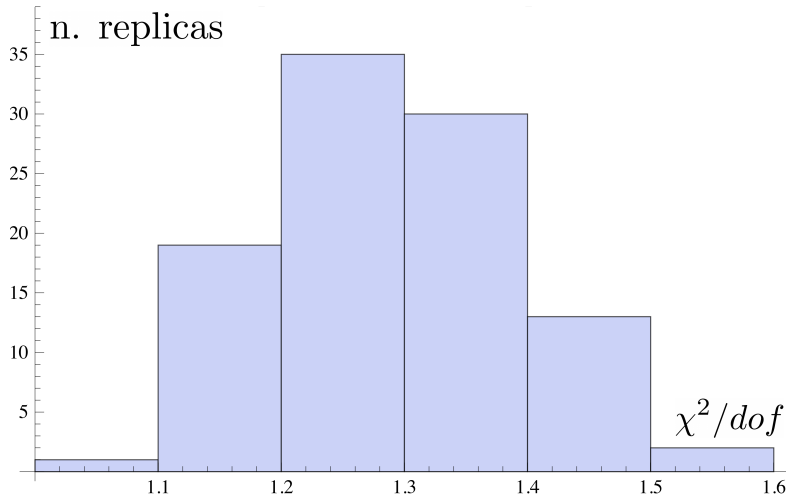

Figure 4.7: Histogram of χ^2/dof values: 100 replicas, partial analysis.

and 4.11) histograms related to $\langle p^2_{T,up}\rangle$, $\langle p^2_{T,down}\rangle$, $\langle K^2_{T,fav}\rangle$ and $\langle K^2_{T,unf}\rangle$ are, respectively, shown. This phenomenological study suggests that the TMD Gaussian related to the up quark is *wider* than the down quark one. Also, the two mean square transverse momenta are statistically incompatible. Mean values of favoured and unfavoured mean squared transverse momenta, instead, are similar (with the unfavoured one slightly bigger) but statistically

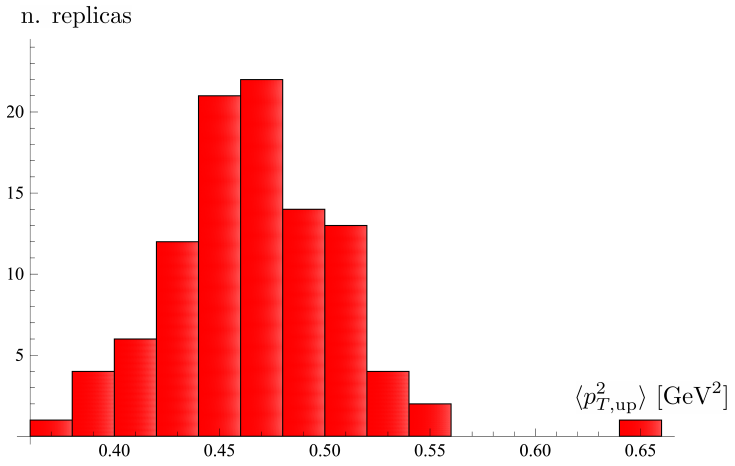

Figure 4.8: Histogram of $\langle p^2_{T,up} \rangle$ values: 100 replicas, partial analysis.

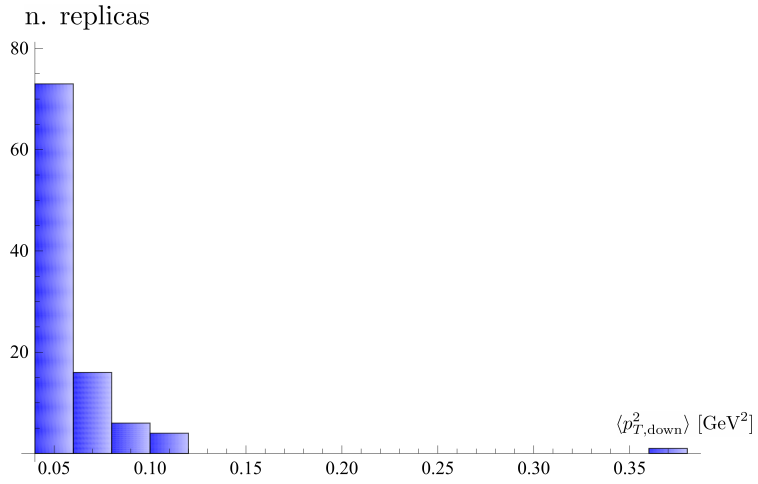

Figure 4.9: Histogram of $\langle p^2_{T,down} \rangle$ values: 100 replicas, partial analysis.

incompatible, as in the previous case (see tab. 4.3).
This behaviour (up wider than down, unfavored wider than favored) is confirmed also in the global analysis, but average values will turn out to be

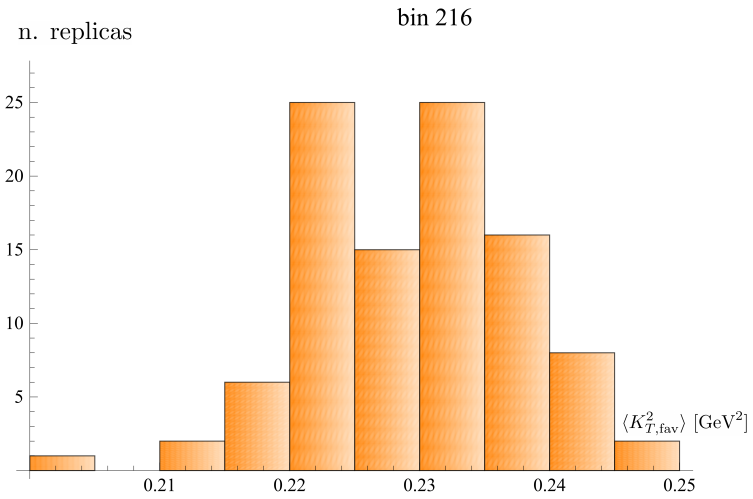

Figure 4.10: Histogram of $\langle K^2_{T,fav}\rangle$ values for $\langle z^2\rangle = 0.3$: 100 replicas, partial analysis.

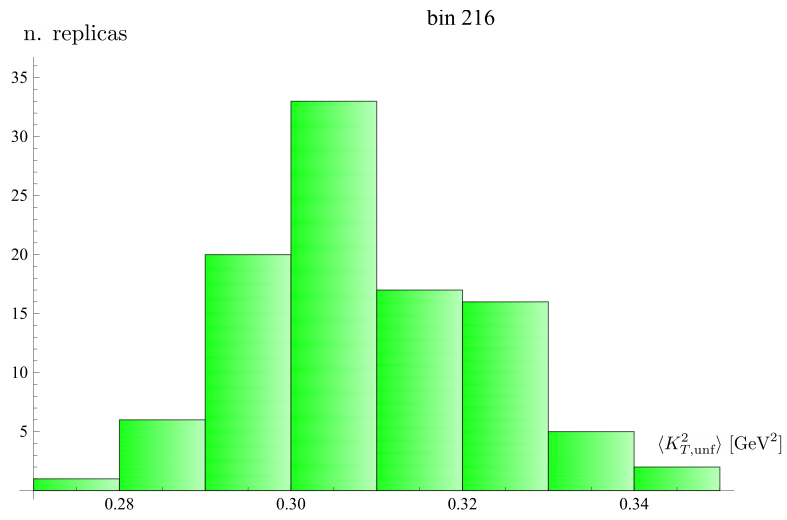

Figure 4.11: Histogram of $\langle K^2_{T,unf}\rangle$ values for $\langle z^2\rangle = 0.3$: 100 replicas, partial analysis.

statistically compatible. The correlations between $\langle p^2_{T,up}\rangle$ and $\langle p^2_{T,down}\rangle$ and

Table 4.3: Mean values and statistical errors for up, down, favored and unfavored variances over the $\mathcal{M} = 100$ replicas (partial analysis).

	$\langle p^2_{T,up} \rangle$	$\langle p^2_{T,down} \rangle$	$\langle K^2_{T,fav} \rangle$	$\langle K^2_{T,unf} \rangle$
Mean val. [GeV^2]	0.46	0.06	0.229	0.307
stat. error [GeV^2]	0.04	0.03	0.008	0.014

between $\langle K^2_{T,fav} \rangle$ and $\langle K^2_{T,unf} \rangle$ are tabulated in tab. 4.4: PDFs variances are correlated (if one increases also the other one increases), FFs variances, instead, are anti-correlated (if one increases, the other one decreases). The

Table 4.4: Correlation indeces for PDFs variances and FFs variances (partial analysis).

	$\rho(\ \langle p^2_{T,up} \rangle, \langle p^2_{T,down} \rangle\ )$	$\rho(\ \langle K^2_{T,fav} \rangle, \langle K^2_{T,unf} \rangle\ )$
corr. index	0.55	-0.34

agreement between our model and data is shown in figs. 4.12 - 4.17, for bins 216, 217, 218 [30]. Since the agreement is very good, we are plotting data without Gaussian noise (even if the fit is performed adding the rumor).
Red bands are related to π^+ multiplicities and green ones are related to π^- multiplicities. Our result are almost always inside the 1σ error band.

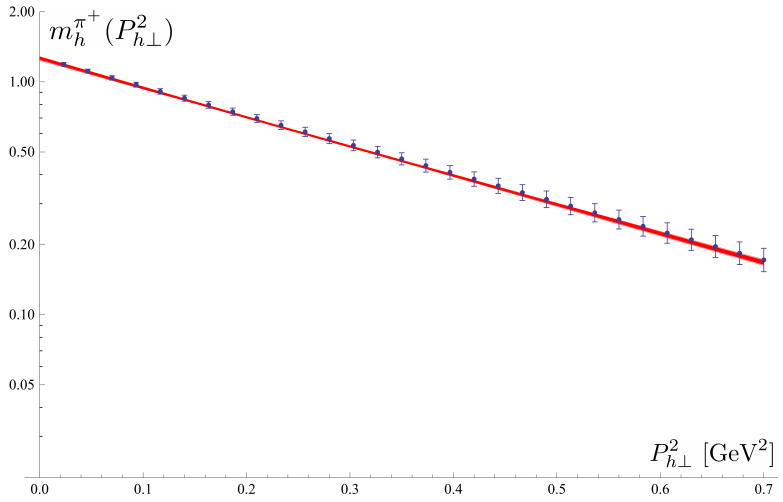

Figure 4.12: Flavour dependent fit of preliminary COMPASS data ($\sigma^{(4)}_{SIDIS}/\sigma^{(2)}_{DIS}$ vs $P^2_{h\perp,i}$) for π^+ in kinematic bin 216.

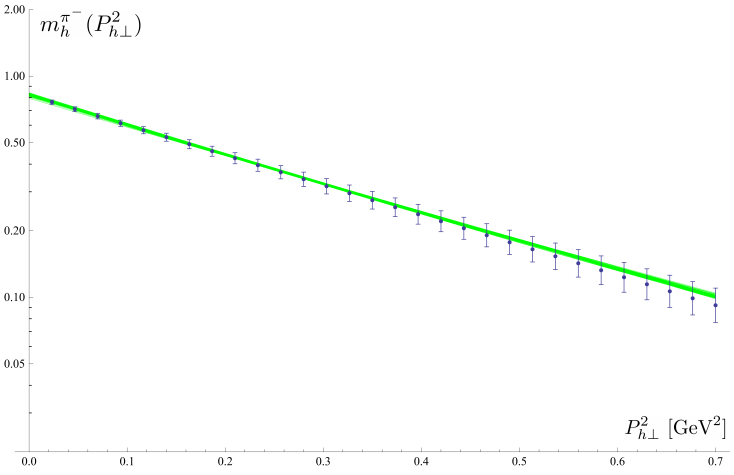

Figure 4.13: Flavour dependent fit of preliminary COMPASS data ($\sigma^{(4)}_{SIDIS}/\sigma^{(2)}_{DIS}$ vs $P^2_{h\perp,i}$) for π^- in kinematic bin 216.

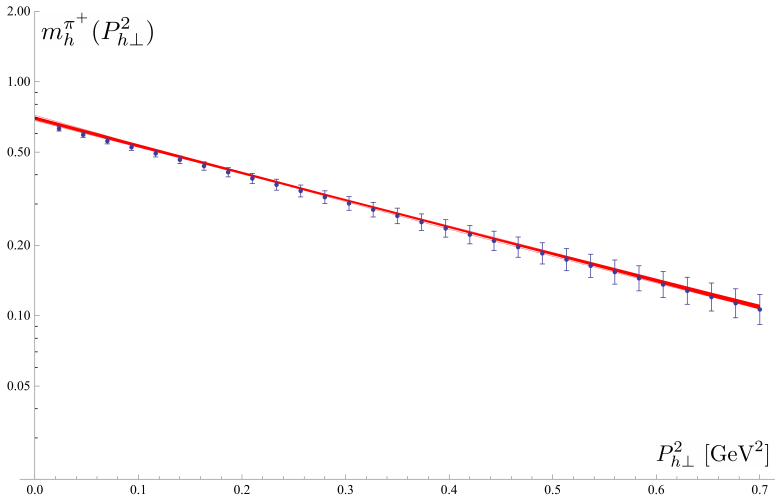

Figure 4.14: Flavour dependent fit of preliminary COMPASS data ($\sigma^{(4)}_{SIDIS}/\sigma^{(2)}_{DIS}$ vs $P^2_{h\perp,i}$) for π^+ in kinematic bin 217.

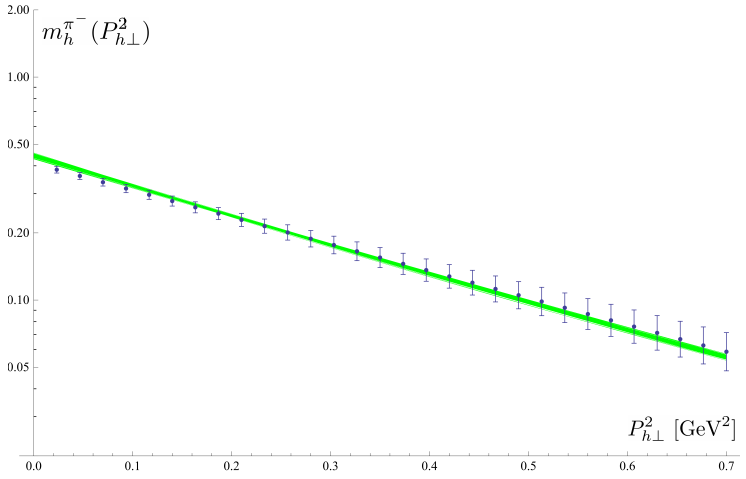

Figure 4.15: Flavour dependent fit of preliminary COMPASS data ($\sigma^{(4)}_{SIDIS}/\sigma^{(2)}_{DIS}$ vs $P^2_{h\perp,i}$) for π^- in kinematic bin 217.

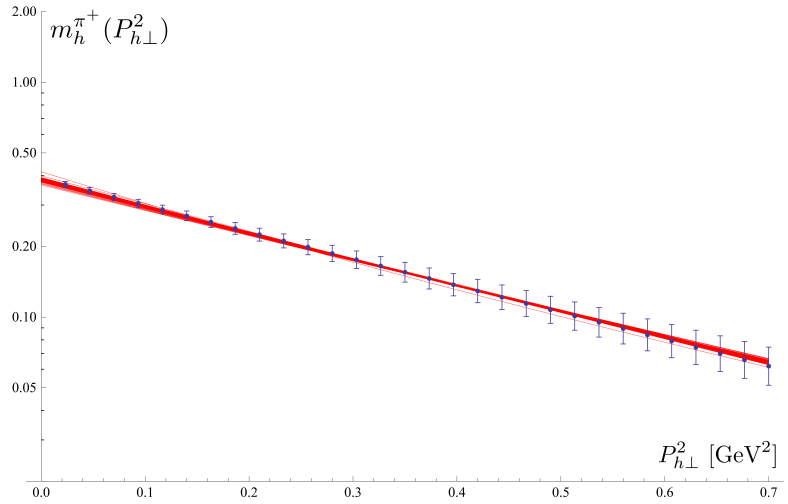

Figure 4.16: Flavour dependent fit of preliminary COMPASS data ($\sigma^{(4)}_{SIDIS}/\sigma^{(2)}_{DIS}$ vs $P^2_{h\perp,i}$) for π^+ in kinematic bin 218.

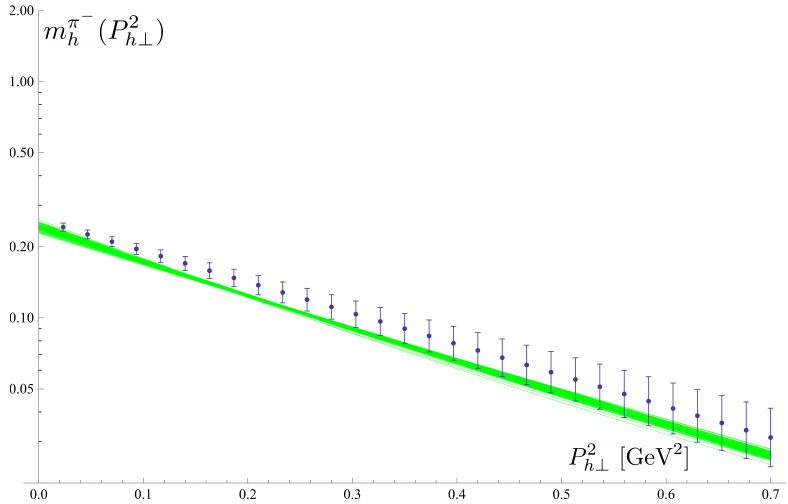

Figure 4.17: Flavour dependent fit of preliminary COMPASS data ($\sigma^{(4)}_{SIDIS}/\sigma^{(2)}_{DIS}$ vs $P^2_{h\perp,i}$) for π^- in kinematic bin 218.

4.1.3 Global analysis

Let's now turn to a *global* analysis. After describing the differences with the partial study of the previous paragraph, we will present results on both flavor dependent and independent tests.

The main differences with the partial analysis consist of the larger number of data considered, the introduction of systematic errors and the application of a selection criterion on COMPASS Gaussian fits: in order to reasonably describe the preliminary SIDIS data, we will disregard the kinematic bins in the COMPASS analysis with $\chi^2/dof > 4$ (where the χ^2 is related to the agreement between preliminary data and their Gaussian fits). As a consequence, the first two Q^2 bins ($1 < Q^2 < 1.5$ GeV^2 and $1.2 < Q^2 < 1.7$ GeV^2), the first two x bins in the third Q^2 bin ($1.5 < Q^2 < 2.5$ GeV^2, $0.008 < x < 0.025$) and all the kinematic bins with $\langle z^2 \rangle = 0.05$ (see fig. 3.4) will not be included.

Since now x and Q^2 are available, we need the complete form of the variances:

$$\begin{aligned} \langle p_{T,a}^2 \rangle (x, Q^2) &= N_a x^{\alpha_a} (1-x)^{\beta_a} \left(1 + \ln \frac{Q^2}{Q_0^2} \right) \\ \langle K_{T,b}^2 \rangle (z, Q^2) &= N_b z^{\alpha_b} (1-z)^{\beta_b} \left(1 + \ln \frac{Q^2}{Q_0^2} \right) . \end{aligned} \tag{4.14}$$

The scale Q_0^2 in the TMD evolution term is fixed [3] to 1 GeV^2. From eq.(4.14) we deduce that the parameters involved in the fit are:

- $\boldsymbol{n}$: the overall normalization
- $\boldsymbol{N_u}$: the normalization factor in $\langle p_{T,up}^2 \rangle$, the variance of the Gaussian TMD PDF for up quarks (valence)
- $\boldsymbol{\alpha_u}$: the exponent in the x-dependent term in $\langle p_{T,up}^2 \rangle$
- $\boldsymbol{\beta_u}$: the exponent in the $(1-x)$-dependent term in $\langle p_{T,up}^2 \rangle$
- $\boldsymbol{N_d}$: the normalization factor in $\langle p_{T,down}^2 \rangle$, the variance of the Gaussian TMD PDF for down quarks (valence)
- $\boldsymbol{\alpha_d}$: the exponent in the x-dependent term in $\langle p_{T,down}^2 \rangle$
- $\boldsymbol{\beta_d}$: the exponent in the $(1-x)$-dependent term in $\langle p_{T,down}^2 \rangle$

[3] Results are not very sensitive on Q_0^2. Lower values for this parameters are also allowed.

- $\boldsymbol{N_s}$: the normalization factor in $\langle p^2_{T,sea}\rangle$, the variance of the Gaussian TMD PDF for sea quarks
- $\boldsymbol{\alpha_s}$: the exponent in the x-dependent term in $\langle p^2_{T,sea}\rangle$
- $\boldsymbol{\beta_s}$: the exponent in the $(1-x)$-dependent term in $\langle p^2_{T,sea}\rangle$
- $\boldsymbol{N_{fav}}$: the normalization factor in $\langle K^2_{T,fav}\rangle$, the variance of the Gaussian TMD FF for favored processes
- $\boldsymbol{\alpha_{fav}}$: the exponent in the z-dependent term in $\langle K^2_{T,fav}\rangle$
- $\boldsymbol{\beta_{fav}}$: the exponent in the $(1-z)$-dependent term in $\langle K^2_{T,fav}\rangle$
- $\boldsymbol{N_{unf}}$: the normalization factor in $\langle K^2_{T,unf}\rangle$, the variance of the Gaussian TMD FF for unfavored processes
- $\boldsymbol{\alpha_{unf}}$: the exponent in the z-dependent term in $\langle K^2_{T,unf}\rangle$
- $\boldsymbol{\beta_{unf}}$: the exponent in the $(1-z)$-dependent term in $\langle K^2_{T,unf}\rangle$

The number of degrees of freedom is:

$$dof = 2 \cdot 30 \cdot 7 \cdot (2+4+4+2) - 16 = 5024\ . \tag{4.15}$$

As in the partial analysis, the factor 2 is related to π^+ and π^- detection, 30 is the number of $P^2_{h\perp}$ values considered, 7 is the number of z bins and the numbers in parentheses are related to the quantity of x bins for every Q^2 bin considered.

4.2 Test of flavour dependence

In this section we will show results obtained by setting $b = 2$ and $\lambda = 0.1$ in eq.(4.5).

$$b = 2 \qquad\qquad \lambda = 0.1 \ . \tag{4.16}$$

Namely, we are doubling the statistic error on Gaussian fits and we are introducing a systematic error of 10% of the value of the hadron multiplicity.

The agreement between the flavour dependent model and pseudo-data generated from Gaussian fits is acceptable for the present explorative study: the mean value of the χ^2/dof is 2.3. However, this value strictly depends on the choice of b and λ. As a consequence, it is an important indicator about the quality of the fit, but it is not "absolute". Moreover, we can compare only results obtained with the same choice for b and λ (e.g., this same study, replicated with $\lambda = 0.15$, is characterized by a mean χ^2/dof around 1.5).

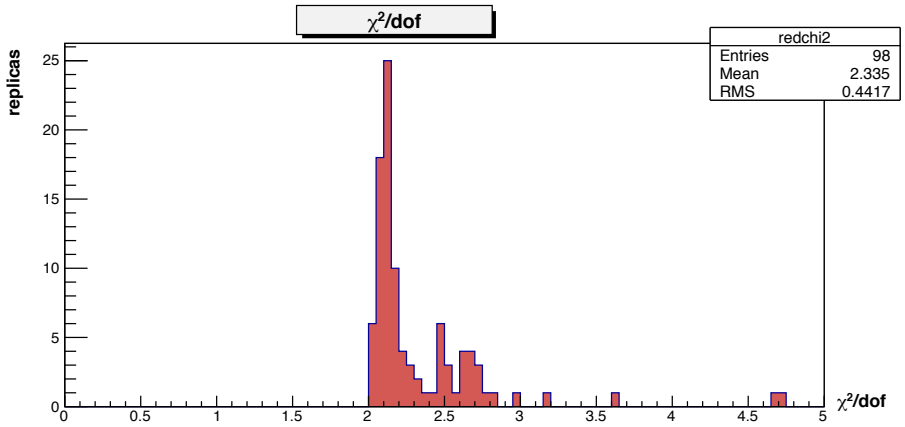

Figure 4.18: Histogram collecting the values of the random variable χ^2/dof ($b = 2$, $\lambda = 0.1$). 2 of the 100 replicas were rejected by the selection criterion implemented in the code ($\chi^2/dof < 5$).

The overall normalization factor is described by a distribution peaked around its mean value (see fig. 4.20). The mean value and the statistical error of the overall normalization are:

$$n = 1.35 \pm 0.05 \ , \tag{4.17}$$

similar to values obtained in the previous section.

From our analysis, a global tendency for the average mean square transverse momentum of the up quark to be *larger* than the down one emerges over all the considered kinematic intervals in x and Q^2. Moreover, sea quarks are characterized by Gaussian distributions with average mean transverse momenta larger than or similar to the up ones. This could be explained observing that all the considered x values are $\lesssim 0.1$, a region where sea quarks give a significant contribution. In the final section "Plots", histograms related to $\langle p_T^2 \rangle$ and $\langle K_T^2 \rangle$ values are presented for different kinematic regions. Values of the transverse momentum of the struck parton for different x intervals, with $3.5 < Q^2 < 6$ GeV^2, are shown in figs. from (4.21) to (4.29). In figs. from (4.30) to (4.38) we present the transverse momentum of the struck parton for different Q^2 values, with $0.04 < x < 0.07$ (the "mean" interval for the x variable). Results related to mean values and standard deviations are collected in tab. (4.5) and tab. (4.6).

	up [GeV^2]	down [GeV^2]	sea [GeV^2]
$1.5 < Q^2 < 2.5$ GeV^2	0.55 ± 0.53	0.54 ± 0.98	0.67 ± 0.54
$2.5 < Q^2 < 3.5$ GeV^2	0.60 ± 0.53	0.56 ± 0.99	0.69 ± 0.58
$3.5 < Q^2 < 6$ GeV^2	0.73 ± 0.65	0.64 ± 1.12	0.84 ± 0.71

Table 4.5: Mean values and standard deviations for $\langle p_T^2 \rangle (x, Q^2)$ for different Q^2 values, with $0.04 < x < 0.07$.

x_B	up [GeV^2]	down [GeV^2]	sea [GeV^2]
$0.018 - 0.025$	1.09 ± 1.28	0.32 ± 1.04	1.46 ± 1.09
$0.025 - 0.04$	0.94 ± 0.99	0.81 ± 1.59	1.21 ± 0.90
$0.07 - 0.12$	0.45 ± 0.29	0.36 ± 0.50	0.40 ± 0.38

Table 4.6: Mean values and standard deviations for $\langle p_T^2 \rangle (x, Q^2)$ for different x values, with $3.5 < Q^2 < 6$.

Values of $\langle p_T^2 \rangle$ in the histograms sometimes accumulate towards zero. This, however, is unphysical: a lower bound on $\langle p_T^2 \rangle$ values should be imposed (for example $\langle p_T \rangle > 50$ MeV). Unfortunately, it is not easy to implement this requirement into the code. One strategy could be rejecting replicas of the fit procedure with non-physical $\langle p_T^2 \rangle$ values, but it would be very time-consuming from the computational point of view.

From the analysis of $\langle K_T^2 \rangle$ we can notice that the TMD part of the unfavored fragmentation function is wider than the favored one for *low* values

of z^2: this happens in all the Q^2 bins. For *high* values of z^2, instead, the behaviour is opposite (see figs. from (4.39) to (4.44)).

$\langle z^2 \rangle$	$\langle K^2_{T,fav} \rangle$ [GeV^2]	$\langle K^2_{T,unf} \rangle$ [GeV^2]
0.08	0.19 ± 0.08	0.21 ± 0.05
0.14	0.24 ± 0.07	0.23 ± 0.10
0.56	0.20 ± 0.15	0.14 ± 0.15

Table 4.7: Mean values and standard deviations for $\langle K^2_T \rangle (z, Q^2)$ for different z values, with $3.5 < Q^2 < 6$.

In figs. from (4.51) to (4.58) fits of COMPASS pseudo-data with flavour dependent hadron multiplicities are shown. All the curves calculated as best-fit models in each replica of the minimization algorithm are represented. On the contrary, only one set of pseudo-data is plotted. As a consequence, the visual agreement between data and the band of models is "data-set dependent". However the choice of the data-set is not crucial, since all the replicas of data are generated with the same criterion (the Gaussian noise). The qualitative agreement is good and compatible with a $\chi^2/dof \sim 2.3$. Since this is a *combined* fit (over 84 kinematic bins replicated both for π^+ and for π^-), there are very good plots related to bins with low χ^2 and other where the agreement is worse, due to a higher χ^2 value. The mean "degree" of agreement ($\chi^2/dof \sim 2.3$) is built taking into account all the analyzed bins. All the plots are represented using logarithmic scale for the multiplicity m_h. Some best-fit curves are not straight lines in these plots: this is because the structure function $F_{UU,T}$ inside the multiplicity is no more a Gaussian function, but a *sum* of Gaussian functions.

In the next table we show correlation indeces among mean square transverse momenta calculated with $2.5 < Q^2 < 3.5$ GeV^2, $0.025 < x < 0.04$ and $\langle z^2 \rangle = 0.14$:

ρ	$\langle p^2_{T,up} \rangle$	$\langle p^2_{T,down} \rangle$	$\langle p^2_{T,sea} \rangle$	$\langle K^2_{T,fav} \rangle$	$\langle K^2_{T,unf} \rangle$
$\langle p^2_{T,up} \rangle$	1	0.15	-0.36	-0.03	0.43
$\langle p^2_{T,down} \rangle$	0.15	1	-0.20	-0.05	0.06
$\langle p^2_{T,sea} \rangle$	-0.36	-0.20	1	-0.05	-0.71
$\langle K^2_{T,fav} \rangle$	-0.03	-0.05	-0.05	1	0.02
$\langle K^2_{T,unf} \rangle$	0.43	0.06	-0.71	0.02	1

Up and down mean square transverse momenta are correlated, with correlation index 0.15. Favored and unfavored transverse momenta have a very low

correlation index (0.02): this means that when $\langle K_{T,fav} \rangle$ increases, $\langle K_{T,unf} \rangle$ is left almost unchanged (and vice-versa).

4.3 Test of flavour independence

In this section we will show results related to fitting COMPASS pseudo-data with flavour *independent* hadron multiplicities. A worse degree of agreement in comparison to the flavour dependent analysis would be the indication of the need to use flavour dependent unpolarized TMDs.

The flavour independent fit is performed with the same values of b and λ (2 and 0.10) used for the flavour dependence test. The number of replicas is 50. We used the same physical structure for $\langle p_T^2 \rangle$ and $\langle K_T^2 \rangle$ as in the flavour dependent case, with:

$$\begin{aligned} &N_u = N_d = N_{sea} \\ &\alpha_u = \alpha_d = \alpha_{sea} \\ &\beta_u = \beta_d = \beta_{sea} \\ \\ &N_{fav} = N_{unf} \\ &\alpha_{fav} = \alpha_{unf} \\ &\beta_{fav} = \beta_{unf} \ . \end{aligned} \tag{4.18}$$

As a consequence, the mean square transverse momentum is equal for up, down and sea quarks and for favored and unfavored fragmentation processes. Moreover, the number of parameters is reduced to 7 (taking into account also the overall normalization).

From the histogram of χ^2/dof values we can immediately see that the agreement between pseudo-data and the model is *worse* than in the flavour dependent case. The mean value of χ^2/dof is 3 (see fig. 4.19), higher than in the flavour dependent case.

Figs. from (4.59) to (4.62) show the agreement between flavour independent models and one set of pseudo-data. The general behaviour over all the kinematic bins investigated is *worse* than in the flavour dependent fit.

Some of the values calculated for $\langle p_T^2 \rangle$ and $\langle K_T^2 \rangle$ in different bins are shown in the following tables:

$\langle p^2_{T,u=d=sea} \rangle$	$0.018 < x < 0.025$	$0.025 - 0.04$	$0.04 - 0.07$	$0.07 - 0.12$
$3.5 < Q^2 < 6$	1.47 ± 0.26	1.36 ± 0.26	1.12 ± 0.23	0.76 ± 0.15

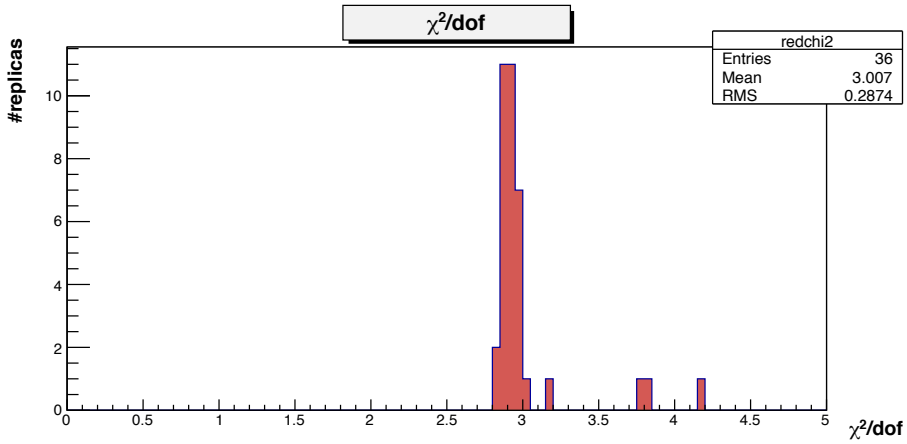

Figure 4.19: Histogram collecting the values of the random variable χ^2/dof. 14 of the 50 replicas were rejected by the selection criterion implemented in the code ($\chi^2/dof < 5$).

$\langle K^2_{T,fav=unf}\rangle$[GeV^2]	$\langle z^2\rangle = 0.08$	$\langle z^2\rangle = 0.11$	$\langle z^2\rangle = 0.14$	$\langle z^2\rangle = 0.20$
$3.5 < Q^2 < 6$[GeV^2]	0.21 ± 0.01	0.16 ± 0.03	0.13 ± 0.04	0.09 ± 0.05

4.4 Plots

4.4.1 Flavour dependence test

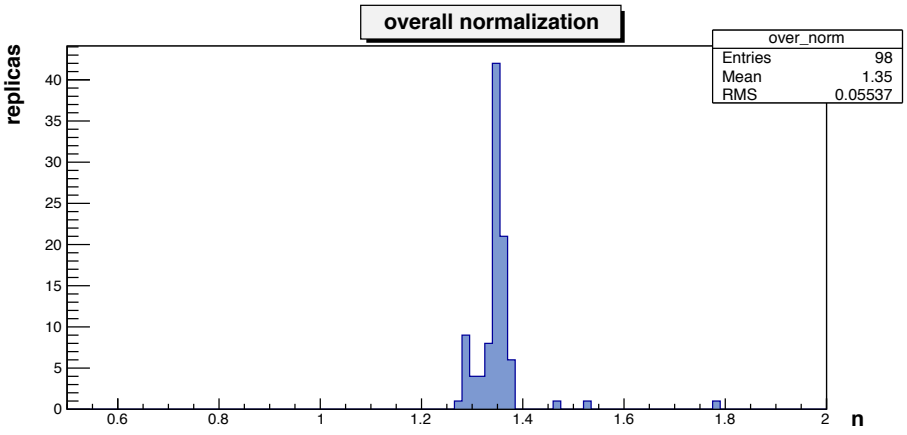

Figure 4.20: Histogram collecting the values of the overall normalization in the hadron multiplicity m_h.

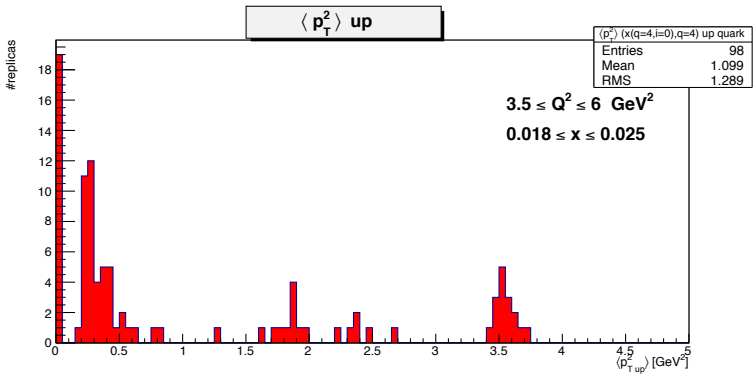

Figure 4.21: Histogram collecting the values of $\langle p^2_{T,up}\rangle(x, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $0.018 < x < 0.025$.

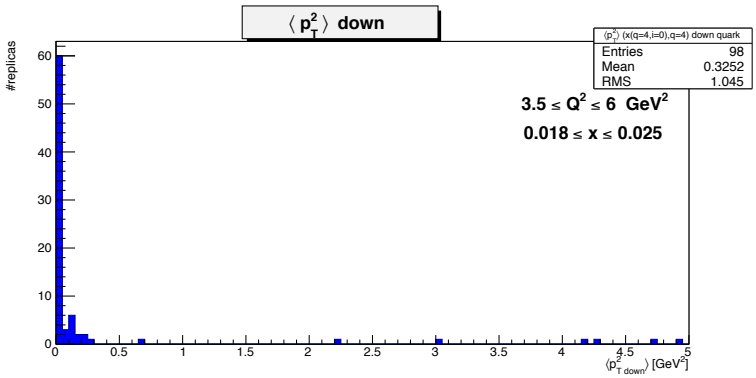

Figure 4.22: Histogram collecting the values of $\langle p^2_{T,down}\rangle(x, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $0.018 < x < 0.025$.

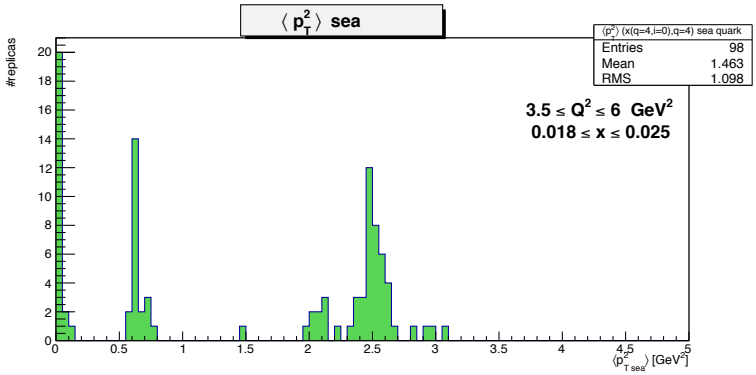

Figure 4.23: Histogram collecting the values of $\langle p^2_{T,sea}\rangle(x, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $0.018 < x < 0.025$.

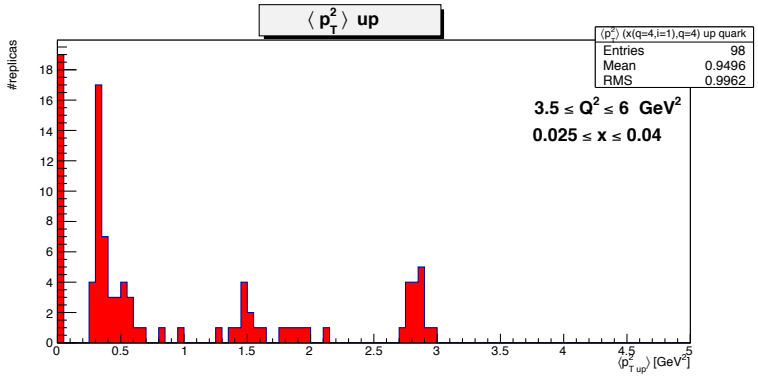

Figure 4.24: Histogram collecting the values of $\langle p^2_{T,up}\rangle(x, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $0.025 < x < 0.04$.

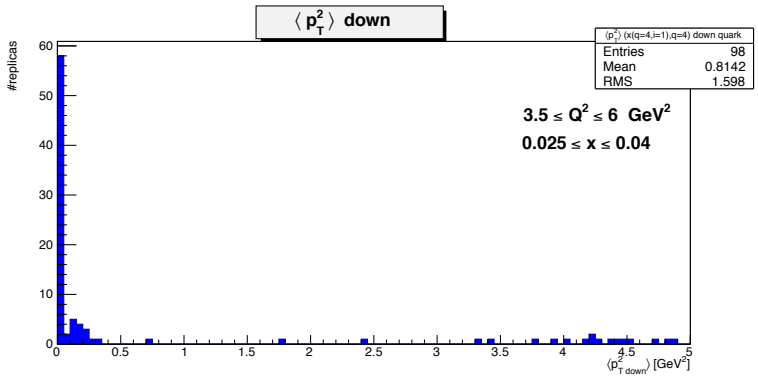

Figure 4.25: Histogram collecting the values of $\langle p^2_{T,down}\rangle(x, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $0.025 < x < 0.04$.

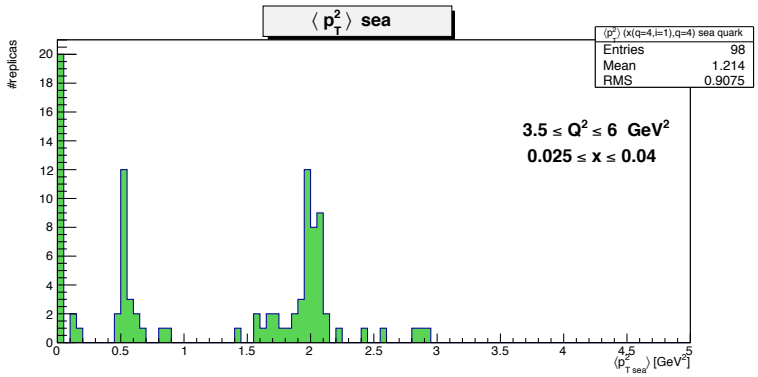

Figure 4.26: Histogram collecting the values of $\langle p^2_{T,sea}\rangle(x, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $0.025 < x < 0.04$.

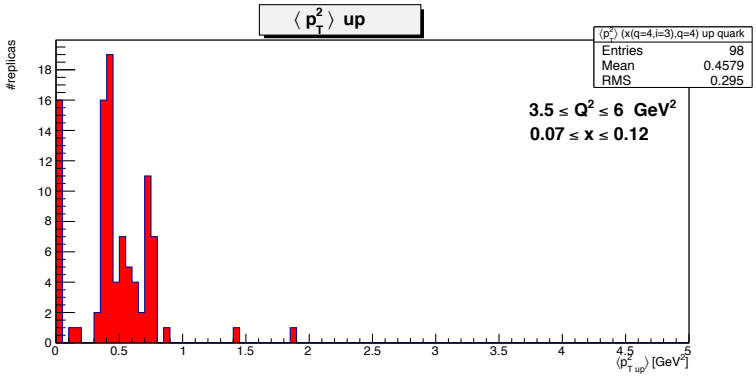

Figure 4.27: Histogram collecting the values of $\langle p^2_{T,up} \rangle (x, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $0.07 < x < 0.12$.

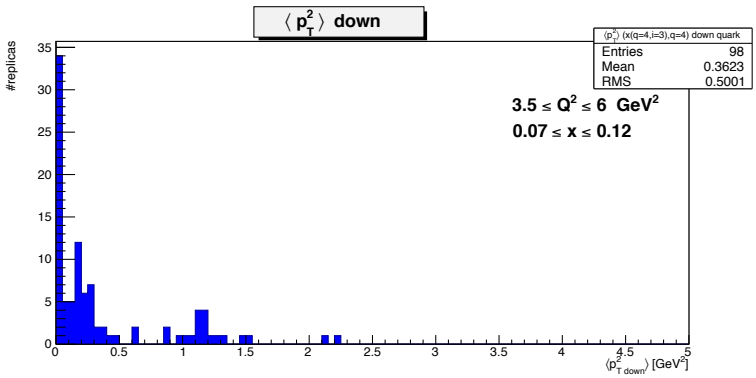

Figure 4.28: Histogram collecting the values of $\langle p^2_{T,down} \rangle (x, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $0.07 < x < 0.12$.

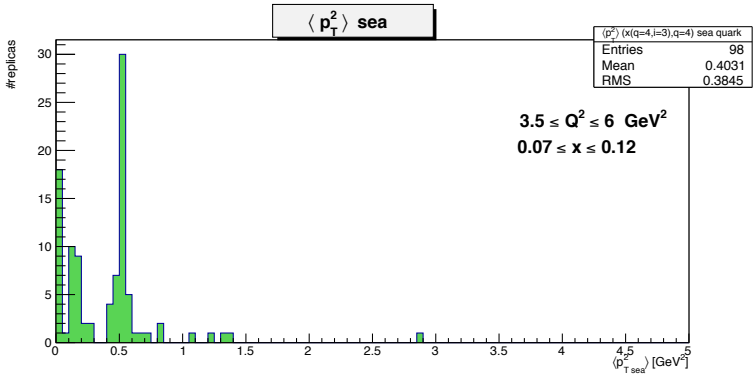

Figure 4.29: Histogram collecting the values of $\langle p^2_{T,sea} \rangle (x, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $0.07 < x < 0.12$.

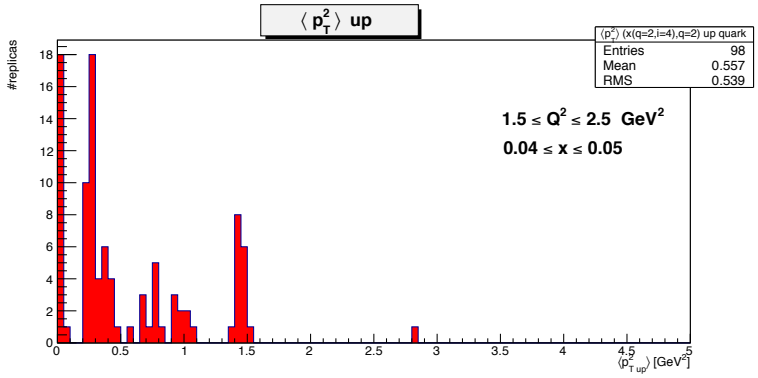

Figure 4.30: Histogram collecting the values of $\langle p^2_{T,up}\rangle(x,Q^2)$, with $1.5 < Q^2 < 2.5$ GeV^2 and $0.04 < x < 0.05$.

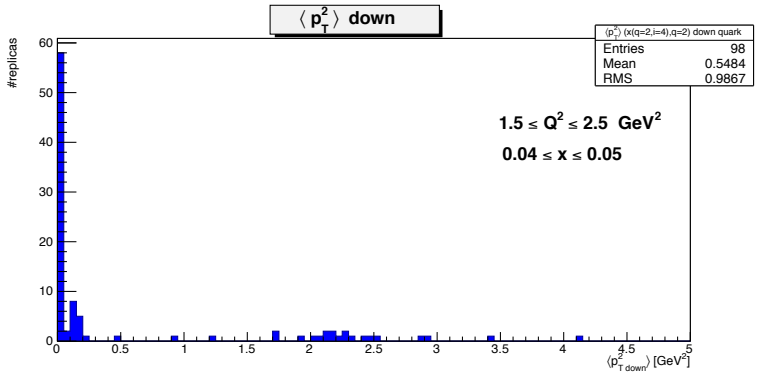

Figure 4.31: Histogram collecting the values of $\langle p^2_{T,down}\rangle(x,Q^2)$, with $1.5 < Q^2 < 2.5$ GeV^2 and $0.04 < x < 0.05$.

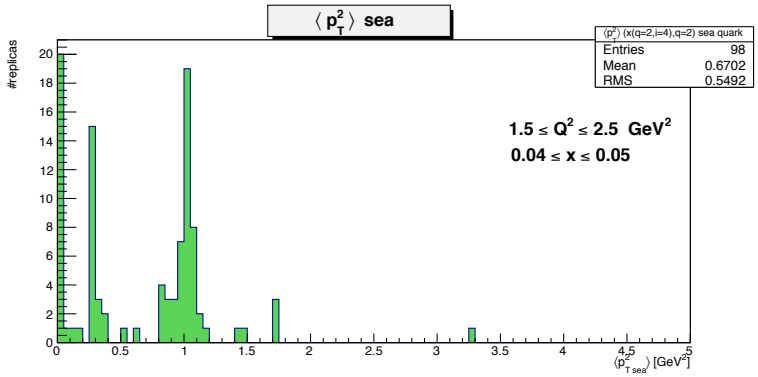

Figure 4.32: Histogram collecting the values of $\langle p^2_{T,sea}\rangle(x,Q^2)$, with $1.5 < Q^2 < 2.5$ GeV^2 and $0.04 < x < 0.05$.

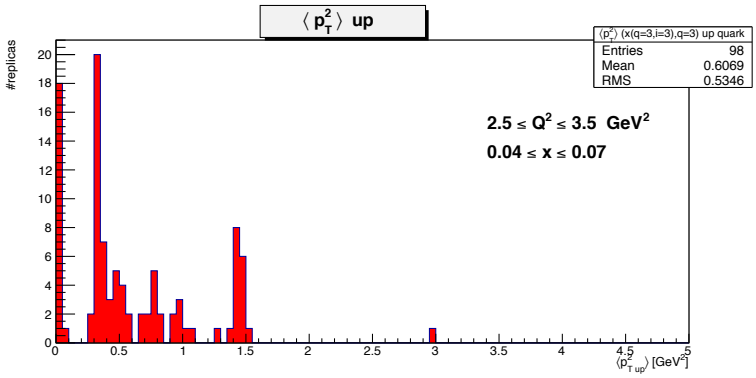

Figure 4.33: Histogram collecting the values of $\langle p^2_{T,up}\rangle(x, Q^2)$, with $2.5 < Q^2 < 3.5$ GeV^2 and $0.04 < x < 0.07$.

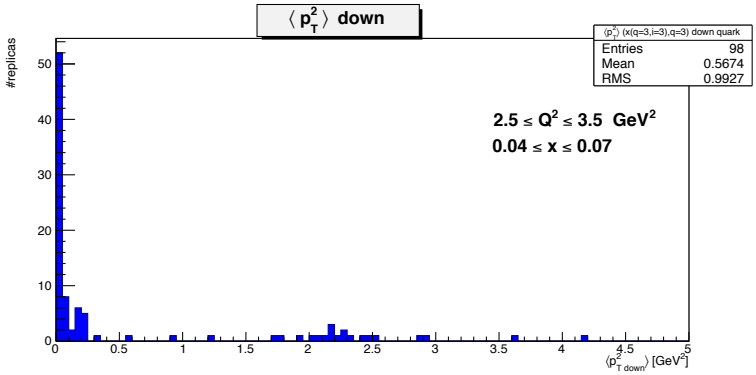

Figure 4.34: Histogram collecting the values of $\langle p^2_{T,down}\rangle(x, Q^2)$, with $2.5 < Q^2 < 3.5$ GeV^2 and $0.04 < x < 0.07$.

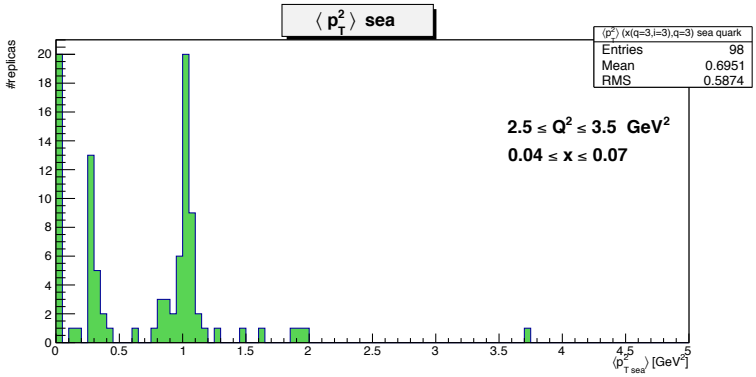

Figure 4.35: Histogram collecting the values of $\langle p^2_{T,sea}\rangle(x, Q^2)$, with $2.5 < Q^2 < 3.5$ GeV^2 and $0.04 < x < 0.07$.

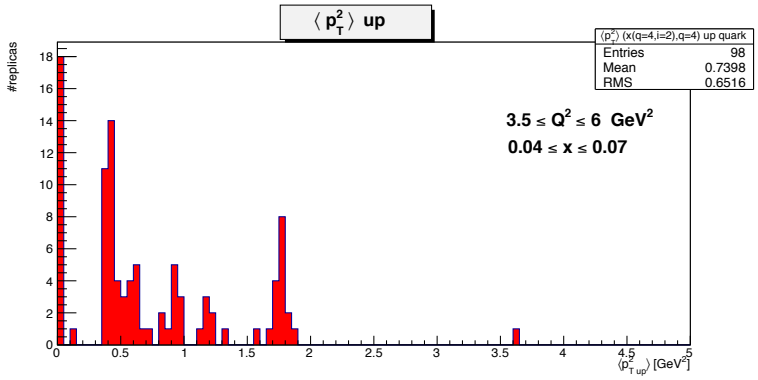

Figure 4.36: Histogram collecting the values of $\langle p^2_{T,up}\rangle(x, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $0.04 < x < 0.07$.

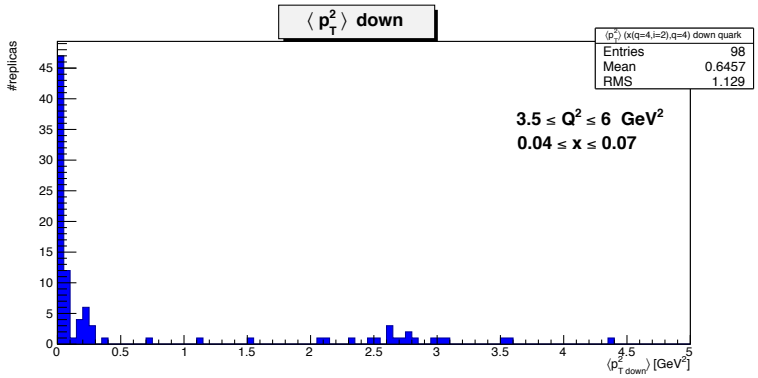

Figure 4.37: Histogram collecting the values of $\langle p^2_{T,down}\rangle(x, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $0.04 < x < 0.07$.

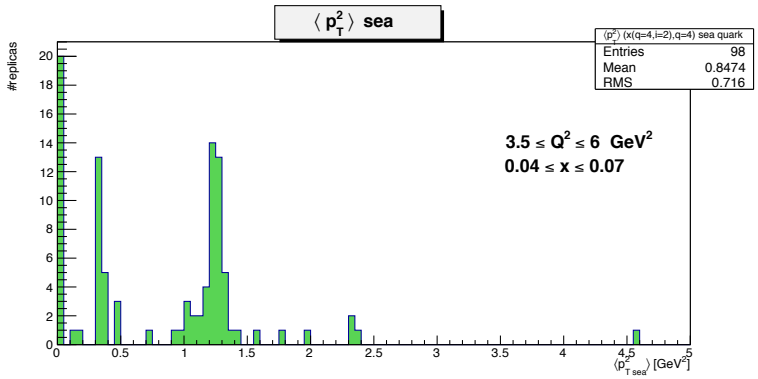

Figure 4.38: Histogram collecting the values of $\langle p^2_{T,sea}\rangle(x, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $0.04 < x < 0.07$.

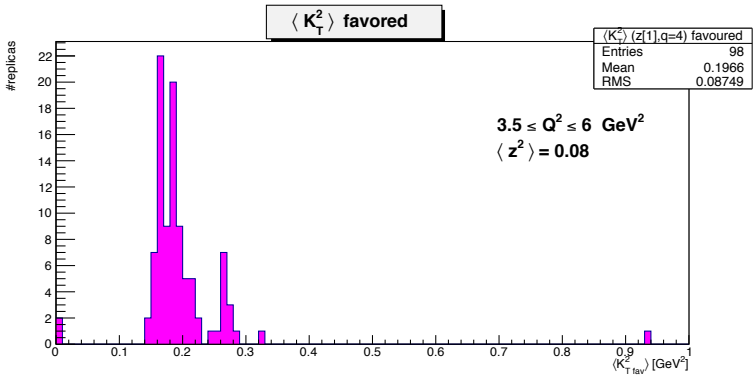

Figure 4.39: Histogram collecting the values of $\langle K^2_{T,fav}\rangle(z, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $\langle z^2\rangle = 0.08$.

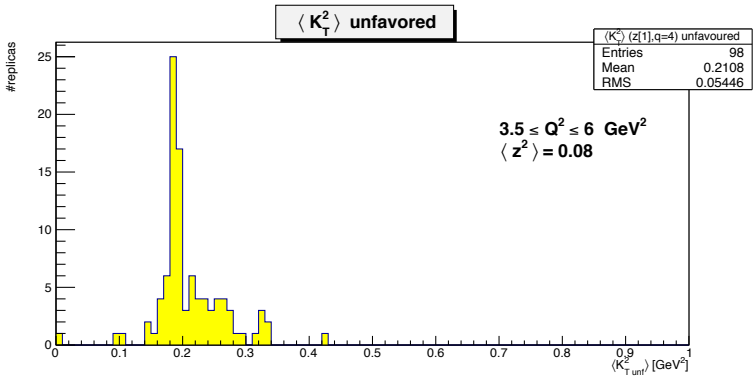

Figure 4.40: Histogram collecting the values of $\langle K^2_{T,unf}\rangle(z, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $\langle z^2\rangle = 0.08$.

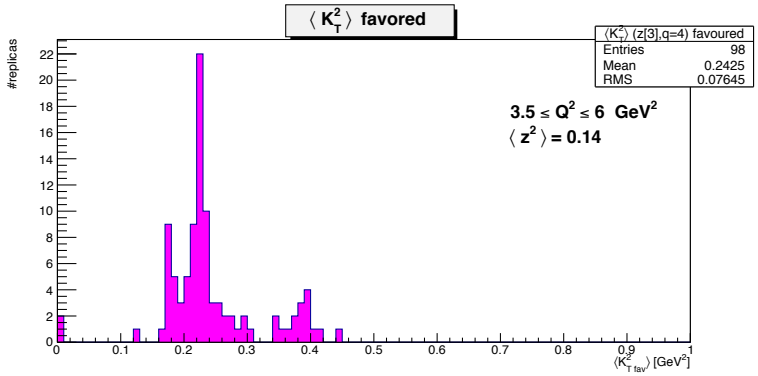

Figure 4.41: Histogram collecting the values of $\langle K^2_{T,fav}\rangle(z, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $\langle z^2\rangle = 0.14$.

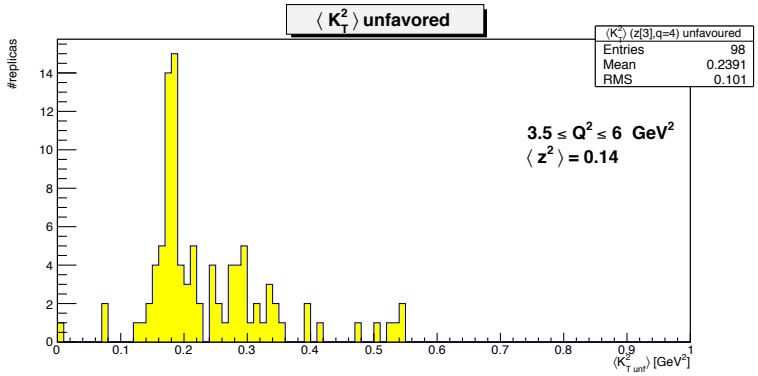

Figure 4.42: Histogram collecting the values of $\langle K^2_{T,unf}\rangle(z, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $\langle z^2\rangle = 0.14$.

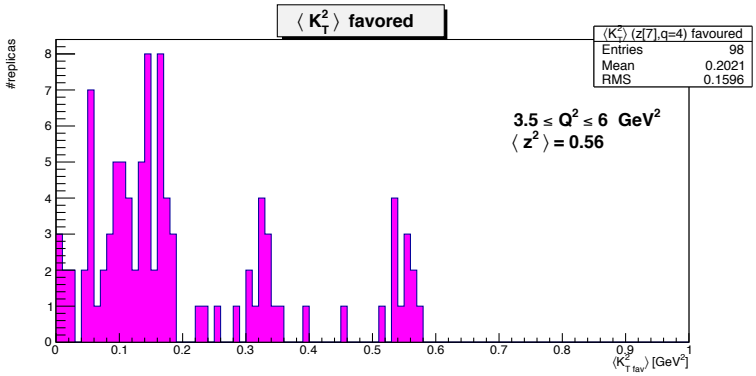

Figure 4.43: Histogram collecting the values of $\langle K^2_{T,fav}\rangle(z, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $\langle z^2\rangle = 0.56$.

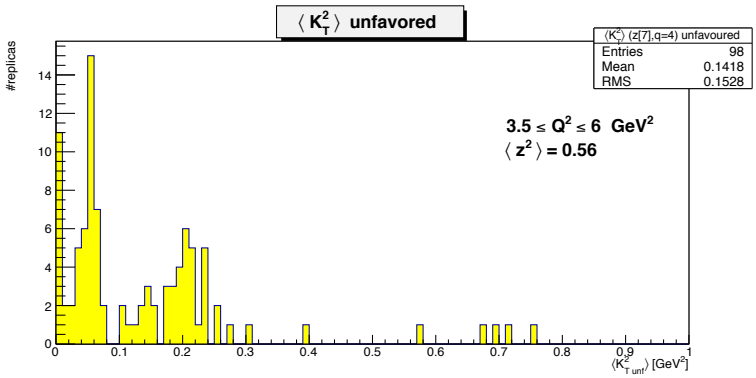

Figure 4.44: Histogram collecting the values of $\langle K^2_{T,unf}\rangle(z, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $\langle z^2\rangle = 0.56$.

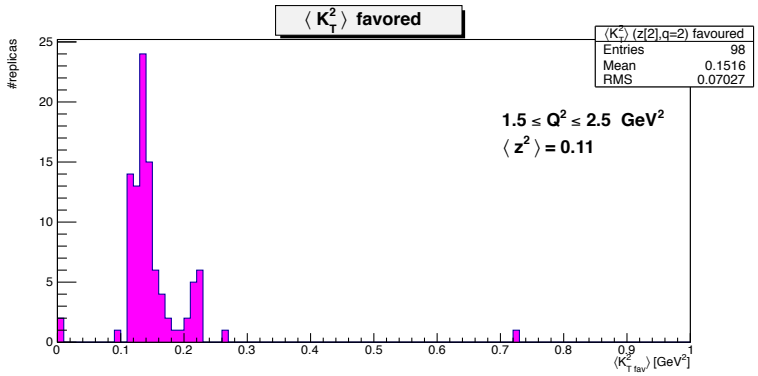

Figure 4.45: Histogram collecting the values of $\langle K^2_{T,fav}\rangle(z, Q^2)$, with $1.5 < Q^2 < 2.5$ GeV^2 and $\langle z^2\rangle = 0.11$.

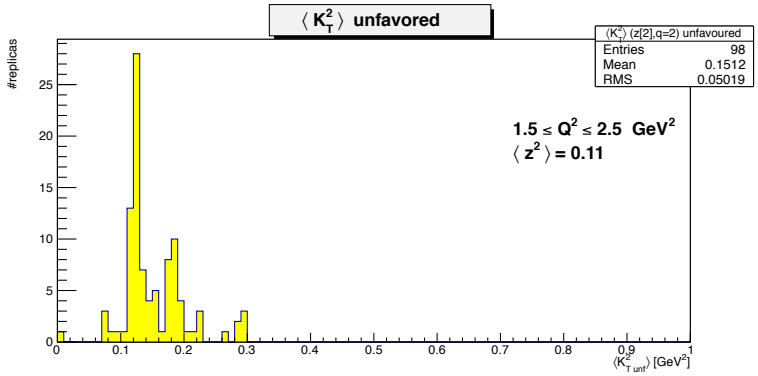

Figure 4.46: Histogram collecting the values of $\langle K^2_{T,unf}\rangle(z, Q^2)$, with $1.5 < Q^2 < 2.5$ GeV^2 and $\langle z^2\rangle = 0.11$.

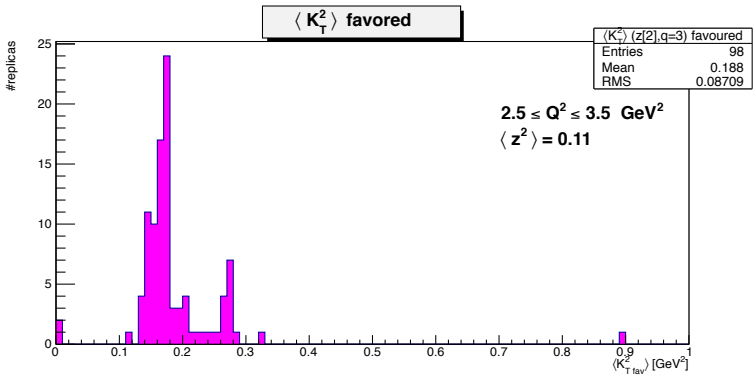

Figure 4.47: Histogram collecting the values of $\langle K^2_{T,fav}\rangle(z, Q^2)$, with $2.5 < Q^2 < 3.5$ GeV^2 and $\langle z^2\rangle = 0.11$.

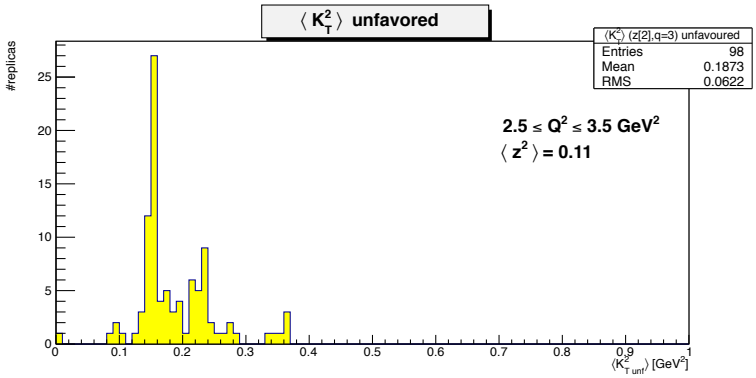

Figure 4.48: Histogram collecting the values of $\langle K^2_{T,unf}\rangle(z, Q^2)$, with $2.5 < Q^2 < 3.5$ GeV^2 and $\langle z^2\rangle = 0.11$.

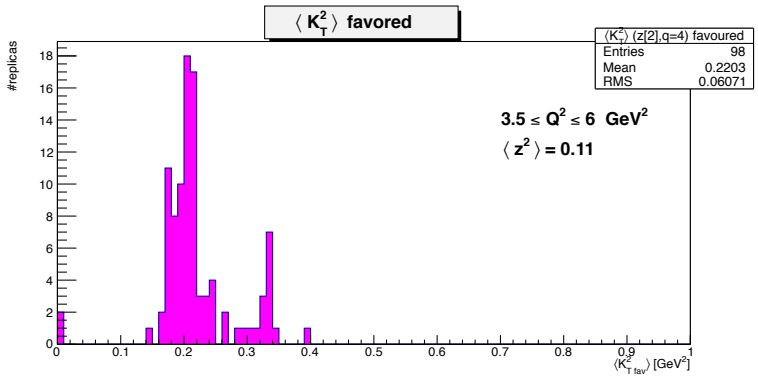

Figure 4.49: Histogram collecting the values of $\langle K^2_{T,fav}\rangle(z, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $\langle z^2\rangle = 0.11$.

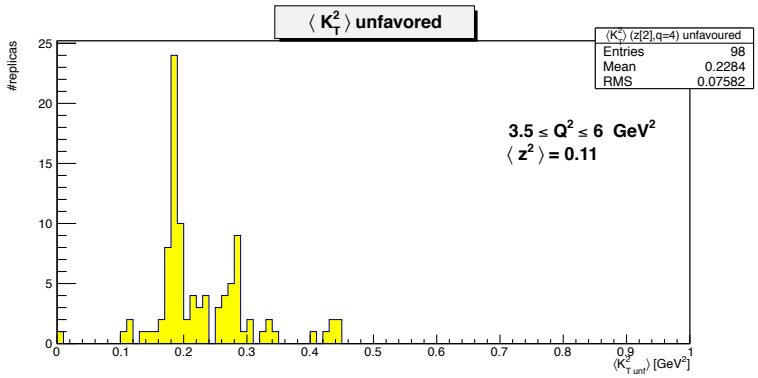

Figure 4.50: Histogram collecting the values of $\langle K^2_{T,unf}\rangle(z, Q^2)$, with $3.5 < Q^2 < 6$ GeV^2 and $\langle z^2\rangle = 0.11$.

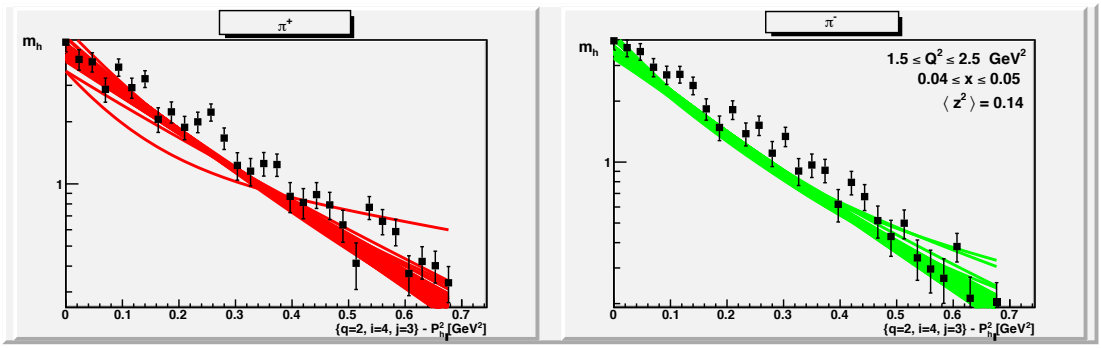

Figure 4.51: Fit of COMPASS pseudo-data (π^+ and π^-) with flavour dependent multiplicities $m_h(P_{h\perp}^2)$ for $1.5 < Q^2 < 2.5$ GeV^2, $0.04 < x < 0.05$, $\langle z^2 \rangle = 0.14$.

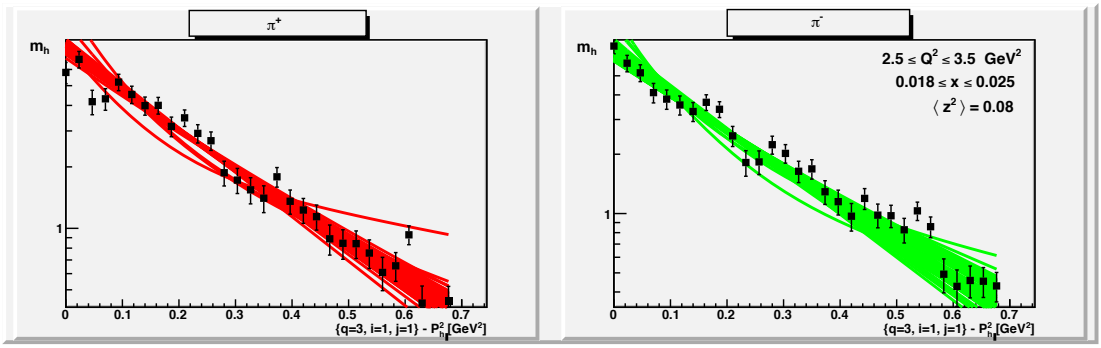

Figure 4.52: Fit of COMPASS pseudo-data (π^+ and π^-) with flavour dependent multiplicities $m_h(P_{h\perp}^2)$ for $2.5 < Q^2 < 3.5$ GeV^2, $0.018 < x < 0.025$, $\langle z^2 \rangle = 0.08$.

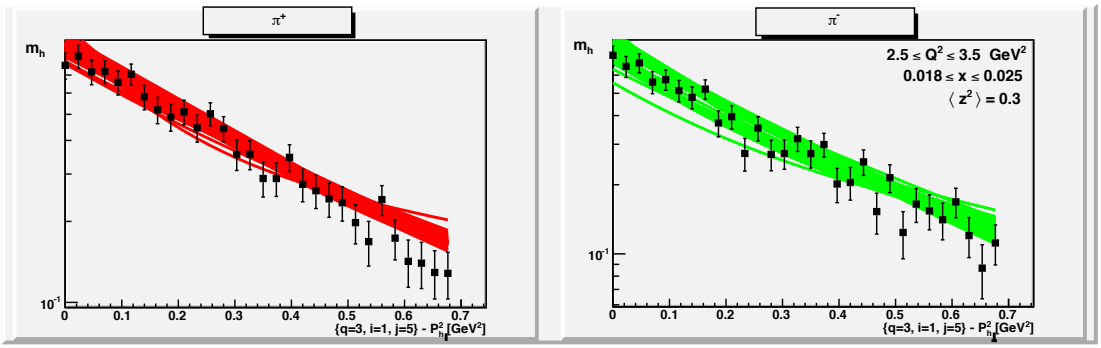

Figure 4.53: Fit of COMPASS pseudo-data (π^+ and π^-) with flavour dependent multiplicities $m_h(P_{h\perp}^2)$ for $2.5 < Q^2 < 3.5$ GeV^2, $0.018 < x < 0.025$, $\langle z^2 \rangle = 0.30$.

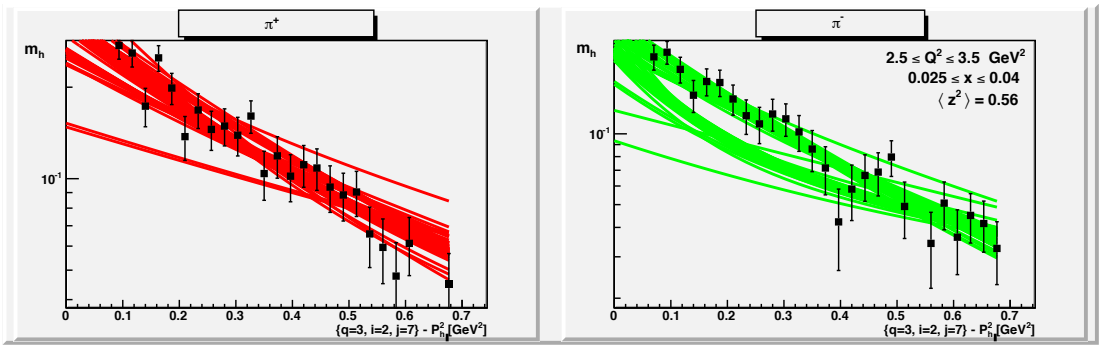

Figure 4.54: Fit of COMPASS pseudo-data (π^+ and π^-) with flavour dependent multiplicities $m_h(P_{h\perp}^2)$ for $2.5 < Q^2 < 3.5$ GeV^2, $0.025 < x < 0.04$, $\langle z^2 \rangle = 0.56$.

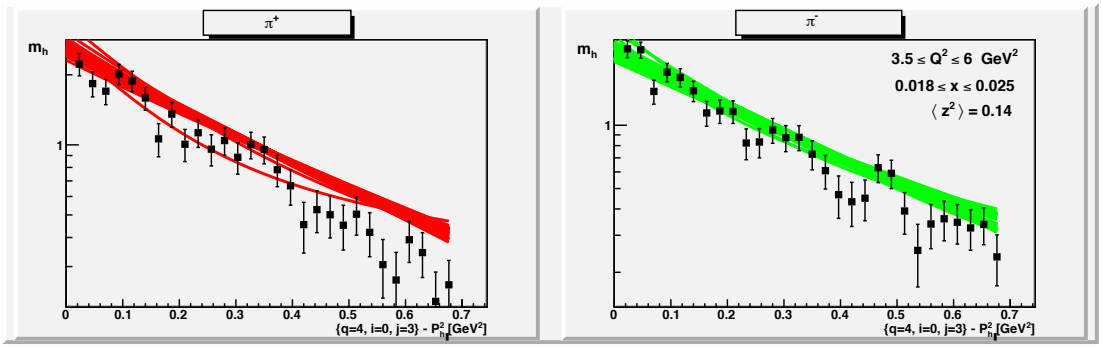

Figure 4.55: Fit of COMPASS pseudo-data (π^+ and π^-) with flavour dependent multiplicities $m_h(P_{h\perp}^2)$ for $3.5 < Q^2 < 6$ GeV^2, $0.018 < x < 0.025$, $\langle z^2 \rangle = 0.14$.

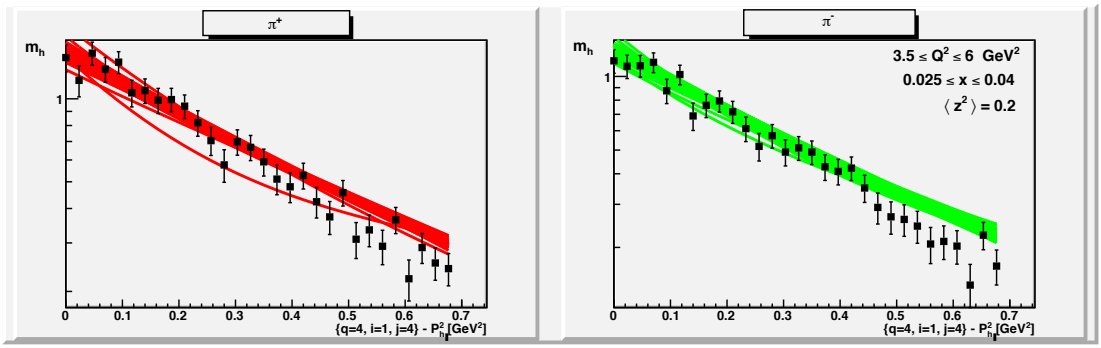

Figure 4.56: Fit of COMPASS pseudo-data (π^+ and π^-) with flavour dependent multiplicities $m_h(P_{h\perp}^2)$ for $3.5 < Q^2 < 6$ GeV^2, $0.025 < x < 0.04$, $\langle z^2 \rangle = 0.20$.

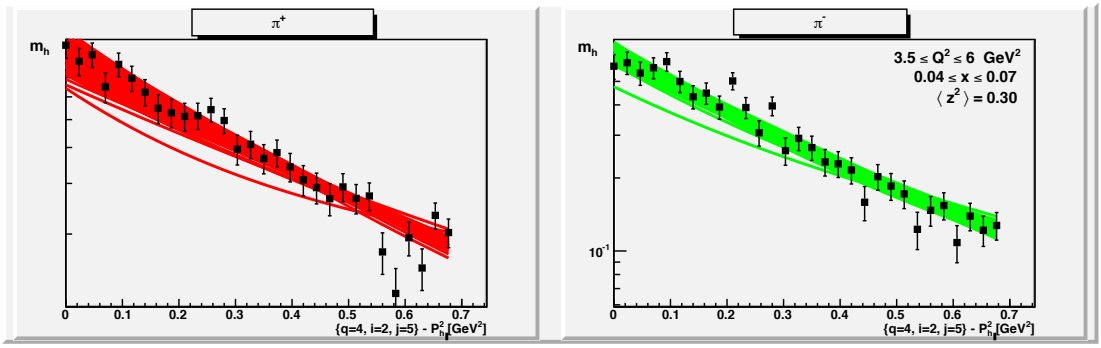

Figure 4.57: Fit of COMPASS pseudo-data (π^+ and π^-) with flavour dependent multiplicities $m_h(P_{h\perp}^2)$ for $3.5 < Q^2 < 6$ GeV^2, $0.04 < x < 0.07$, $\langle z^2 \rangle = 0.30$.

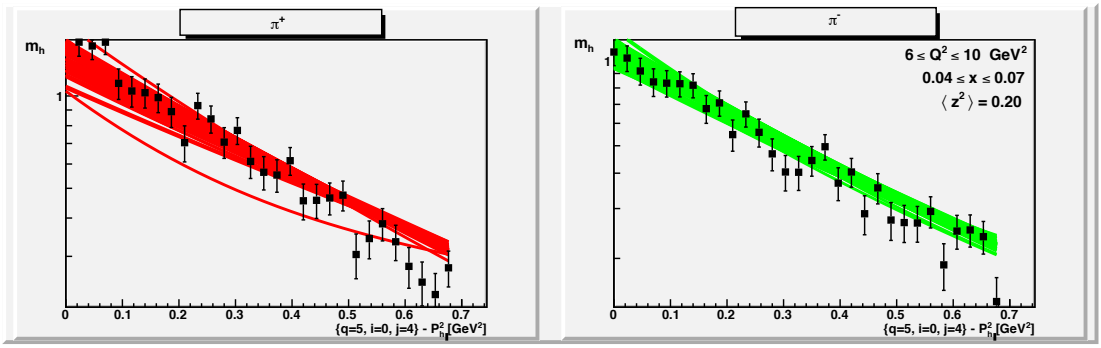

Figure 4.58: Fit of COMPASS pseudo-data (π^+ and π^-) with flavour dependent multiplicities $m_h(P_{h\perp}^2)$ for $6 < Q^2 < 10$ GeV^2, $0.04 < x < 0.07$, $\langle z^2 \rangle = 0.20$.

4.4.2 Flavour independence test

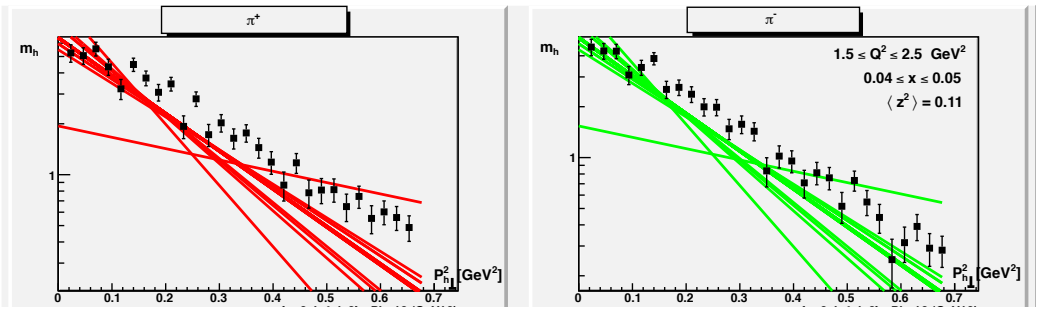

Figure 4.59: Fit of COMPASS pseudo-data (π^+ and π^-) with flavour independent multiplicities $m_h(P_{h\perp}^2)$ for $1.5 < Q^2 < 2.5$ GeV^2, $0.04 < x < 0.05$, $\langle z^2 \rangle = 0.11$.

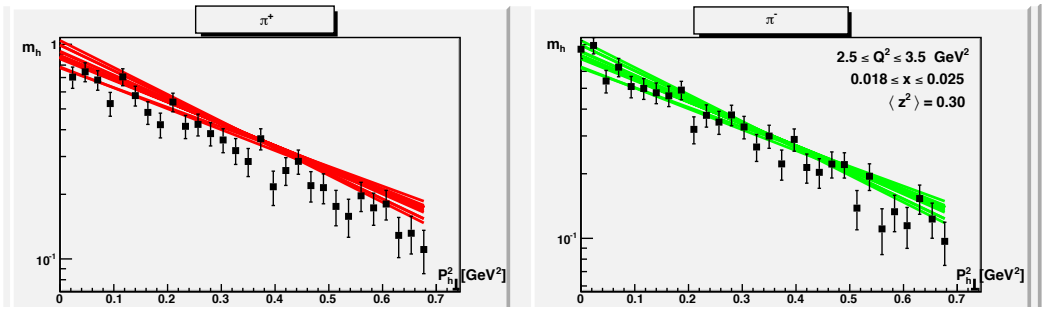

Figure 4.60: Fit of COMPASS pseudo-data (π^+ and π^-) with flavour independent multiplicities $m_h(P_{h\perp}^2)$ for $2.5 < Q^2 < 3.5$ GeV^2, $0.018 < x < 0.025$, $\langle z^2 \rangle = 0.30$.

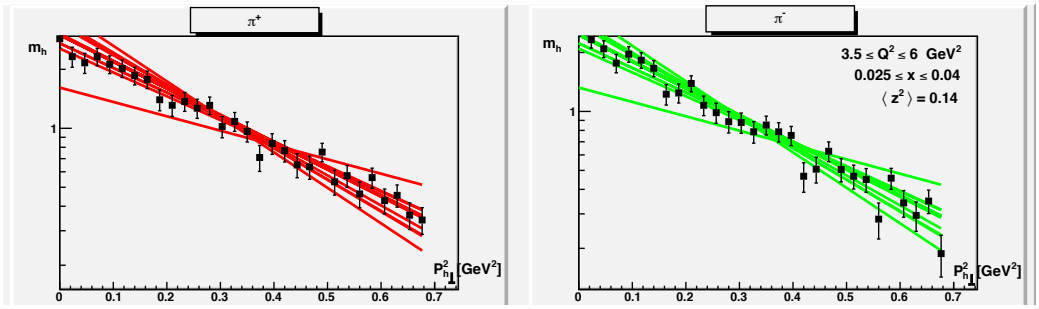

Figure 4.61: Fit of COMPASS pseudo-data (π^+ and π^-) with flavour independent multiplicities $m_h(P_{h\perp}^2)$ for $3.5 < Q^2 < 6$ GeV^2, $0.025 < x < 0.04$, $\langle z^2 \rangle = 0.14$.

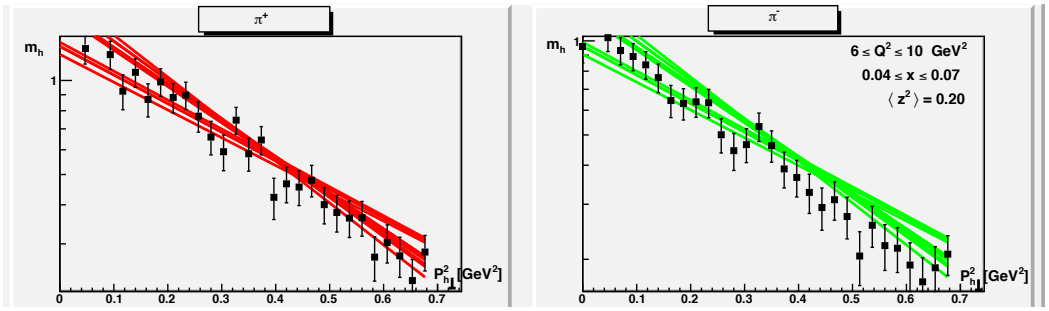

Figure 4.62: Fit of COMPASS pseudo-data (π^+ and π^-) with flavour independent multiplicities $m_h(P_{h\perp}^2)$ for $6 < Q^2 < 10$ GeV^2, $0.04 < x < 0.07$, $\langle z^2 \rangle = 0.20$.

Chapter 5

Conclusions

In this thesis we performed an explorative study of the flavour dependence of unpolarized transverse-momentum-dependent distribution functions and fragmentation functions. Our phenomenological analysis starts from Semi-Inclusive Deep-Inelastic Scattering preliminary data collected by the COMPASS experiment, at CERN. These sets of data are not available, but Jean-Francois Rajotte, a member of the COMPASS collaboration, in his Ph.D. thesis [30] calculated Gaussian fits of hadron multiplicities as a function of the transverse momentum of the detected hadron, $\boldsymbol{P}_{h\perp}$. Starting from these fits we generated $\mathcal{M}$ sets of pseudo-experimental data adding Gaussian noise to the fits of multiplicities. The procedure of fitting pseudo-data with flavour dependent multiplicities was repeated for each of the $\mathcal{M}$ sets of pseudo-data, deducing histograms for values of best-fit parameters.

We performed a combined fit over 84 kinematic bins, both for π^+ and π^-, replicated $\mathcal{M} = 100$ times. The explored 3-dimensional region is:

$$\begin{aligned} &1.5 < Q^2 < 10 \text{ GeV}^2 \\ &0.0045 < x_B < 0.12 \\ &0.08 < \langle z^2 \rangle < 0.56 \end{aligned} \tag{5.1}$$

Results indicate that models with flavour dependent unpolarized TMDs fit data *significantly better* than model without flavour dependence, since χ^2/dof values are smaller (see figs. 5.1 and 5.2). For each considered bin we calculated histograms for the mean square transverse momentum of up, down and sea quarks ($\langle p^2_{T,up} \rangle$, $\langle p^2_{T,down} \rangle$, $\langle p^2_{T,sea} \rangle$) and for the mean square transverse momentum of the fragmenting hadron in favored and unfavored channels ($\langle K^2_{T,fav} \rangle$ and $\langle K^2_{T,unf} \rangle$). For illustrative purposes we quote here results related to the average-Q^2 bin ($3.5 < Q^2 < 6$ GeV^2), as an example of the general behaviour:

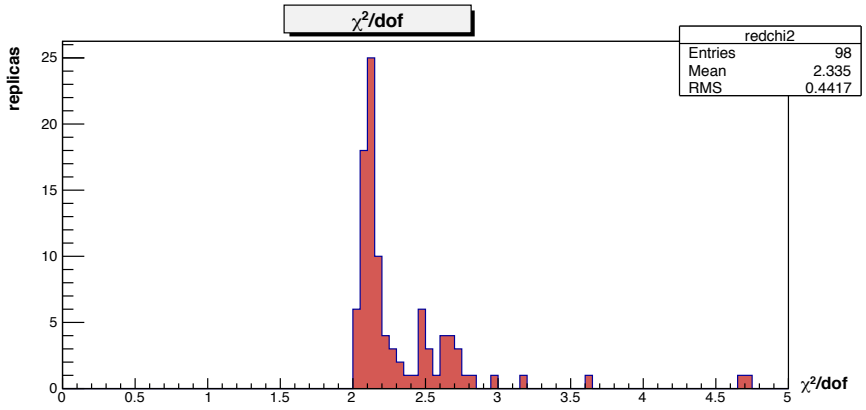

Figure 5.1: χ^2/dof values for the flavour *dependent* fit.

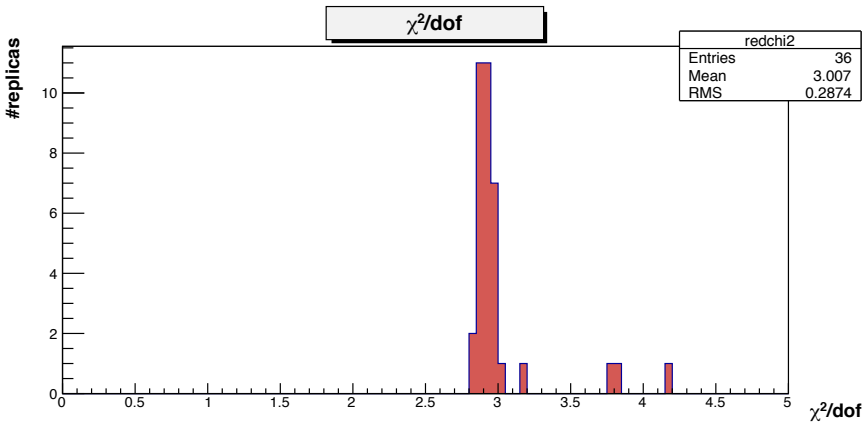

Figure 5.2: χ^2/dof values for the flavour *independent* fit.

x_B	$\langle p^2_{T,up}\rangle$ [GeV^2]	$\langle p^2_{T,down}\rangle$ [GeV^2]	$\langle p^2_{T,sea}\rangle$ [GeV^2]
$0.018 - 0.025$	1.09 ± 1.28	0.32 ± 1.04	1.46 ± 1.09
$0.025 - 0.04$	0.94 ± 0.99	0.81 ± 1.59	1.21 ± 0.90
$0.07 - 0.12$	0.45 ± 0.29	0.36 ± 0.50	0.40 ± 0.38

$\langle z^2\rangle$	$\langle K^2_{T,fav}\rangle$ [GeV^2]	$\langle K^2_{T,unf}\rangle$ [GeV^2]
0.08	0.19 ± 0.08	0.21 ± 0.05
0.14	0.24 ± 0.07	0.23 ± 0.10
0.56	0.20 ± 0.15	0.14 ± 0.15

In this Q^2 bin (but also in all the other explored regions), the *average* mean square transverse momentum of the up quark is found to be bigger than the down one. However, average values of $\langle p_T^2\rangle$ are well compatible within 1σ deviations. Sea quarks are characterized by average mean square transverse momenta bigger than (or at least comparable to) the ones related to valence quarks. For low values of z the *average* mean square transverse momentum for unfavored fragmentations is bigger than the favored one, but the situation is reversed at higher values of z. Average values, as in the case of $\langle p_T^2\rangle$, are well compatible within 1σ deviation.

In the following table, we collect correlation indeces among histograms for mean square transverse momenta, calculated with $2.5 < Q^2 < 3.5$ GeV^2, $0.025 < x < 0.04$ and $\langle z^2\rangle = 0.14$:

ρ	$\langle p^2_{T,up}\rangle$	$\langle p^2_{T,down}\rangle$	$\langle p^2_{T,sea}\rangle$	$\langle K^2_{T,fav}\rangle$	$\langle K^2_{T,unf}\rangle$
$\langle p^2_{T,up}\rangle$	1	0.15	-0.36	-0.03	0.43
$\langle p^2_{T,down}\rangle$	0.15	1	-0.20	-0.05	0.06
$\langle p^2_{T,sea}\rangle$	-0.36	-0.20	1	-0.05	-0.71
$\langle K^2_{T,fav}\rangle$	-0.03	-0.05	-0.05	1	0.02
$\langle K^2_{T,unf}\rangle$	0.43	0.06	-0.71	0.02	1

Some replicas in the set display results compatible with previous studies (down quark "larger" than the up one) but the *average* behaviour is different. Our analysis is statistically more complete than previous ones: replicating $\mathcal{M}$ sets of data we can appreciate the possible best-fit values corresponding to different choices of the Gaussian noise, namely to hypothetical different replicas of the experimental procedure.

From the test of flavour dependence we can observe that the the hadron multiplicity m_h is no more a Gaussian function in the transverse momentum of the detected hadron $\boldsymbol{P}_{h\perp}$. In fact, the structure function $F_{UU,T}$ is built as a *sum* of convolutions of Gaussian TMD PDFs and FFs, resulting in a non-Gaussian function.

Numerical results related to χ^2/dof and to transverse momenta are b and λ-dependent. Considering values for b between 2 and 3 and values for λ between 0.1 and 0.2, however, the relationship between average transverse momenta is always the same (up wider than down).

After this analysis, the flavour dependence of unpolarized TMDs is no more an heuristic hypothesis only: there are also phenomenological hints pointing out its reliability and goodness.

5.1 Comparison with results from Jefferson Lab

Results on the flavour dependence of unpolarized TMDs collected at Jefferson Lab [3] are obtained in the following kinematic region:

$$\begin{aligned} &2 < Q^2 < 4 \text{ GeV}^2 \\ &0.2 < x_B < 0.5 \\ &0.3 < z < 1 \\ &P_{h\perp}^2 < 0.2 \text{ GeV}^2 \ . \end{aligned} \tag{5.2}$$

The fit is performed disregarding sea quarks, since $x_B > 0.2$. The calculated best-fit values (with their error) for up, down, favored and unfavored mean square transverse momenta are:

$$\begin{aligned} &\langle p_{T,up}^2 \rangle = -0.01 \pm 0.04 \text{ GeV}^2 \\ &\langle p_{T,down}^2 \rangle = 0.22 \pm 0.13 \text{ GeV}^2 \\ &\\ &\langle K_{T,fav}^2 \rangle = 0.23 \pm 0.04 \text{ GeV}^2 \\ &\langle K_{T,unf}^2 \rangle = 0.19 \pm 0.04 \text{ GeV}^2 \ . \end{aligned} \tag{5.3}$$

Even if a rigorous comparison is not allowed (since the x_B range explored at Jefferson Lab does not overlap with the COMPASS one), we can make a comparison with our analysis in similar Q^2 regions (e.g., $2.5 < Q^2 < 3.5$ GeV^2). As for $\langle p_T^2 \rangle$, results are in disagreement: at Jefferson Lab the average value for the down quark is found to be bigger than the up one and non-compatible within 1σ. Concerning $\langle K_T^2 \rangle$ values, instead, results are compatible with ours, for high z values. In our analysis statistical errors are bigger than JLab's ones, but this is due to our statistical method (replicas). In the following table we list only an approximate estimate of the correlation indices among the various squared mean transverse momenta:

ρ	$\langle p_{T,up}^2 \rangle$	$\langle p_{T,down}^2 \rangle$	$\langle K_{T,fav}^2 \rangle$	$\langle K_{T,unf}^2 \rangle$
$\langle p_{T,up}^2 \rangle$	1	~ 0	< 0	$-$
$\langle p_{T,down}^2 \rangle$	~ 0	1	$-$	< 0
$\langle K_{T,fav}^2 \rangle$	< 0	$-$	1	> 0
$\langle K_{T,unf}^2 \rangle$	$-$	< 0	> 0	1

where "$-$" means that the value is not estimated. The agreement with our analysis is only weak.

5.2 Comparison with results from Lattice QCD

In [4] properties of unpolarized and polarized TMDs are calculated using lattice QCD. The transverse momentum dependence of unpolarized TMDs is parametrized as a Gaussian function. The estimated numerical values for the mean transverse momentum (integrated over x, but without Q^2-dependence [1]) of up and down quarks are:

$$\begin{aligned} \langle p_{T,up} \rangle &= 0.394 \pm 0.004 \pm 0.027 \text{ GeV} \\ \langle p_{T,down} \rangle &= 0.405 \pm 0.005 \pm 0.027 \text{ GeV} . \end{aligned} \tag{5.4}$$

The first error is statistical. The second one includes an estimate related to the discretization of space-time. Again, this behaviour is opposite to our findings.

5.3 Final comments

A few comments are in order about the disagreement between the results of our analysis and of previous investigations:

1. COMPASS data are still preliminary: results in the final release may be different from the present ones

2. we analyzed fits of data, not data themselves.

We are operating with "approximations of approximated data". Only when SIDIS data will be finally released we will be able to carry out a more reliable analysis and possibly make a confident statement on the flavour dependence of TMDs. Also cross-checks will be of fundamental importance: we will

[1] The connection to the experimental Q^2 scale is not yet well established (see [4]).

analyze SIDIS data from the HERMES collaboration and Jefferson Lab's experiments.

The HERMES measurement, for example, has some remarkable features which are very interesting for our analysis:

- both protons and deuterons are available as target hadrons: this feature allows us to have a clearer access to the flavour structure of unpolarized TMDs. Investigating SIDIS off deuterons will increase the sensitivity to down quarks
- particle identification systems included in the detector will give us information also on the fragmentations of kaons

These features will affect all the best-fit parameters defined in our procedure.

The study of the flavour dependence of *unpolarized* TMDs could have an impact also on the determination of *polarized* transverse-momentum-dependent distribution functions and fragmentation functions. These ones, in fact, are extracted from asymmetries, i.e. ratios of cross-sections involving unpolarized TMDs too. The impact of unpolarized TMDs, however, is not confined to the field of hadronic physics: every high-energy physics experiment involving hadrons which is sensitive to transverse momentum dependence would be affected by the determination of the flavour dependence of unpolarized TMDs.

Acknowledgments

"I care"
don L. Milani

I would like to thank my supervisor Alessandro Bacchetta for all the help, the patience and the care that he reserved to me during these months. Thanks also to all the other members of the "Hadronic Structure and QCD" group of the University of Pavia. Barbara Pasquini, Marco Guagnelli, Marco Radici, Sigfrido Boffi: it was a pleasure to work in your research group. I enjoyed a both hard-working and friendly setting in the Department of Physics at University of Pavia: Oreste Nicrosini, Guido Montagna and Fulvio Piccinini, thank you for your confidence and expertise. I would like to thank also all the HERMES collaboration and, most of all, Gunar Schnell and Aram Movsisyan: your help during the DESY Summer School was fundamental in order to complete this project. Thanks to Marco Cagnotti, with whom I discovered the importance of scientific writing and all the care needed in order to make science going from academias to the society.

A special mention to my family: thank you for supporting me during these years. Without your help, I would not be what I am. Thanks also to my friends from Cremona and Pavia: I am aware of your importance, caring me as a friend (and as a physicist). Last but not least, thanks to Summer-School-mates met in Hamburg: I spent a terrific summer with you. It is the proof of how enriching and funny is meeting people from all over the world.

Bibliography

[1] A. Bacchetta and M. Contalbrigo, "The proton in 3D," *Il Nuovo Saggiatore* **28** (2012) 16–27.

[2] F. Conti, A. Bacchetta, and M. Radici, "Transverse-momentum-dependent parton distributions in a diquark spectator model," *Phys.Rev.* **D78** (2008) 074010.

[3] H. Mkrtchyan, P. Bosted, G. Adams, A. Ahmidouch, T. Angelescu, *et al.*, "Transverse momentum dependence of semi-inclusive pion production," *Phys.Lett.* **B665** (2008) 20–25, `arXiv:0709.3020 [hep-ph]`.

[4] B. U. Musch, P. Hagler, J. W. Negele, and A. Schafer, "Exploring quark transverse momentum distributions with lattice QCD," *Phys.Rev.* **D83** (2011) 094507, `arXiv:1011.1213 [hep-lat]`.

[5] W. J. den Dunnen, D. Boer, C. Pisano, M. Schlegel, and W. Vogelsang, "Linearly polarized Gluons and the Higgs Transverse Momentum Distribution," `arXiv:1205.6931 [hep-ph]`.

[6] M. Peskin and D. Schroeder, *An introduction to Quantum Field Theory*. Westview Press, 1995.

[7] J. Collins, *Foundations of perturbative QCD*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, 2011.

[8] **CTEQ** Collaboration, R. Brock *et al.*, "Handbook of perturbative QCD: Version 1.0," *Rev.Mod.Phys.* **67** (1995) 157–248.

[9] P. J. Mulders, *Transverse momentum dependence in structure functions in hard scattering processes*. 2007. Lecture Notes.

[10] A. Bacchetta, *Transverse momentum distributions*. 2012. Lecture Notes.

[11] A. Bacchetta, M. Diehl, K. Goeke, A. Metz, P. J. Mulders, *et al.*, "Semi-inclusive deep inelastic scattering at small transverse momentum," *JHEP* **0702** (2007) 093, arXiv:hep-ph/0611265.

[12] I. Akushevich, N. Shumeiko, and A. Soroko, "Radiative effects in the processes of hadron electroproduction," *Eur.Phys.J.* **C10** (1999) 681–687, arXiv:hep-ph/9903325.

[13] M. Schlegel and A. Metz, "Two-Photon Exchange in (Semi-)Inclusive DIS," *AIP Conf.Proc.* **1149** (2009) 543–546, arXiv:0902.0781 [hep-ph].

[14] J. P. Ralston and D. E. Soper, "Production of Dimuons from High-Energy Polarized Proton Proton Collisions," *Nucl.Phys.* **B152** (1979) 109.

[15] P. Mulders and R. Tangerman, "The Complete tree level result up to order 1/Q for polarized deep inelastic leptoproduction," *Nucl.Phys.* **B461** (1996) 197–237, arXiv:hep-ph/9510301.

[16] K. Goeke, A. Metz, and M. Schlegel, "Parameterization of the quark-quark correlator of a spin-1/2 hadron," *Phys.Lett.* **B618** (2005) 90–96, arXiv:hep-ph/0504130.

[17] R. L. Jaffe, "Spin, twist and hadron structure in deep inelastic processes," arXiv:hep-ph/9602236.

[18] R. Jaffe, "Can transversity be measured?," arXiv:hep-ph/9710465.

[19] A. Bacchetta, U. D'Alesio, M. Diehl, and C. A. Miller, "Single-spin asymmetries: The Trento conventions," *Phys.Rev.* **D70** (2004) 117504, arXiv:hep-ph/0410050.

[20] J. C. Collins, "Light cone variables, rapidity and all that," arXiv:hep-ph/9705393.

[21] D. Boer and P. Mulders, "Time reversal odd distribution functions in leptoproduction," *Phys.Rev.* **D57** (1998) 5780–5786, arXiv:hep-ph/9711485.

[22] J. C. Collins, "Fragmentation of transversely polarized quarks probed in transverse momentum distributions," *Nucl.Phys.* **B396** (1993) 161–182, arXiv:hep-ph/9208213.

[23] J. B. Kogut and D. E. Soper, "Quantum Electrodynamics in the Infinite Momentum Frame," *Phys.Rev.* **D1** (1970) 2901–2913.

[24] S. M. Aybat and T. C. Rogers, "TMD Parton Distribution and Fragmentation Functions with QCD Evolution," *Phys.Rev.* **D83** (2011) 114042, `arXiv:1101.5057 [hep-ph]`.

[25] **COMPASS** Collaboration, G. Baum *et al.*, "COMPASS: A Proposal for a Common Muon and Proton Apparatus for Structure and Spectroscopy,".

[26] X.-d. Ji, J.-P. Ma, and F. Yuan, "QCD factorization for spin-dependent cross sections in DIS and Drell-Yan processes at low transverse momentum," *Phys.Lett.* **B597** (2004) 299–308, `arXiv:hep-ph/0405085`.

[27] F. Landry, R. Brock, P. M. Nadolsky, and C. Yuan, "Tevatron Run-1 Z boson data and Collins-Soper-Sterman resummation formalism," *Phys.Rev.* **D67** (2003) 073016, `arXiv:hep-ph/0212159`.

[28] C. P. Hays, B. Jayatilaka, A. Kotwal, L. Nodulman, O. Stelzer-Chilton, *et al.*, "First measurement of the W boson mass with CDF in run II," *J.Phys.Conf.Ser.* **110** (2008) 122019.

[29] U. D'Alesio and F. Murgia, "Parton intrinsic motion in inclusive particle production: Unpolarized cross sections, single spin asymmetries and the Sivers effect," *Phys.Rev.* **D70** (2004) 074009, `arXiv:hep-ph/0408092`.

[30] J. F. Rajotte, "Hadron muoproduction at the COMPASS experiment. PhD Thesis (Munchen, 2010), unpublished,".

[31] M. Whalley, D. Bourilkov, and R. Group, "The Les Houches accord PDFs (LHAPDF) and LHAGLUE," `arXiv:hep-ph/0508110`.

[32] D. de Florian, R. Sassot, and M. Stratmann, "Global analysis of fragmentation functions for pions and kaons and their uncertainties," *Phys.Rev.* **D75** (2007) 114010, `arXiv:hep-ph/0703242`.